Didáctica de la Estadística

Book · January 2001

DOI: 10.13140/RG.2.1.3946.7044

1 author:

Carmen Batanero
University of Granada

Didáctica de la Estadística

Carmen Batanero

Departamento de Didáctica de la Matemática
Universidad de Granada

Didáctica de la Estadística

Carmen Batanero

G rupo de
E ducación
E stadística
U niversidad de
G ranada

Departamento de Didáctica de la Matemática
Universidad de Granada

Depósito Legal: GR-150/ 2001

ISBN: 84-699-4295-6

Publica:
Grupo de Investigación en Educación Estadística
Departamento de Didáctica de la Matemática
Universidad de Granada

Imprime:
Servicio de Reprografía de la Facultad de Ciencias
Universidad de Granada
Avda. Fuentenueva s/n 18071 Granada

Financiación:
Proyecto BSO2000-1507, DGES, Ministerio de Educación y Ciencia.
Grupo de Investigación FQM-126. Consejería de Educación. Junta de Andalucía.

ÍNDICE

INTRODUCCIÓN

Este libro está pensado para mis alumnos del curso "Didáctica de la Estadística", que se ha venido ofreciendo como asignatura optativa dentro del Plan de Estudios de la Licenciatura en Ciencias y Técnicas Estadísticas de la Universidad de Granada, desde el comienzo de dicho plan de estudios. Incluye también parte del material que a lo largo de doce años he venido elaborando, como parte de mi trabajo en diversos grupos de investigación en educación estadística. Finalmente, el libro trata de reflejar las diversas facetas de la estadística: como ciencia, como herramienta de investigación en áreas diversas, como campo de investigación didáctica, tanto para la formación de niños, como de profesionales, investigadores y profesores.

Al tratar de reflexionar sobre la formación didáctica que sería necesario impartir a estadísticos o a matemáticos, quienes tienen una formación suficientemente sólida y actualizadas en los métodos y técnicas de esta materia, surge la necesidad de concretar lo que, en el lenguaje didáctico, conocemos como conocimiento del contenido didáctico. ¿Cuál es este conocimiento, una vez que se dominan las técnicas matemáticas? ¿Cómo resumirlo y hacerlo útil e interesante para los futuros profesores? ¿Qué tipo de situaciones didácticas podemos usar para la enseñanza del contenido didáctico, si queremos ser consecuentes con los principios constructivistas del aprendizaje, con la importancia de la interacción social y del trabajo en grupo del alumno?

El material que presenta este libro trata de dar respuesta a estas preguntas en un área particular de las matemáticas -la estadística- en la que existen pocos - o ningún- precedente de textos dedicados a la formación didáctica de los profesores. Está dirigido a los profesores de educación secundaria, tanto estadísticos, como matemáticos, quienes suponemos tienen ya una base sólida de los conceptos estadísticos elementales. Podría ser también útil a los profesores de estadística en los cursos introductorios en la universidad en áreas como humanidades, ciencias sociales, ciencias de la actividad física y deporte, educación y psicología, y otras especialidades en las que los alumnos no siempre han cursado un Bachillerato científico.

En primer lugar, presentamos una breve panorámica del estado actual de la educación estadística, que proviene de diversas áreas de conocimiento y no sólo de la educación matemática, y que en los últimos años, con la creación de la *IASE* (Sociedad Internacional de Educación Estadística) empieza a configurarse como una disciplina con entidad propia.

El segundo capítulo lo dedicamos a la reflexión epistemológica. Mientras que los conceptos estadísticos son sencillos, desde un punto de vista matemático, existen numerosas dificultades de tipo filosófico ligadas a la interpretación de estos conceptos y su aplicación a situaciones prácticas. El profesor debe ser consciente de la pluralidad de significados atribuibles a conceptos como el de aleatoriedad o probabilidad y de las controversias existentes en torno a la inferencia estadística, puesto que las dificultades epistemológicas se reproducen con frecuencia en el aprendizaje de los alumnos.

La investigación sobre comprensión y aprendizaje de los conceptos estadísticos elementales está aún poco sistematizada y proviene de diversas áreas de conocimiento que han usado diferentes paradigmas de investigación y marcos teórico. En el capítulo 3 hemos tratado de hacer una síntesis personal de los resultados más estables y que son aceptados con generalidad por los investigadores en educación estadística.

En el capítulo 4 presentamos diversos elementos a tener en cuenta en el diseño curricular, incluyendo las disposiciones curriculares, reflexión sobre los fines de la educación estadística, sus campos de aplicación, el papel y métodos de evaluación y los materiales didácticos, entre los cuales la tecnología ocupa un papel importante.

En estos cuatro capítulos la información presentada se intercala con un conjunto de actividades que he venido utilizando en cursos de didáctica de la estadística: textos breves o puntos para discutir, análisis de ítems usados en la evaluación del razonamiento estocástico, problemas paradójicos o con solución contraintuitiva, búsqueda, recogida o análisis de datos, análisis de tareas o procedimientos, juegos, etc.

Estas actividades están pensadas como medio para contextualizar y provocar la reflexión y discusión sobre los diferentes conocimientos presentados, -sean epistemológicos, psicológicos, o didácticos- en cursos destinados a la formación didáctica de los profesores. Muchas de estas actividades pueden ser también usadas en las clases de estadística de secundaria, o en la universidad para hacer reflexionar a los alumnos sobre sus intuiciones incorrectas, para plantearles problemas originales y aumentar su motivación.

Finalmente, en el último capítulo presentamos algunos ejemplos de proyectos desarrollados para trabajar en la clase de estadística, junto con sugerencias de temas para otros proyectos. La finalidad es contextualizar la enseñanza de la estadística, dentro del proceso más general de investigación y presentar un desarrollo de un curso de estadística para secundaria o primeros cursos de universidad apoyado en el trabajo con proyectos.

1

SITUACIÓN ACTUAL Y PERSPECTIVAS FUTURAS DE LA DIDÁCTICA DE LA ESTADÍSTICA

1.1. Introducción

Aunque hace unos años pocos investigadores se interesaban por los problemas de la enseñanza y aprendizaje de la estadística, en la actualidad asistimos a un aumento notable de las publicaciones, diseños curriculares e investigación relacionados con este tema. El propósito de este primer capítulo es hacer un breve resumen histórico del origen de la educación estadística y reflexionar sobre la situación actual y perspectivas futuras.

Recientemente la estadística se ha incorporado, en forma generalizada al currículo de matemáticas de la enseñanza primaria y secundaria y de las diferentes especialidades universitarias en la mayoría de países desarrollados.

Ello ha impulsado la investigación y el desarrollo curricular en el campo especifico de la estadística. Ejemplos de proyectos curriculares desarrollados de acuerdo a estas ideas son, por ejemplo, los del Schools Council Project on Statistical Education en el Reino Unido (1957-1981) y el Quantitative Literacy Project (1985-98) y Data Driven Mathematics (1996-2000) en Estados Unidos. Los materiales didácticos, el software educativo, investigaciones, revistas, reuniones y congresos sobre la enseñanza de la estadística han crecido espectacularmente en los últimos años.

Este interés, sin embargo, no es exclusivo de la comunidad de educación matemática. La preocupación por las cuestiones didácticas y por la formación de profesionales y usuarios de la estadística ha sido una constante de los propios estadísticos, y las investigaciones sobre el razonamiento estocástico han tenido un gran auge en el campo de la psicología. En lo que sigue analizamos los trabajos sobre educación estadística llevados a cabo en estos tres campos.

1.2. Didáctica de la estadística dentro de la estadística

La relación entre el desarrollo de un país y el grado en que su sistema estadístico produce estadísticas completas y fiables es clara, porque esta información es necesaria para la toma de decisiones acertadas de tipo económico, social y político. La educación estadística, no sólo de los técnicos que producen estas estadísticas, sino de los profesionales y ciudadanos que deben interpretarlas y tomar a su vez decisiones basadas en esta información, así como de los que deben colaborar en la obtención de los datos requeridos es, por tanto, un motor del desarrollo.

La educación estadística ha asido una preocupación crucial *del Instituto Internacional de Estadística* (*ISI*) desde su fundación en 1885, y esta preocupación se concretó oficialmente en 1948 en el establecimiento del *Comité de Educación*, encargado de promover la formación estadística, colaborando, para este fin, con la *UNESCO* y otros organismos internacionales, en un momento histórico en que era prioritario mejorar la información estadística en los países en vías de desarrollo, lo que implicaba la necesidad de preparar suficiente numero de técnicos estadísticos en estos países. Las responsabilidades del Comité de Educación incluyeron el desarrollo de diplomaturas y licenciaturas en estadística en los que se formarían los profesores y técnicos estadísticos. Una de las primeras actividades de este comité fue la creación de los *Centros de Internacionales de Educación Estadística (ISEC)* en Calcuta y Beirut, para atender las necesidades formativas de los países de su respectivo entorno geográfico. Este año se han cumplido 50 años de funcionamiento ininterrumpido del *ISEC* de Calcuta que estuvo dirigido por estadísticos tan prestigiosos como Mahalanobis y Rao.

Asimismo, el comité ha colaborado en la producción y difusión de ayudas para la enseñanza, por ejemplo, la preparación de libros de texto universitarios, bibliografías específicas y diccionarios de términos estadísticos. Subcomités especiales se dedicaron a impulsar la introducción de la estadística en las escuelas, el papel de la mujer en la estadística, y la promoción de conferencias sobre la educación estadística, dando origen, en particular a los *ICOTS* (*International Conference on Statistical Education*) que se iniciaron en 1982 en la Universidad de Sheffield y han continuado cada cuatro años.

Otro tipo de conferencias iniciadas por el comité de educación, como satélites del *ICME* (*International Congress of Mathematics Education*), son *las Round Table Conferences* sobre temas especifico de educación estadística, que han sido los siguientes: "Estadística en la escuela" (Viena, 1973; Varsovia, 1975, Calcuta, 1977), "La enseñanza universitaria de la estadística en los países en vías de desarrollo (La Haya, 1968), "Enseñanza de la estadística y ordenadores", (Oisterwijk, 1970; Camberra, 1984), y "Formación de profesores" (Budapest, 1988).

En 1991 el ISI decide crear una nueva sección, a la que se transferirían las responsabilidades y objetivos que hasta entonces había tenido el Comité de Educación. Nace así *IASE (International Association for Statistical Education),* con igualdad de derechos y obligaciones que el resto de las secciones del Instituto, participando en la elaboración de sus revistas y organización de sus Sesiones bianuales, contribuyendo a su financiación y teniendo representación en sus organismos directivos. El objetivo principal de *IASE* es el desarrollo y mejora de la educación estadística en el ámbito internacional. Sus miembros son personas interesadas en la enseñanza de la estadística en cualquiera de los niveles educativos, el desarrollo de software estadístico, la enseñanza de la estadística en empresas o industria, preparación de expertos estadísticos para las unidades estadísticas en el gobierno y el desarrollo curricular, libros de texto y materiales didáctico.

Entre las responsabilidades asumidas, se encuentran la organización del ICOTS a (ICOTS IV, Marrakesh 1994; ICOTS V, Singapur, 1998) y de las *Round Table Conference* asociadas al *ICME,* habiendo organizado hasta la fechas las dedicadas a "Enseñanza del análisis de datos" (Quebec, 1992), "Impacto de las nuevas tecnologías en la investigación" (Granada, 1996) y "Formación de los investigadores en el uso de la estadística" (Tokio, 2000).

1.3. La perspectiva psicológica: investigación sobre el razonamiento estocástico

La influencia que, en la psicología, han tenido las investigaciones sobre el razonamiento estocástico es tal que Pérez Echeverría (1990) habla de "revolución probabilista" para referirse a este impacto, equiparándolo al que ha tenido la perspectiva cognitiva. Esta importancia se debe al giro que estos estudios han implicado en los trabajos sobre razonamiento humano, donde se ha pasado de un modelo de actuación acorde a la lógica formal, a concebir un decisor que actúa de acuerdo a un sistema probabilístico complejo, utilizando heurísticas adquiridas en su relación empírica con lo cotidiano. Trabajos como los recogidos en Kahneman y cols. (1982), que tocan entre otros puntos el razonamiento correlacional, la inferencia, la probabilidad condicional y regla de Bayes, han contribuido a este cambio de paradigma en los estudios psicológicos.

Una heurística es una estrategia inconsciente que reduce la complejidad de un problema probabilístico, suprimiendo parte de la información. Aunque las heurísticas ayudan en muchos casos a obtener una solución aproximada al problema, en otros producen sesgos en las conclusiones obtenidas, con las consiguientes implicaciones en las decisiones tomadas. Consideremos, por ejemplo, el siguiente ítem adaptado de Kahneman y cols. (1982).

Ejemplo 1.1 En un hospital maternal se lleva un registro del sexo de los recién nacidos. ¿Cuál de los dos sucesos siguientes te parece que tiene más probabilidad?

A. Que entre los próximos 10 recién nacidos haya más de un 70 % de niñas.

B. Que entre los próximos 100 recién nacidos haya más de un 70 % de niñas.

C. Las dos cosas me parecen igual de probables.

La respuesta correcta a este ítem es la A, aunque la mayoría de los sujetos suelen considerar correcta la respuesta C. Esto es debido a que sólo tienen en cuenta que la proporción de niñas en las dos muestras es la misma, sin tener en cuenta el tamaño de las muestras, ya que las muestras pequeñas son más variables que las grandes. Al suprimir un dato (el tamaño de la muestra) se ha reducido la complejidad del problema, pero se ha producido una solución incorrecta. Razonamientos como el mostrado en el ejemplo (la heurística de la representatividad) han sido estudiados por los psicólogos, en diversos contextos de aplicación, tales como diagnóstico médico o juicios legales.

Por otro lado, y a partir de los estudios de Piaget e Inhelder (1951), la adquisición de las ideas de aleatoriedad y probabilidad, del razonamiento combinatorio, de la intuición de la frecuencia relativa, distribución y convergencia, así como de la capacidad de cuantificación de probabilidades ha sido analizada en los niños desde sus primeros años a la adolescencia, deteminándose, en consecuencia diferentes etapas en el desarrollo del razonamiento probabilístico. Otros autores han estudiado también la influencia de creencias previas y concepciones animistas de los niños sobre su capacidad de percepción de lo aleatorio. La importancia que estos trabajos tienen para los profesores es que permiten seleccionar de una forma racional el tipo de tareas probabilísticas que podemos proponer a nuestros alumnos en función de su edad. Los instrumentos de evaluación construidos en estas investigaciones son también útiles para valorar los conocimientos y modos de razonamientos de nuestros alumnos.

Mención particular merecen los trabajos de Fischbein (1975) y posteriores, ya que constituyeron uno de los primeros puentes de unión entre la psicología y la educación matemática. Interesado no solo por la formación de los conceptos formales, sino por la aparición de intuiciones parciales sobre los conceptos estocásticos, se preocupó también del efecto de la instrucción. Sus investigaciones apoyan decididamente la conveniencia de adelantar la educación estocástica y también muestran que, sin instrucción, es difícil que se desarrolle un razonamiento estocástico adecuado, incluso una vez que se alcanza la etapa de las operaciones formales. Fischbein ha sido uno de los fundadores del grupo PME (Psychology of Mathematics Education) que en el año 2000 ha celebrado su 24 reunión anual y es, en la actualidad el principal foro de investigadores en educación matemática. En 1994 se crea un grupo de trabajo sobre estocástica dentro de PME.

1.4. Especificidad de la estadística dentro de la didáctica de la matemática

El interés por la enseñanza de la estadística, dentro de la educación matemática, viene ligado al rápido desarrollo de la estadística como ciencia y como útil en la investigación, la técnica y la vida profesional, impulsado por la difusión de los ordenadores, el crecimiento de su potencia y rapidez de calculo y las posibilidades de comunicación.

Todo ello ha facilitado el uso de la estadística a un numero creciente de personas, provocando una gran demanda de formación básica en esta materia, que ha sido encomendada, en los niveles no universitarios, a los profesores de matemáticas. Los nuevos currículos de educación primaria y secundaria incluyen en forma generalizada recomendaciones sobre la enseñanza de la estadística. Sin embargo, en la práctica son todavía pocos los profesores que incluyen este tema y en otros casos se trata muy brevemente o en forma excesivamente formalizada. Analizaremos, a continuación, la problemática que, para muchos profesores supone la enseñanza de la estadística.

Una primera dificultad proviene de los cambios progresivos que la estadística está experimentando en nuestros días, tanto desde el punto de vista de su contenido, como del punto de vista de las demandas de formación. Estamos caminando hacia una sociedad cada vez mas informatizada y una comprensión de las técnicas básicas de análisis de datos y su interpretación adecuada son cada día mas importantes. Esto nos lleva a tener que enseñar estadística a alumnos con capacidades y actitudes variables, e incluso a los que siguen un bachillerato no científico, que no disponen de la misma base de conocimientos de cálculo que sus compañeros.

Al mismo tiempo, la estadística como ciencia, atraviesa un periodo de notable expansión, siendo cada vez mas numerosos los procedimientos disponibles, alejándose cada vez mas de la matemática pura y convirtiéndose en una "ciencia de los datos", lo que implica la dificultad de enseñar un tema en continuo cambio y crecimiento. Por ejemplo, todo profesor que ha tratado de incorporar las calculadoras gráficas o el ordenador en su clase de estadística, conoce bien el trabajo añadido que supone la continua puesta el día en el manejo de estos recursos.

Por otro lado, el numero de investigaciones sobre la didáctica de la estadística es aun muy escaso, en comparación con las existentes en otras ramas de las matemáticas. Por ello, no se conocen aun cuales son las principales dificultades de los alumnos en muchos conceptos importantes. Sería también preciso experimentar y evaluar métodos de enseñanza adaptados a la naturaleza especifica de la estadística, a la que no siempre se pueden transferir los principios generales de la enseñanza de las matemáticas. Las investigaciones existentes no son muy conocidas por los profesores, ya que falta todavía mucha labor de difusión, especialmente de trabajos realizados fuera de nuestro país. Precisamente este libro pretende contribuir a llenar este hueco.

La misma naturaleza de la estadística es muy diferente de la cultura determinista tradicional en clase de matemáticas. Un indicador de ello es que aun hoy día prosiguen las controversias filosóficas sobre la interpretación y aplicación de conceptos tan básicos como los de probabilidad, aleatoriedad, independencia o contraste de hipótesis, mientras que estas controversias no existen en álgebra o geometría. Las dimensiones políticas y éticas del uso y posible abuso de la estadística y la información estadística contribuyen, asimismo, a la especificidad del campo.

La formación específica de los profesores en este ámbito especifico es prácticamente inexistente. En España sólo muy recientemente se ha iniciado una asignatura especifica de didáctica de la estadística en la Licenciatura en Ciencias y Técnicas estadísticas de la Universidad de Granada y creemos que este tipo de asignatura es prácticamente inexistente en otras universidades o licenciaturas.

Por otro lado, aunque existen libros de texto excelentes, la investigación didáctica está comenzando a mostrar como algunos errores conceptuales y pedagogía inadecuada se transmiten con una frecuencia mayor de lo que seria deseable en los libros de texto (Sánchez-Cobo, 1996; Ortiz, 1999).

Un ultimo punto es la naturaleza interdisciplinar del tema, que hace que los conceptos estadísticos aparezcan en otras materias, como ciencias sociales, biología, geografía, etc., donde los profesores, a veces se ven obligados a enseñar estadística, lo que puede ocasionar conflictos cuando las definiciones o propiedades presentadas de los conceptos no coinciden con las impartidas en la clase de matemáticas.

Parece, en consecuencia, necesaria una mejor preparación previa y formación permanente del profesorado y un apoyo de los departamentos universitarios y grupos de investigación implicados para lograrlo. El papel de las sociedades profesionales, como *IASE* es también decisivo, especialmente a partir de la constitución de grupos locales activos que sirvan de intermediarios entre los profesores, estadísticos profesionales e investigadores en educación estadística en sus distintas vertientes.

1.5. ¿Hacia donde va la educación estadística?

En los apartados anteriores hemos descrito el panorama actual de la educación estadística. Es indiscutible que el siglo XX ha sido el siglo de la estadística, que ha pasado a considerarse una de las ciencias metodológicas fundamentales y base del método científico experimental. La enseñanza de la estadística, sin embargo, aún se encuentra en sus comienzos, aunque como hemos descrito parece avanzar de una forma imparable. ¿Será el siglo XXI el siglo de la educación estadística? Analizaremos a continuación algunos indicadores que parecen dar una respuesta positiva a esta pregunta, así como los cambios previsibles en nuestros métodos de enseñanza de esta materia.

Un primer indicador de la expansión futura de la educación estadística son los trabajos previstos por *IASE*, como el congreso *ICOTS VI* en el 2002 y las sesiones de educación en la Conferencia *ISI* de Corea en 2001. Oras sociedades de estadística o de educación están también organizando de secciones especificas de educación estadística, como, por ejemplo, la *ASA* (*American Statistical Association*) *AERA* (*American Educational Research Association, Royal Statistical Society*, en Inglaterra, *Sociedad estadística Japonesa, la Sociedad Española de Investigación en Educación Matemática*, etc..

Las revistas orientadas a los profesores de estadística sugieren una problemática docente y un interés de los profesores por mejorar su acción docente. El mejor exponente lo tenemos en *Teaching Statistics*, que ha cumplido ya 22 años de existencia durante los cuales se ha ido desarrollando y adquiriendo una identidad y calidad internacional reconocida. Otras revistas similares son *Induzioni* y *Journal of Statistical Education*, que es una revista de educación estadística a nivel universitario con un servicio de información asociado.

Algunas de las asociaciones que hemos nombrado preparan boletines de noticias que distribuyen por Internet con un sistema semejante al de las revistas electrónicas, como la *Newsletter del Statistical Education Research Group*. A nivel docente podemos citar el boletín del *Statistics Teacher Network*, que es una asociación de profesores de estadística en los niveles de primaria y secundaria.

A la vista de todas estas posibilidades, surge la pregunta de hacia donde va la educación estadística y que tipo de enseñanza tendrá· lugar en el futuro (Hawkins, 1997). Es difícil dar una respuesta, aunque los libros de texto se empiezan a transformar a ediciones electrónicas e incluso en formato accesibles a la consulta, modificación y sugerencias a través de Internet. Es también sencillo obtener datos de todo tipo para que los estudiantes puedan realizar investigaciones sobre casi cualquier tema, incluso con pocos recursos disponibles. El profesor puede cargar estos conjuntos de datos desde Internet e introducirlos en los ordenadores o las calculadoras gráficas de los alumnos que tienen una difusión mucho mayor. De este modo los alumnos pueden trabajar con los datos en casa o exportarlos a otros ordenadores o calculadoras. también pueden combinar diferentes conjuntos de datos en un mismo proyecto o "enviar" a la red sus propias colecciones de datos para que sean usadas por nuevos estudiantes en cualquier rincón del planeta.

Las listas de discusión entre profesores o entre alumnos, la "tutoría" de alumnos a distancia, cuando el trabajo del alumno no permite la comunicación directa con el profesor son ya hechos cada vez mas cercanos y ya están siendo implementados en forma experimental en algunas escuelas y universidades, como, por ejemplo, la experiencia australiana de formación a distancia de profesores (Watson y Baxter, 1998). La rapidez del cambio tecnológico hace previsible la extensión de estas nuevas formas de enseñanza y aprendizaje en un plazo de tiempo no muy lejano.

2

FUNDAMENTOS EPISTEMOLÓGICOS

2.1. Introducción

Una vez situada la educación estadística, seguiremos con un estudio epistemológico, que es bastante importante para la didáctica. La estadística, a pesar de contar con una axiomática satisfactoria, es quizás la única rama de las matemáticas donde prosiguen hoy día las discusiones sobre la interpretación de conceptos básicos. Esta controversia no es de tipo técnico, ya que desde el punto de vista matemático, cualquier concepto estadístico queda determinado por su definición. Por ejemplo, la probabilidad sería cualquier función medible normada de un algebra de sucesos en el intervalo [0, 1]. Los problemas filosóficos que la axiomatización no ha resuelto se refieren a las posibilidades de aplicación de los conceptos estadísticos y la interpretación de los mismos en diferentes circunstancias.

Si el profesor no es consciente de esta problemática, difícilmente podrá comprender algunas dificultades de sus estudiantes, quienes necesitan materializar en ejemplos concretos los conceptos y modelos matemáticos. En el ejemplo dado, la definición no resuelve el problema de asignar probabilidad a sucesos como “mañana lloverá” o “habrá 3 ganadores máximos en la próxima quiniela”.

En lo que sigue discutiremos algunos conceptos básicos, comenzando por la misma noción de estadística, a partir de las reflexiones realizadas por estadísticos, filósofos, psicólogos e investigadores en Didáctica de la Matemática preocupados por este problema.

2.2. Estadística

Son muchas las definiciones posibles de estadística, y entre ellas hemos elegido la siguiente que refleja bien nuestra conceppción del tema:

> *"La estadística estudia el comportamiento de los fenómenos llamados de colectivo. Está caracterizada por una <u>información</u> acerca de un <u>colectivo</u> o universo, lo que constituye su <u>objeto material</u>; un modo propio de razonamiento, el <u>método estadístico</u>, lo que constituye su <u>objeto formal</u> y unas <u>previsiones</u> de cara al futuro, lo que implica un ambiente de incertidumbre, que constituyen su <u>objeto o causa final</u>." (Cabriá, 1994).*

Los orígenes de la estadística son muy antiguos, ya que se han encontrado pruebas de recogida de datos sobre población, bienes y producción en las civilizaciones china (aproximadamente 1000 años a. c.), sumeria y egipcia. Incluso en la Biblia, en el libro de Números aparecen referencias al recuento de los israelitas en edad de servicio militar. No olvidemos que precisamente fué un censo lo que motivó del viaje de José y María a Belén, según el Evangelio. Los censos propiamente dichos eran ya una institución el siglo IV a.C. en el imperio romano.

Sin embargo sólo muy recientemente la estadística ha adquirido la categoría de ciencia. En el siglo XVII surge la aritmética política, desde la escuela alemana de

Conring, quien imparte un curso son este título en la universidad de Helmsted. Posteriormente su discípulo Achenwall orienta su trabajo a la recogida y análisis de datos numéricos, con fines específicos y en base a los cuales se hacen estimaciones y conjeturas, es decir se observa ya los elementos básicos del método estadístico. Para los aritméticos políticos de los siglos XVII y XVIII la estadística era el arte de gobernar; su función era la de servir de ojos y oídos al gobierno.

La proliferación de tablas numéricas permitió observar la frecuencia de distintos sucesos y el descubrimiento de leyes estadísticas. Son ejemplos notables los estudios de Graunt sobre tablas de mortalidad y esperanza de vida a partir de los registros estadísticos de Londres desde 1592 a 1603 o los de Halley entre 1687 y 1691, para resolver el problema de las rentas vitalicias en las compañías de seguros. En el siglo XIX aparecen las leyes de los grandes números con Bernouilli y Poisson.

Otro problema que recibe gran interés por parte de los matemáticos de su tiempo, como Euler, Simpson, Lagrange, Laplace, Legendre y Gauss es el del ajuste de curvas a los datos. La estadística logra con estos descubrimientos una relevancia científica creciente, siendo reconocida por la British Association for the Advancement of Science, como una sección en 1834, naciendo así la Royal Statistical Society. En el momento de su fundación se definió la estadística como *"conjunto de hechos, en relación con el hombre, susceptibles de ser expresados en números, y lo suficiente numerosos para ser representados por leyes"*.

Se crearon poco a poco sociedades estadísticas y oficinas estadísticas para organizar la recogida de datos estadísticos; la primera de ellas en Francia en 1800. Como consecuencia, fue posible comparar las estadísticas de cada país en relación con los demás, para determinar los factores determinantes del crecimiento económico y comenzaron los congresos internacionales, con el fin de homogeneizar los métodos usados. El primero de ellos fue organizado por Quetelet en Bruselas en 1853. Posteriormente, se decidió crear una sociedad estadística internacional, naciendo en 1885 el *Instituto Internacional de Estadística (ISI)* que, desde entonces celebra reuniones bianuales. Su finalidad especifica es conseguir uniformidad en los métodos de recopilación y abstracción de resultados e invitar a los gobiernos al uso correcto de la estadística en la solución de los problemas políticos y sociales. En la actualidad el ISI cuenta con 5 secciones, una de las cuales, la IASE, fundada en 1991, se dedica a la promoción de la Educación Estadística.

- *Panorama actual*

Aunque es difícil dividir la estadística en partes separadas, una división clásica hasta hace unos 30 años ha sido entre *estadística descriptiva y estadística inferencial.*

La *estadística descriptiva* tiene como fin presentar resúmenes de un conjunto de datos y poner de manifiesto sus características, mediante representaciones gráficas. Los datos se usan para fines comparativos, y no se usan principios de probabilidad. El interés se centra en describir el conjunto dado de datos y no se plantea el extender las conclusiones a otros datos diferentes o a una población.

La *inferencia estadística,* por el contrario, estudia los resúmenes de datos con referencia a un modelo de distribución probabilistico o una familia de modelos, determinando márgenes de incertidumbre en las estimación de los parámetros desconocidos del mismo. Se supone que el conjunto de datos analizados es una muestra de una población y el interés principal es predecir el comportamiento de la población, a partir de los resultados en la muestra.

Esta división es hoy demasiado simple y han surgido diferentes corrientes dentro de la estadística. Por ejemplo, es común hablar de *análisis de datos*. Pero este término tiene diferentes significados. Analizaremos con más detalle el análisis exploratorio de datos, inferencia estadística y análisis multivariante de datos. Antes de ello estudiaremos la problemática epistemológica asociada a la noción de aleatoriedad, así como a las ideas estocásticas fundamentales.

Actividad 2.1. Discutir las siguientes afirmaciones extraídas del artículo de D. S. Moore: "Teaching Statistics as a respectable subject". En F. Gordon y S. Gordon (eds.), Statistics for the Twenty-First Century, (pp. 14-25).

1. *La estadística es una disciplina científica autónoma, que tiene sus métodos específicos de razonamiento;*
2. *Aunque es una ciencia matemática, no es un subcampo de la Matemática. La estadística no ha surgido de la matemática;*
3. *Aunque es una disciplina metodológica, no es una colección de métodos;*
4. *La estadística es la ciencia de los datos. Con más precisión, el objeto de la estadística es el razonamiento a partir de datos empíricos. Los datos no son números, sino números en un contexto;*
5. *En los últimos años la tecnología ha hecho que la investigación y la práctica estadística se distancie cada vez más de la matemática;*
6. *La estadística tiene sus propias controversias, que están muy alejadas de las controversias relacionadas con los fundamentos de las matemáticas;*
7. *La posición que un estadístico toma sobre las cuestiones de fundamentos tiene un impacto inmediato en su práctica estadística;*
8. *La relación entre estadística y matemáticas se produce en un único sentido (no es biunívoca); la estadística toma conceptos matemáticos para el desarrollo de sus métodos, en cambio la matemática no toma conceptos estadísticos.*
9. *La probabilidad es más cercana a la matemática que la estadística.*

¿Cuales son las implicaciones para la enseñanza de la estadística en la escuela primaria y secundaria?

2.3. Aleatoriedad

El Cálculo de Probabilidades se ocupa del estudio de los fenómenos aleatorios. Ahora bien. ¿Qué es la aleatoriedad? ¿Es una propiedad de los fenómenos a los que aplicamos esta calificativo? ¿De donde surge esta idea? A continuación discutimos este concepto, partiendo del trabajo de Batanero y Serrano (1995).

Actividad 2.2. Se pidió a algunos niños lanzar una moneda 150 veces. Algunos lo hicieron correctamente. Otros hicieron trampas. Anotaron con la letra C la aparición de una cara y con X una cruz. Estos son los resultados de Daniel y Diana:

Daniel: c+c++cc++cc+c+c++c++c+ccc+++ccc++c++c+c+c++cc+ccc+
c+c+cc+++cc++c+c++cc+c++cc+c++cc+cc+c+++c++cc++c++
c+c+cc+c++cc+c+c++ccc+cc++c+c++cc+++c+++c+c++ccc++

Diana: +cc+++c++++c+cc+++cc+cc+++cc+ccc+++c++++++c+c+c+c+

+++cccccc+ccc+c+cc+ccccc+ccc++ccc+c+ccccccccc++c+

ccccccc+++++cccc++c+c+cc+cc+cc++++++c+cc++ccc++ccc

¿Hicieron trampas Daniel o Diana? ¿Por qué?.

¿Cual es la solución correcta al problema? ¿Cómo lo resolverías? ¿Podrías dar una lista de variables que podrías cambiar en el ítem 1 para hacer variar su dificultad (variables del campo de problemas de donde surge la idea de aleatoriedad)? ¿Qué otros colectivos de personas se interesarían por estos problemas? ¿Qué harían para resolverlo?

2.2.1. Aleatoriedad y causalidad

Una primera acepción de lo aleatorio se recoge en el diccionario de M. Moliner (1983):

> *"Incierto. Se dice de aquello que depende de la suerte o del azar", siendo el azar "la supuesta causa de los sucesos no debidos a una necesidad natural ni a una intervención humana ni divina*".

En esta definición aleatorio es aquello cuyas causas son desconocidas y el "azar" estaría personificado como la causa de los fenómenos aleatorios. Esta interpretación corresponde a la primera fase histórica en el desarrollo de la idea de aleatoriedad que finaliza al comienzo de la Edad Media. En esta etapa los dados o huesos de astrálago se usaron para echar a suertes, tomar decisiones y en ceremonias religiosas. puesto que el azar suprimía influencia de la voluntad, inteligencia o conocimiento del hombre. Según Poincaré (1936), los filósofos clásicos diferenciaban los fenómenos que parecían obedecer a leyes conocidas y los aleatorios, que no podían preverse porque se rebelaban ante toda ley. La aleatoriedad tenía, además, un sentido preciso, objetivo. Lo que era aleatorio lo era para cualquier persona.

Otra interpretación es suponer que todo fenómeno tiene una causa y que el azar es debido a nuestra ignorancia: *"Nada sucede por azar sino que todo ocurre por una razón y por una necesidad*" (Leucippus, siglo V a.c.). Esta es una interpretación subjetiva porque lo que es aleatorio para una persona puede no serlo para otra con más conocimientos:

> *"El azar no es más que la medida de nuestra ignorancia. Los fenómenos fortuitos son, por definición aquellos cuyas leyes ignoramos".* (Poincaré, 1936, pg. 69 de la reproducción en el libro de Newmann).

Que esta interpretación no es adecuada se deduce de la existencia de fenómenos deterministas cuyas leyes no conocemos, como la muerte y de otros fenómenos aleatorios, como la transmisión de caracteres genéticos que tienen causas conocidas. En nuestra moderna concepción, los fenómenos aleatorios son aquellos a los que podemos aplicar el cálculo de probabilidades, que seguirá siendo válido el día que encontremos sus reglas. Así, el jefe de ventas de una empresa ignora cuando o cuanto comprará cada uno de sus clientes. Sin embargo, sus beneficios seguirán siendo los mismos, incluso cuando lo conociese, porque la distribución del consumo en la población sería la misma, independientemente de este conocimiento.

Cuando una causa pequeña produce un efecto considerable, decimos que el resultado es aleatorio, porque la predicción resulta imposible, como en los juegos de azar, terremotos u otros desastres naturales. Es precisamente la complejidad o multitud de las causas lo que muchas veces determina un resultado aleatorio.

Actividad 2.3. Elaborar una lista de sinónimos del término "aleatoriedad. ¿Son totalmente equivalentes? Buscar en diversos diccionarios (Diccionario de la Lengua, de uso del español, sinónimos, etc.) y encontrar matices diferenciadores. Elaborar una lista de términos relacionados.

Actividad 2.4. ¿Como está presente el azar en los juegos, narrativas y actividades infantiles?

Actividad 2.5. Estudiar algunas creencias culturales infundadas sobre la posibilidad de control de lo aleatorio? ¿Como se ponen de manifiesto? ¿Podrías aportar datos al respecto?

2.2.2. Aleatoriedad y probabilidad

Cuando comienza el cálculo de probabilidades, por ejemplo en el *Liber de Ludo Aleae* de Cardano, se relaciona la aleatoriedad con la equiprobabilidad de los diferentes resultados, es decir, un fenómeno sería aleatorio si todos sus resultados son equiprobables. Esta interpretación se aceptó con facilidad, debido a que los primeros desarrollos del cálculo de probabilidades estuvieron muy ligados a los juegos de azar, en los que el número de posibilidades es finito y el principio de indiferencia de las diferentes posibilidades puede considerarse razonable.

Hacia el final del siglo XVIII y principios del XIX se amplía el número de situaciones consideradas aleatorias, incluyendo no sólo los juegos de azar, sino muchos fenómenos naturales. Paralelamente, se produce un cambio en el concepto de aleatoriedad, que se hace progresivamente más formalizado, introduciendo la idea de "independencia", que se considera imprescindible para asegurar la aleatoriedad de un suceso en experimentos repetidos.

- **Concepción clásica**

La noción de aleatoriedad ha estado ligada a las diferentes concepciones sobre la probabilidad (Véase Godino, Batanero y Cañizares, 1987). En una *concepción clásica*, la probabilidad de un suceso es el "*cociente entre el número de casos favorables al suceso y el número de casos posibles, siempre que todos sean equiprobables·*.

En esta acepción de la probabilidad, consideramos que un objeto (o un suceso) es un miembro aleatorio de una cierta clase de objetos (población), si la probabilidad de obtener este objeto (en un sorteo u otro experimento) es igual que la de cualquier otro miembro de su clase. En la lotería nacional o el sorteo de excedentes de cupo en el servicio militar, cada número (o cada mozo) sería un miembro aleatorio del conjunto de miembros en el sorteo.

Esta definición de aleatoriedad fue considerada suficiente en una primera etapa, donde el cálculo de probabilidades se limitaba a los juegos de azar basados en dados, monedas, cartas, extracción de bolas en urnas, etc, Sin embargo, pronto se le aplicaron algunas críticas, así como a la concepción de probabilidad en que se apoya.

Hay en esta definición una circularidad difícil de evitar, lo que haría también difícil usarla para discriminar un miembro aleatorio o no aleatorio en una clase dada. Por ejemplo, ¿como sabemos que una ruleta dada o un dado, no están ligeramente sesgados? ¿O qué el sorteo de excedentes de cupo es realmente equitativo? Además sólo podríamos decir que un objeto es un miembro aleatorio de una clase, si la clase es finita. Si fuese infinita, la probabilidad de cada miembro de la clase siempre sería nula (por tanto idéntica), aunque el método de selección fuese sesgado.

¿Que ocurriría, finalmente, en el caso de sucesos elementales no equiprobables? Si aplicamos esta definición, no podríamos considerar que el color del pelo o el grupo sanguíneo sea una característica aleatoria, ya que hay mayor proporción de personas morenas que rubias en la población y hay menos personas con Rh negativo. A pesar de estos problemas, la idea de equiprobabilidad se usa todavía para definir la aleatoriedad en situaciones como definir una muestras aleatoria -todos los elementos tienen la misma probabilidad de ser elegidos- o la asignación aleatoria de sujetos a los experimentos -cada sujeto tiene la misma probabilidad de ser asignado al grupo experimental o al control.

- **Concepción frecuencial**

Cuando queremos aplicar la idea de probabilidad a situaciones del mundo físico o natural, como la meteorología, el resultado de elecciones, accidentes, etc. nos encontramos con que no podemos aplicar el principio de equiprobabilidad Usando la *concepción frecuencial de la probabilidad*, podríamos considerar que un objeto es un miembro aleatorio de una clase si pudiéramos elegirlo mediante un método que proporcionase a cada miembro de la clase una cierta frecuencia relativa "a priori" a la larga.

Esta definición es muy útil cuando disponemos de datos estadísticos sobre un gran número de casos, como en los ejemplos citados, aunque tenemos el problema teórico de decidir cuántos experimentos se necesitan para considerar que, a partir de este número, habríamos probado suficientemente el carácter aleatorio del objeto. Esta definición de la probabilidad no proporciona, además, un valor exacto de la probabilidad, sino sólo una estimación del mismo.

- **Concepción subjetiva**

En los dos casos anteriores la aleatoriedad y la probabilidad son propiedades "objetiva" que se asigna al suceso o elemento de una clase, como podría asignársele una profesión, estado civil o nacionalidad, si se trata de una persona. Kyburg (1974) critica esta visión e indica que la idea de aleatoriedad está compuesta de cuatro términos:

- el objeto que se supone es miembro aleatorio de una clase;
- el conjunto del cual el objeto es un miembro aleatorio (población o colectivo);
- la propiedad con respecto a la cual el objeto es un miembro aleatorio de la clase dada;
- el conocimiento de la persona que decide si el objeto es aleatorio o que asigna una probabilidad.

La decisión sobre si consideramos que un objeto es o no un miembro aleatorio de una clase, depende de nuestro conocimiento sobre el mismo. Lo que puede ser aleatorio para una persona puede no serlo para otra. y la aleatoriedad no es una propiedad física "objetiva", sino que tiene un carácter subjetivo. Utilizamos ahora la *concepción subjetiva* de la probabilidad, por la que todas las probabilidades serían condicionales y es más conveniente en las situaciones en que poseemos cierta información que puede cambiar nuestro juicio sobre la aleatoriedad o la probabilidad de un suceso.

Hay algunas situaciones que todo el mundo consideraría aleatorias y donde el uso de la idea de equiprobabilidad para definir un suceso aleatorio parece claro y no controvertido. Por ejemplo, en el caso de un dado equilibrado, cualquier lanzamiento es simplemente un ejemplo de cualquier otro posible lanzamiento. No hay nada nuevo que podamos conocer acerca del dado que nos permita predecir otra probabilidad diferente de 1/6 para un resultado particular del dado.

En otros casos la situación no es tan clara. Consideremos, por ejemplo, la probabilidad de que un chico apruebe el examen de conducir. Si tenemos que dar un valor para esta probabilidad, nuestras estimaciones serán muy diferentes en caso del examinador o de su profesor de la autoescuela, que ha estado observando como conduce a lo largo de los últimos dos meses.

Actividad 2.6. En los siguientes sucesos indica cuales consideras aleatorios y qué concepción de aleatoriedad puede aplicársele

1. *Germinación de una semilla plantada con cuidado*
2. *El resultado obtenido en un dado que yo he lanzado y puedo ver, pero tu no*
3. *El número ganador de la lotería el próximo mes*
4. *El color del próximo jersey que te compres*
5. *Si llovió en Melbourne el 3 de Agosto pasado*
6. *Si lloverá en Melbourne el 3 de Agosto próximo*
7. *Que cojas la gripe este año*
8. *Que estés expuesto a coger la gripe este año*
9. *El resultado de lanzar una moneda, si la he lanzado 10 veces y he obtenido 10 caras*

2.2.3. Procesos y secuencias aleatorias

Según Zabell (1992) la idea de aleatoriedad contiene dos aspectos distintos que, a veces, pueden no coincidir, que son:

- el proceso de generación, que es lo que, matemáticamente se conoce como experimento aleatorio;
- el patrón de la secuencia aleatoria producida como consecuencia del experimento.

Muchas personas piensan que estos dos aspectos están ligados entre sí y esperan que sólo un proceso aleatorio proporcione un patrón de resultados aleatorios. Sin embargo, puede haber secuencias que, aparentemente, parezcan aleatorias, siendo producidas por un proceso completamente determinista, como cualquier trozo de la secuencia de números decimales del número Pi u otro número irracional.

Por otro lado, incluso una secuencia aparentemente regular y determinista, como C+C+ puede obtenerse como resultado de lanzar cuatro veces una moneda y esta secuencia particular, en contra de la creencia popular es tan probable como CCC+ o ++C+.

A final del siglo XIX Edgeworth, Galton, Pearson, Fisher y sus colaboradores inician el estudio de métodos de inferencia, basados en la elección de muestras aleatorias de datos. Hasta esa fecha los estudios estadísticos eran puramente descriptivos, analizando grandes masas de datos empíricos. En los casos en que se utilizaban muestras, estas no se elegían al azar, porque no se era consciente de la importancia de elegir muestras aleatorias para poder realizar inferencias válidas y precisas sobre la población. Cuando los nuevos estudios muestran la importancia de las muestras aleatorias, empieza el interés por encontrar modelos de procesos que aseguren la consecución de largas secuencias de dígitos aleatorios, a partir de los cuales se tomarían las muestras.

Se puede generar resultados aleatorios con dispositivos físicos, como extracción de bolas, dados o ruletas. Es el método que usamos en juegos y sorteos, como la lotería. Sin embargo, es muy difícil conseguir construir dispositivos que aseguren la aleatoriedad física perfecta y es un procedimiento lento. Las tablas de números aleatorios, iniciadas por Lord Kelvin, Fisher, Yates y Kendall surgen para sustituir estos dispositivos físicos. Este tipo de tablas puede también ser usado en clase con los alumnos de educación secundaria.

Para generar números aleatorios podemos también usar una fórmula (los llamados "números pseudo-aleatorios"). Con ayuda de un algoritmo un ordenador o calculadora produce una secuencia numérica que puede ser empleada como aleatoria para los propósitos prácticos. Asimismo, podemos obtener sucesiones aleatorias extraídas de distintos tipos de distribución teórica, como la distribución normal con una cierta media y varianza.

Actividad 2.7. Investigar algunos algoritmos de generación de números aleatorios ¿Cómo podrías asegurar la "calidad" de un algoritmo o una tabla de números aleatorios? ¿Recuerdas algunos test de aleatoriedad clásicos?

2.2.4. Formalización de la Idea de Aleatoriedad

Los autores de las primeras tablas de números aleatorios se preocuparon por asegurar su "calidad". Puesto que era posible obtener números pseudo-aleatorios con ayuda de algoritmos deterministas, debía examinarse la sucesión producida, independientemente del proceso que la genera.

En el siglo XX varios estadísticos estudian este problema, finalizando con la formalización del concepto de aleatoriedad. Una idea intuitiva es que con una sucesión aleatoria es imposible inventar un método que nos permita ganar en un juego de azar cuyos resultados fuesen los de la sucesión; Otra idea intuitiva es que la sucesión aleatoria como altamente irregular o compleja.

- **Enfoque de los algoritmos de selección**

Von Mises comenzó su estudio definiendo la población o colectivo, como fenómeno de masas, suceso repetitivo o larga serie de observaciones, para las cuales hay razones suficientes para creer que la frecuencia relativa de un suceso tiende a un límite fijo.

A partir de esta idea, considera que una secuencia de observaciones es aleatoria si la frecuencia relativa de cada posible suceso en la sucesión no varía en ninguna parte de la misma. Dicho de otra forma, en una secuencia aleatoria no puedo encontrar un algoritmo por el cual seleccionar una subsecuencia en la cual la frecuencia relativa de uno de los resultados se vea afectada.

Supongamos, por ejemplo, que al elegir todos los elementos pares, o cada diez elementos, o todos los elementos primos en una sucesión de caras y cruces obtenidas al lanzar una moneda obtuviese una secuencia en la que la probabilidad de cara fuese 3/5. Consideraría, entonces, que la secuencia dada no era realmente aleatoria, ya que podría obtener ventaja en un juego de azar apostando a favor de la cara cada dos, diez jugadas o en las jugadas número primo.

Esta definición de aleatoriedad se usó para inventar contrastes de aleatoriedad que se aplicaban a las tablas de números aleatorios, antes de ponerlas en uso en la comunidad científica. Estos contrastes tratan de detectar las posibles regularidades de la sucesión. Sin embargo, puesto que todo contraste de hipótesis siempre lleva asociado una posible probabilidad de error, nunca podríamos tener total seguridad de que una sucesión dada, a pesar de haber pasado todas las pruebas, tuviese algún tipo de patrón que nos hubiera pasado desapercibido y, por tanto, no fuese totalmente aleatoria. Además, en toda sucesión finita podríamos tomar una subsucesión que alterase la frecuencia de uno de los resultados, por lo que esta definición sólo vale para sucesiones infinitas.

- **Enfoque de la complejidad absoluta**

Otro intento de definir la aleatoriedad de una sucesión, debido a Kolmogorov, se basa en su complejidad. La complejidad de una sucesión es la dificultad de describirla (o almacenarla en un ordenador) mediante un código que nos permita reconstruirla más tarde. El mínimo número de signos necesarios para codificar esta sucesión proporciona una escala para medir la dificultad de almacenarla, o, lo que es lo mismo, su complejidad. Por ejemplo, si comparamos las secuencias:

CC++C+CC++
C+C+C+C+C+

vemos que puedo codificar mas abreviadamente la primera como CC+2+CC++, mientras que la segunda la podríamos codificar como 4C+. Por lo tanto la primera sucesión es más compleja que la segunda, ya que precisa más signos para codificarse.

Bajo este enfoque, una secuencia sería aleatoria si no se puede codificar en forma más abreviada, por lo que la ausencia de patrones periódicos sería su característica esencial. Es decir, una secuencia sería aleatoria si sólo podemos describirla listando uno tras otros todos sus componentes. Esta definición establece una jerarquía en los grados de aleatoriedad de diferentes sucesiones y la aleatoriedad perfecta sería un concepto teórico o límite.

Como resumen, podemos decir que el concepto de aleatoriedad se puede interpretar desde un punto de vista formal o informal. Desde el punto de vista informal, hablamos del azar, como patrón que explica aquellos efectos para los que no conocemos la causa o que no son predecibles. Esta interpretación refleja muchos aspectos filosóficos epistemológicos sociales y personales y es una importante fuente de confrontación entre las probabilidades subjetivas y objetivas. Desde el punto de vista formal la idea central es la sucesión de resultados de un mismo experimento realizado repetida e independientemente, que lleva, en su caso más sencillo a la sucesión de ensayos de Bernoulli.

Actividad 2.8. Estudiar la complejidad de las secuencias propuestas en el ítem 1 y decidir cual sería más aleatoria en este enfoque. Generar 10 secuencias de dígitos aleatorios con la calculadora y ordenarlas según su complejidad de codificación.

Actividad 2.9. El modelo más simple de secuencia de resultados aleatorios es la secuencia de ensayos de Bernoulli. Dar una lista de modelos aleatorios que se deducen de este tipo de secuencia. Analizar las secuencias presentadas en el ítem 1, desde el punto de vista de este modelo. ¿Tendríamos ahora argumentos para afirmar que una de las niñas hace trampas?

Actividad 2.10. Pablo juega a un juego con 16 fichas numeradas 1, 2, 3, 4,...16. Pablo pone las fichas en una caja y la mueve con fuerza. Toma una ficha al azar y es el 7. Pone una cruz en la casilla 7.

1	*2*	*3*	*4*
5	*5*	*7 x*	*8*
9	*10*	*11*	*12*
13	*14*	*15*	*16*

Otros niños jugaron al mismo juego, aunque algunos hicieron trampas.

X	*x*	*x*	*X*
x	*x*	*X*	*X*
X	*x*	*X*	*x*
x	*x*	*X*	*X*

X X x		*X*	
	x	*X X*	*X xX*
x	*x*		*x X*
	x x		

Luis

Jaime

x x	*X X*		
		x X	*X X*
	X X		*X X*
	x x		*x x*

		x	*xx*
		X x X	*X X*
x	*x X*		
x	*XxxX*		

Jesus

Maria

¿Qué modelo aleatorio podemos aplicar a estas situaciones? ¿Como podrías justificar cuales de los niños hacen trampas? Elabora una lista de situaciones de la vida real en que el modelo que has encontrado sea útil y en donde algunas personas se interesen por resolver un problema similar al dado.

2. 4. Ideas estocásticas fundamentales

Además de la ideas básicas de aleatoriedad hay una serie de conceptos sobre los cuáles se apoya todo el cálculo de probabilidades y la estadística. Es importante identificar estos conceptos, que son los que debemos enseñar en los niveles no universitarios. Para elaborar una lista de ideas estocásticas fundamentales, analizaremos el artículo de Heitele (1975), quien se basa en el libro de *Bruner "El proceso de educación"* , publicado en 1960, que sostiene las tesis siguientes:

1. El principio decisivo de instrucción en un tema es la transmisión de las ideas fundamentales.
2. Cualquier tema puede ser enseñado efectivamente en alguna forma correcta a cualquier niño en cualquier estado de desarrollo, lo que implica que las ideas fundamentales son una guía necesaria desde la escuela primaria a la universidad para garantizar una cierta continuidad.

3. La ideas fundamentales son aquellas ideas y conceptos que se usarán en diferentes niveles cognitivos y lingüísticos en una "espiral curricular".
4. La transición a un nivel cognitivo superior se facilita si el tema subyacente ha sido preparado en una representación conveniente en etapas cognitivas anteriores.

Las ideas fundamentales proporcionan al niño modelos explicativos eficientes en cada etapa de su desarrollo, que difieren en los distintos niveles cognitivos, no de una forma estructural, sino sólo de una forma lingüística y en su nivel de profundización. Por ejemplo, si un niño, al lanzar dos dados concede más probabilidad al 7, porque hay más sumas con éste valor, tiene un modelo apropiado, que puede evolucionar al más complejo de aplicación de la regla de Laplace, e incluso al de variable aleatoria y moda de su distribución de probabilidad. Por el contrario, si un niño al lanzar dos monedas explica la mayor proporción de casos "mixtos" argumentando que tras "cara" es más probable "cruz", usa un modelo de explicación que le satisface, pero no permite continuación a un estadio más elaborado. Se trata de una "intuición errónea".

En el campo de la probabilidad la intuición juega un papel muy importante. Los modelos intuitivos explicatorios tienen dos funciones:

- En una edad temprana ayudan al niño a entender su entorno por sus propios medios, mucho antes de que sea capaz de comprender la complejidad del modelo matemático subyacente.
- Preparan el conocimiento analítico posterior.

El gran número de paradojas estocásticas puede confundir incluso a los expertos. Por ello, es más importante construir intuiciones correctas en este campo que en ningún otro. Como señala Feller, la intuición estadística puede ser entrenada incluso en los adultos, aunque si un niño adquiere "intuiciones erróneas" cuando es muy pequeño, esto le puede impedir mas tarde la adquisición de un conocimiento adecuado. Por ello parece necesario ofrecer a los niños actividades estocásticas, en forma de juegos y experimentos en edad temprana. Sin embargo, las actividades a desarrollar no deben escogerse al azar. Es preciso un principio de organización, en la forma de la espiral curricular, escalonada por las ideas fundamentales. A continuación analizamos los 10 grupos de ideas fundamentales en estocástica que propone Heitele.

- **La probabilidad como normalización de nuestras creencias**.

La primera idea fundamental es asignar números a los sucesos aleatorios, de forma que estos números reflejen nuestro grado de creencia en su verificación. En el lenguaje ordinario se usan expresiones del tipo de "casi cierto", "más probable que", etc., para comparar los sucesos aleatorios, pero estas expresiones son poco precisas, porque diferentes personas les conceden diferente valor.

Se normalizan estas expresiones asignándoles un valor en la escala de la probabilidad. De este modo ponemos en correspondencia la multidimensionalidad del complejo mundo a nuestro alrededor con el intervalo [0,1], y se hace accesible a los dispositivos matemáticos. Podemos comparar sucesos muy dispares, en base a su mayor o menor probabilidad. En la idea de probabilidad, tan sencilla, pero tan potente, encontramos un ejemplo de cómo el hombre idea modelos matemáticos para comprender y predecir la realidad.

Si en la idea de probabilidad (asignar números a los sucesos aleatorios de modo que sean comparables) incluimos todos los procesos de cálculo que llevan a asignar estos números (probabilidad a priori, a posteriori, subjetivas, etc.), habremos incluido toda la teoría de la probabilidad en esta primera idea. No obstante, en el estudio que hemos hecho con anterioridad, vemos que, a pesar de su aparente simplicidad, esta idea no está libre de controversias.

Actividad 2.11. Se toman 3 fichas de la misma forma y tamaño, de las cuales una es roja por ambas caras; otra es azul por una cara y roja por la otra, y la tercera es azul por las dos caras. El profesor coloca las tres fichas en una caja, que agita convenientemente, antes de seleccionar una de las tres fichas, al azar. Muestra, a continuación, una de las caras de la ficha elegida, manteniendo la otra tapada, pidiendo a sus alumnos que adivinen el color de la cara oculta. Una vez hechas las apuestas, el profesor muestra la cara oculta. Cada alumno que haya acertado en la predicción efectuada, consigue un punto. Se trata de buscar la mejor estrategia en este juego.
¿Qué tipo de razonamiento has dado (o darías) para validar que tu estrategia es la óptima? ¿Es igualmente válido el argumento que se basa en la experimentación que el basado en consideraciones lógicas y combinatorias? ¿Podrías probar que tu estrategia es la mejor, sólo con la experimentación?

Actividad 2.12. A partir de las noticias en la prensa del último Domingo, elaborar una lista de sucesos que podrían ocurrir (o no) durante el mes en curso. Asignar un valor de probabilidad a cada uno de estos sucesos. Revisar las probabilidades iniciales, al cabo de una semana, indicando las posibles informaciones que han influido en los cambios de asignación de probabilidades.

- **El espacio muestral como conjunto de todas las posibilidades**

No menos fundamental es la idea debida a Kolmogorov de asignar un espacio muestral de sucesos observables a cada experimento aleatorio y representar los sucesos observables como subjconjuntos del espacio muestral, dando una interpretación probabilística a las operaciones con sucesos. Es decir, inventariar todos los posible sucesos elementales asignados a un experimento, considerarlo como un conjunto de referencia o universal y aplicar toda la potencia del álgebra de conjuntos para poder definir los demás sucesos, a partir de los sucesos elementales

Esta idea permitió axiomatizar la probabilidad, como medida normada aditiva, sobre el álgebra de conjuntos, puesto que las operaciones en este álgebra de conjuntos permitían definir operaciones sobre la misma probabilidad. Puesto que todo suceso elemental forma parte del conjunto de referencia, se dota de sentido al muestreo, ya que al observar repetidamente una serie de repeticiones del experimento, siempre observaremos elementos del espacio muestral.

Por otro lado, si un subconjunto está incluido en un conjunto mayor, la probabilidad del primero es menor que la del segundo. Esta propiedad permite, desde la escuela elemental realizar actividades de probabilidad comparativa, incluso sin cuantificación.

El inventariar todos los posibles resultados del experimento tiene, sin embargo, a veces dificultades para los niños que no han alcanzado un nivel de razonamiento combinatorio suficiente. La dificultad está en que hay que considerar no sólo el suceso que ha ocurrido realmente o incluso el suceso de interés sino todos los sucesos que *podrían ocurrir.* En la investigación del desarrollo del concepto de azar se ha mostrado que los niños, como algunos adultos supersticiosos están confinados estrechamente al determinismo: Creen en fuerzas ocultas que expliquen los fenómenos aleatorios. Fischbein y Cohen explican esto por que se concentran en un sólo suceso, en lugar de hacerlo sobre la totalidad.

Actividad 2.13. Ruletas no transitivas: Supongamos que tenemos tres ruletas. Con la primera siempre obtenemos el número 3. Con la segunda obtenemos el número 1 con probabilidad 0.52 y el número 5 con probabilidad 0.48. Con la tercera obtenemos el número 0 con probabilidad 0.25 y el número 4 con probabilidad 0.75.
Jugamos a un juego en el que dos jugadores eligen una ruleta cada uno y gana aquél que consiga el número mayor al girar la ruleta. ¿Cuál jugador tiene ventaja, el que elige la ruleta en primer o segundo lugar?

- **Regla de adición de probabilidades**

Una regla general en matemáticas es construir modelos complejos a partir de otros más simples. Aunque las probabilidades pueden asignarse fácilmente en espacios muestrales sencillos, la regla de Laplace es difícil de aplicar directamente en casos más complejos, como obtener tres números iguales al lanzar 3 dados.

La regla de adición permite obtener este tipo de probabilidades. Consiste en calcular la probabilidad de un suceso compuesto calculando por separado la probabilidad de cada uno de los sucesos simples que la componen y luego sumando estas probabilidades. Otra aplicación de esta regla sería calcular la probabilidad del suceso contrario a uno dado. La regla de adición se generaliza posteriormente a conjuntos infinitos, por ejemplo en las distribuciones de probabilidad geométrica o en las distribuciones continuas. Vemos por tanto que esta idea puede aprenderse a diversos niveles de profundidad y que el aprendizaje temprano para casos finitos prepara la comprensión del caso general.

- **Independencia y regla del producto**

La característica de la Matemática en general y de la estocástica en particular de composición y descomposición de modelos, se distingue más claramente si los mismos experimentos aleatorios se combinan entre sí y se asignan probabilidades a los experimentos compuestos. El producto cartesiano permite construir el espacio muestral del experimento compuesto. De nuevo vemos como a partir de la idea de experimento simple pasamos a la de experimento compuesto y posteriormente a la idea de proceso estocástico en tiempo discreto (más tarde al proceso estocástico en tiempo continuo).

En todos estos modelos es fundamental, en primer lugar, el concepto de probabilidad condicional como medida de como cambia nuestro grado de creencia con la nueva información. Esta idea de probabilidad condicional será utilizada con mucha frecuencia en inferencia y otros métodos estadísticos, como en el estudio de la asociación entre variables.

Más importante aún en la definición de los experimentos compuestos y en el cálculo de probabilidades es la idea de independencia. Aunque matemáticamente la idea de independencia puede deducirse de la regla del producto de probabilidades y se relaciona con la de probabilidad condicional, desde un punto de vista didáctico merece un análisis independiente. Por un lado, mientras que la definición de independencia de sucesos de un mismo experimento se entiende con facilidad cuando se usa la regla de producto, es más complejo de explicar cuando se trata de independencia de sucesos de distintos experimentos, porque en este caso, tenemos que identificar con claridad que entendemos por la intersección de los sucesos intervinientes en la fórmula del producto.

Por otro lado, la idea de que un experimento aleatorio se puede repetir en las mismas condiciones y que los resultados de cada experimento son independientes es un buen ejemplo de la diferencia que hay entre entender un modelo teórico y saber aplicarlo en una situación concreta. Por un lado, el modelo matemático de la independencia, expresado por la regla del producto es fácil de entender. Por otro, ocurre que en la vida cotidiana mucha gente, incluso científicos, no son capaces de aplicar esta idea consecuentemente en las situaciones prácticas.

Actividad 2.14. Monedas dependientes. Una bolsa contiene 7 monedas de 100p, 50p, 50p, 50p, 10p, 10p, 10p. Sacamos dos monedas al azar. ¿Cuál es el valor esperado de su suma? ¿Depende este valor esperado de si la primera moneda es o no reemplazada? ¿Por qué?

Actividad 2.15. Paradoja de Blythe. Tenemos tres ruletas: La primera siempre da como resultado el número 3. La segunda da como resultado 2 con probabilidad 0.51, 4 con probabilidad 0.29 y 6 con probabilidad 0.20. La tercera da como resultados 1 con probabilidad 0.52 y 5 con probabilidad 0.48. Si cada uno de dos jugadores tiene que elegir una ruleta, y gana el que obtenga el número mayor, ¿cual es la mejor elección para el primer jugador? ¿Cambie esta elección si son tres los jugadores?

- **Equidistribución y simetría**

Las cuatro primeras ideas fundamentales no dan reglas prácticas de cómo calcular las probabilidades. Es una idea estratégica descubrir y usar las simetrías físicas o de otro tipo en las situaciones problemáticas, para decidir que ninguno de los resultados posibles tiene mayor ventaja que el resto y que, por lo tanto, podemos asignarles la misma probabilidad. Una vez que se acepta esta conclusión, a partir de los axiomas se llega con facilidad al cálculo de las probabilidades de los sucesos elementales en los espacios muestrales finitos.

Por ejemplo, al lanzar un dado, la simetría supone que ninguna cara se distingue de las demás. Esto es tomado como argumento para aceptar la igualdad de probabilidad de cada resultados y llegar a la regla de Laplace, que nos permite asignar una probabilidad de 1/6 a cada uno de los resultados posibles. Una vez calculadas estas probabilidades elementales, podremos calcular la probabilidad de sucesos más complejos como obtener un número par o obtener una suma par al lanzar 2 o 3 dados.

Hay que recalcar que la equidistribución (igualdad de probabilidad) de los sucesos elementales de un experimento no puede ser separada de la simetría estadística, es decir, la simetría confirmada por los registros estadísticos de resultados del experimento. El que un dado u otro dispositivo generador de resultados aleatorios cumpla las condiciones de simetría no es un hecho que pueda deducirse de la teoría matemática, sino de la experiencia. De hecho, se han establecido algunos principios, como que la simetría física implica la simetría estadística, aunque este principio es insuficiente, puesto que, aunque el dado esté bien construido, podría haber un sesgo en el jugador que lo lanza. En muchas situaciones, es la hipótesis más adecuada, pero sólo puede ser contrastada a posteriori, por medio de la adecuación del modelo.

Parece ser que la idea de simetría es difícil de enseñar a los niños, por éste motivo y porque los niños tienen creencias sobre que algunos resultados son más fáciles que otros, a pesar de la simetría física. Sólo con el trabajo repetido de ejemplos de diversos materiales simétricos y no simétricos se irá desarrollando esta idea.

- **Combinatoria**

El utilizar técnicas de recuento para calcular el número de elementos favorables y desfavorables a un suceso y usar estos números para calcular las probabilidades es otra idea fundamental, sobre todo en el cálculo de probabilidades complejas. Sin embargo, es demasiado simple considerar la combinatoria tan sólo como auxiliar de la probabilidad, como puede parecer a la vista de su estructura matemática. La extracción al azar de una urna de tres objetos entre cuatro posibles es un experimento aleatorio en tres fases, que puede ser interpretado significativamente en el espacio muestral de las variaciones $V_{4,3}$. Además las operaciones combinatorias pueden definirse, mediante experimentos aleatorios (extracción con o sin reemplazamiento, ordenada o no ordenada). Esta conexión entre los experimentos compuestos y la combinatoria se clarifica con el uso de diagramas en árbol.

El diagrama en árbol es una representación icónica fundamental, porque visualiza la estructura multi-paso del experimento compuesto. Por ello, las operaciones combinatorias, más que ser algoritmos de cálculo de probabilidades en espacios probabilísticos complejos, proporcionan una interpretación clara de la estructura interior de los experimentos y el encadenamiento de sucesivos experimentos en un complejo mayor.

Este punto de vista es soportado por los resultados de psicología del desarrollo, en particular por los trabajos de Piaget, quien sostiene que el camino de la comprensión de los fenómenos de azar pasa por el de las operaciones combinatorias básicas. Además las operaciones combinatorias son un componente fundamental de pensamiento formal que opera mediante combinaciones de las posibilidades que descubre. Finalmente señalamos la presencia de las operaciones combinatorias en las definiciones de las distribuciones de probabilidad discreta, así como en el estudio de los procesos estocásticos discretos.

Actividad 2.16. Las operaciones combinatorias se pueden definir mediante un modelo de muestreo o selección. Analizar los tipos diferentes de problemas que podemos tener según la selección sea o no ordenada y si se permite el reemplazamiento. Proponer un problema combinatorio de cada uno de estos tipos.

Actividad 2.17. Otros problemas combinatorios se refieren a la colocación de una serie de n objetos en m celdas, como en el ítem siguiente:

Disponemos de tres cartas iguales. Deseamos colocarlas en cuatro sobres de diferentes colores: amarillo, blanco, crema y dorado. Si cada sobre sólo puede contener, a lo sumo, una carta. ¿De cuántas formas podemos colocar las tres cartas en los cuatro sobres diferentes? Ejemplo: podemos colocar una carta en el sobre amarillo, otra en el blanco y otra en el crema.

Analizar los diferentes tipos de problema que podemos generar en el modelo de colocación, proponiendo ejemplo de enunciados. ¿Es cada problema de colocación directamente traducible a *un problema de selección?*

Actividad 2. 18. Otro tipo de problema combinatorio es el de partición de un conjunto en partes. Por ejemplo, ¿De cuantas formas diferentes puede descomponerse el número 5 en sumandos? Analizar los diferentes tipos de problemas de partición y estudiar su posible equivalente con los problemas de selección y colocación.

- **Modelos de urnas y simulación**

La palabra clave "simulación" en estadística significa algo parecido al isomorfismo en otras ramas de las matemáticas. Consiste en poner en correspondencia dos experimentos aleatorios diferentes. La condición es que a cada suceso elemental del primer experimento le corresponda un suceso elemental del segundo y sólo uno, de forma que los sucesos puestos en correspondencia en ambos experimentos sean equiprobables. Un ejemplo sería "simular" el experimento aleatorio consistente en observar el sexo de un recién nacido por el experimento aleatorio consistente en lanzar una moneda al aire.

Lo importante de la simulación es que podemos operar y observar resultados del segundo experimento y utilizarlos para obtener información del primero. Por ejemplo, si queremos saber cual es la probabilidad que entre 100 recién nacidos hay más de un 60% de varones, podemos lanzar, por ejemplo 1000 veces 100 monedas al aire, estudiar en cada uno de los 1000 experimentos si hubo o no más de un 60% de nacimientos y obtener una estimación para la probabilidad pedida. La ventaja de la simulación es obvia, incluso en este ejemplo tan sencillo, pues permite condensar el experimento en un tiempo y espacio concreto.

Otro hecho importante es que, en principio es posible asignar un modelo de urnas (el experimento consistente en extraer al azar una bola de una urna con una cierta composición de bolas de colores) a la mayor parte de experimentos aleatorios, al menos a aquellos con espacio muestral numerable. Además, no sólo el experimento original pueden componerse para formar nuevos experimentos, sino también los modelos de urnas asociados. Esto permite, como ha mostrado Polya, simular procesos aleatorios complejos mediante una secuencia de extracciones de bolas en urnas.

Por otro lado, la idea de la urna es fundamental, ya que en estocástica hay unos pocos conceptos firmemente enraizados que desafían toda definición rigurosa, como "muestreo al azar", que sólo pueden ser descritos concretizándolos en un modelo de urnas.

- **La variable aleatoria**

La idea de variable aleatoria ha sido responsable de las múltiples aplicaciones actuales del cálculo de probabilidades, puesto que el cálculo de probabilidades pasó de ocuparse del estudio de la probabilidad de sucesos aislados al estudio de las distribuciones de probabilidad (y posteriormente al de los procesos estocásticos). La variable aleatoria y su distribución, así como el estudio de las familias de distribuciones y sus propiedades son una herramienta muy potente, porque permite trabajar con el aparato del análisis matemático. Hay seria razones que justifican por qué los matemáticos pasados, que no conocían esta idea, tuvieron serios problemas con diversas paradojas matemáticas y, por ejemplo, Bernouilli necesitara 20 años para descubrir y probar su ley débil de los grandes números, temas que hoy se explica en unas pocas líneas.

Desde el punto de vista elemental, manejamos intuitivamente la idea de variable aleatoria cuando nos encontramos con juegos y experimentos en los que usamos dados, monedas, etc. Así mismo tenemos experiencias cotidianas con variables aleatorias continuas, como el tiempo de espera del autobús, o el necesario para llegar de nuestra casa al trabajo. Por el refuerzo de las múltiples, a veces inconscientes, experiencias con variables aleatorias en la vida cotidiana, la intuición de la magnitud aleatoria y de valor esperado, a veces aparece antes que la de probabilidad.

Por ejemplo, algunos psicólogos sostienen que la habilidad para estimar la esperanza matemática de las variables aleatorias es puramente biológica; mediante la experiencia llegamos a estimar el tiempo medio que tardaremos en preparar la comida o arreglarnos, lo que gastaremos en promedio al hacer la compra. En un plano formal, la esperanza matemática se interpreta como la media aritmética de los valores de una variable aleatoria, si el experimento se repitiese suficientemente en idéntica condiciones.

En el modelo de variable aleatoria hay tres conceptos básicos: su distribución, media y varianza. Mientras que la idea de media (esperanza matemática) es muy intuitiva, lo es menos la idea de distribución, especialmente cuando unos valores son más probables que otros. Para los niños pequeños, todas las variables aleatorias, tienen en principio distribución uniforme. Esto plantea el dilema de si debería restringirse la enseñanza o no a este tipo de variable.

Sea cual sea la respuesta, no debe negarse el hecho de que la distribución normal tenga una parte fundamental para explicar el mundo que nos rodea, ya que la encontramos en muchos fenómenos físicos, psicológicos o de otro tipo. Un modelo matemático convincente para explicar la presencia de la distribución normal en tantos fenómenos es el teorema central del límite que, por otra parte, no es accesible en su forma deductiva en ningún nivel por debajo de la universidad. Esto no excluye que sea interesante con los alumnos de secundaria considerar alguna aproximación intuitiva y experimental-inductiva, por ejemplo, usando el aparato de Galton u otro dispositivo de simulación.

- **Las leyes de los grandes números**

La convergencia estocástica hace posible el estudio de los fenómenos aleatorios en su conjunto, ya que individualmente son impredecibles. Para analizar la dificultad de comprensión de la convergencia, hay que distinguir entre las leyes empíricas de los grandes números (la que se observa al recoger datos estadísticos sobre un cierto fenómeno) y las correspondientes leyes matemática deducidas en forma de teoremas por diferentes probabilistas y que pueden ser demostradas formalmente.

La convergencia empírica es observable en la realidad, por ejemplo, al observar las gotas de lluvia sobre un pavimento, o la proporción de recién nacidos varones en un hospital a lo largo del año. Es filosóficamente interesante que una regularidad global surja de la variabilidad local, que parece inherente al curso de la naturaleza "libertad individual bajo restricciones colectivas". Esta convergencia empírica hace que las correspondientes leyes matemáticas de los grandes números se justifiquen como un buen modelo para los fenómenos aleatorios, aunque no contesta la pregunta de si es posible que los alumnos sean capaces de diferenciar entre el modelo y realidad, ya que de hecho vemos que con frecuencia se espera una convergencia empírica demasiado rápida o demasiado exacta.

Las oportunidades didácticas de experiencias empíricas sobre la convergencia serían deseables desde la escuela, para preparar la comprensión posterior de los teoremas matemáticos.

Sin embargo, estas oportunidades son más restringidas de lo que nos dicen los textos escolares. Las sucesiones aleatorias obtenidas en clase convergen lentamente y a veces fallan, cuando se precisan para una demostración, lo que puede ser contraproducente.

- **Muestreo**

La última idea fundamental es la de muestra que nos introduce en la inferencia y establece otro nuevo puente entre estadística y probabilidad. Esta idea es muy importante porque todo nuestro conocimiento y juicios sobre el mundo o las personas están basado en el muestreo. El conocimiento científico se adquiere a partir de las experiencias empíricas y estas son siempre limitadas, por lo que las conclusiones deben ser más amplia de los datos que obtenemos en las observaciones. La idea de muestra tiene en si dos características contradictorias: representatividad y variabilidad. La representatividad nos indica que la muestra se parece, en cierto modo a la población y debe ser una característica importante, ya que un prejuicio es sólo juicio basado en una muestra no representativa.

Por otro lado, la variabilidad indica que una muestra puede ser diferente de otra, por lo que al enjuiciar, pensar e inferir sólo es posible a base de muestras, la gente debería ser cauta y crítica al argumentar. Al igual que un estadístico profesional, todo el mundo debiera considerar el muestreo como modelo para explicar la realidad y entender la naturaleza estadística de sus conclusiones en cada caso particular, y las consecuencias de una decisión equivocada.

2.5. Análisis exploratorio de datos

Hasta los comienzos del siglo XX la estadística se restringía a la estadística descriptiva, que, a pesar de sus limitaciones, hizo grandes aportaciones al desarrollo de las ciencias experimentales. A partir de esa época, comenzaría la inferencia estadística clásica, con los trabajos de Fisher, Pearson y sus colaboradores y progresivamente se incorporaría la aportación de la escuela bayesiana.

Cabriá (1994) señala que los avances del cálculo de probabilidades llevaron a la creación de la estadística teórica, que en cierto modo, se alejó de las ideas estadísticas primitivas centradas en el análisis y recogida de datos. De este modo, en los años 60, la mayor parte de los libros de texto se ocupaban especialmente de los modelos inferenciales clásicos o bayesianos con respecto a conjunto simple de datos y hubo una tendencia a la matematización, junto con un descuido en la enseñanza de los aspectos prácticos del análisis de datos.

Con el desarrollo espectacular de la informática en la segunda mitad del siglo XX y la posibilidad de manejar rápidamente grandes masas de datos, se produjo, por un lado, una reacción ante tanta matematización, y por otro, disminuyó la importancia de los estudios muestrales.

Puesto que era fácil analizar grandes muestras ya no había por qué limitarse a los métodos estadísticos basados en distribuciones conocidas, cuya principal aplicación eran las pequeñas muestras. Tampoco había por qué restringirse a analizar una o unas pocas variables, porque el tiempo de cálculo se había eliminado y era preferible aprovechar toda la información disponible.

Como consecuencia, durante las últimas décadas se han desarrollado una serie de tipos de análisis de datos que se sitúan entre la estadística descriptiva y la inferencia o estadística teórica. Entre estos tipos se encuentran el análisis exploratorio de datos desarrollado por Tukey entre 1960 y 1980 y el análisis multivariante.

- **Análisis exploratorio de datos**

Las capacidades de cálculo y representación gráfica de los ordenadores actuales, permiten la obtención de una amplia variedad de gráficos y estadísticos diferentes de una forma sencilla y han hecho posible la aparición de una nueva filosofía en los estudios estadísticos: el análisis exploratorio de datos, introducido por Tukey, quien compara la labor del estadístico con la de un detective.

Anteriormente a este enfoque, el análisis de datos se basaba fundamentalmente en el cálculo de estadísticos (medias, varianza, coeficientes de correlación), conduciendo a dos consecuencias. En primer lugar se disminuía la importancia visual de la representación de los datos, dándosela exclusivamente a los cálculos y en segundo se pensaba que para obtener conclusiones de los datos era preciso recurrir a la inferencia (modelo confirmatorio).

En inferencia, el conjunto de valores observados se supone que se ajusta a un modelo preestablecido; por ejemplo, se supone que los datos se han obtenido de una población normal con media y desviación típica desconocidas. Debido a que las en muchas aplicaciones de la inferencia los datos son difíciles de recoger, se trabaja con pequeñas muestras y los métodos estadísticos correspondientes requieren la normalidad.

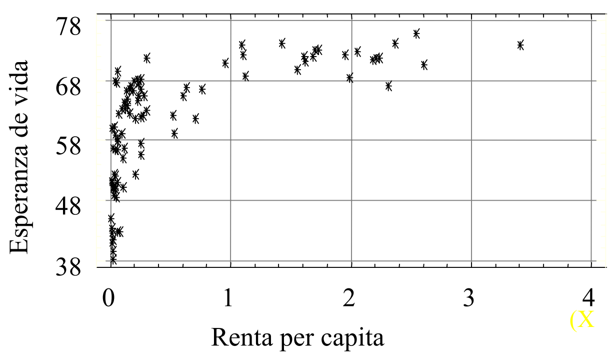
Partiendo de esta hipótesis, que es previa a la recogida de datos, se calculan los estadísticos (media y cuasivarianza de la muestra) para aceptar o no la hipótesis. En realidad los datos se recogen con el único propósito de poner tal hipótesis a prueba. Al contemplar solamente dos alternativas, (confirmación o no de la hipótesis), los datos no se exploraban para extraer cualquier otra información que pueda deducirse de los mismos. Por ejemplo, no tratamos de ver si la distribución es uniforme o en forma de U.

Para entender los principios por los que se guía el análisis exploratorio, se ha de tener en cuenta que los datos están constituidos por dos partes: la "regularidad" o tendencia y las "desviaciones" o variabilidad. Por regularidad entendemos la estructura simplificada de un conjunto de observaciones (en una nube de puntos, por ejemplo, es la recta a la cual se ajusta). Las diferencias de los datos con respecto a esta estructura (diferencia en nuestro caso respecto a la recta), representan las desviaciones o residuos de los datos, que usualmente no tienen por qué presentar una estructura determinada.

Tradicionalmente el estudio se ha concentrado en la búsqueda de un modelo que exprese la regularidad de las observaciones. Por el contrario, el análisis exploratorio de datos es básicamente el desglose de los mismos en las dos partes que hemos citado. En lugar de imponer, en hipótesis, un modelo a las observaciones, se genera dicho modelo desde las mismas.

__Ejemplo 2.1.__ Cuando se estudian las relaciones entre dos variables, el investigador no solamente necesita ajustar los puntos a una línea recta, sino que estudia los estadísticos, compara la línea con los residuos, estudia la significación estadística del coeficiente de correlación u otros parámetros para descubrir si la relación entre las variables se debe o no al azar.

Figura 2.1. Esperanza de vida en función de la renta per cápita



En la Figura 2.1. relacionamos la renta per cápita con la esperanza de vida en 97 países. Aunque los estadísticos calculados presenten un valor estadísticamente significativo (en el ejemplo, el coeficiente de correlación sea significativamente distinto de cero), la relación entre las variables puede no ajustarse bien a una línea recta. En este caso, si no se representasen gráficamente los datos al investigador le faltaría descubrir algo importante: el modelo latente no es el esperado.

Actividad 2.19. Analizar algunos de los tipos de gráficos propuestos por Tukey para el análisis exploratorio y comparar con otros gráficos clásicos, como el histograma de frecuencias o la curva acumulativa de frecuencias. ¿Qué características de los primeros les dan una potencia exploratoria?

Actividad 2.20. Analizar las actividades recomendadas en los nuevos currículos de enseñanza secundaria sobre análisis exploratorio de datos. ¿Cuáles dificultades cognitivas y didácticas pueden preverse para implantar estas actividades?

Actividad 2.21. Al análisis de datos se basa en el método de elaboración de proyectos por parte de los estudiantes. Describe algunos proyectos sencillos en los que los alumnos de secundaria puedan recoger datos significativos y apropiados para el aprendizaje de conceptos elementales de análisis de datos.

- **Características educativas del análisis exploratorio de datos**

Como hemos indicado, nos encontramos ante una nueva filosofía en la aplicación de los métodos de análisis de datos, aunque unida a ella se han desarrollado también algunas técnicas concretas para su aplicación, como el gráfico de la caja o el de tallo y hojas. Esta filosofía consiste en el estudio de los datos desde todas las perspectivas y con todas las herramientas posibles, incluso las ya existentes. El propósito es extraer cuanta información sea posible, generar "hipótesis" nuevas, en el sentido de conjeturar sobre las observaciones de las que disponemos.

Al considerar la conveniencia o no de incluir un tema como objeto de enseñanza hemos de tener en cuenta su utilidad y que este tema se halle al alcance de los alumnos. Además de la utilidad, ya razonada, el análisis exploratorio de datos tiene las siguientes características que lo hacen un tema apropiado de estudio:

- *Posibilidad de generar situaciones de aprendizaje referidas a temas de interés al alumno: l*o usual es trabajar sobre un fichero de datos que han sido codificados previamente e introducidos en el ordenador, ya que se pretende estudiarlos mediante cuantas perspectivas y técnicas tengamos a nuestro alcance. Estos conjuntos de datos pueden ser obtenidos por los mismos estudiantes, mediante la realización de una encuesta a sus compañeros sobre temas diversos, como características físicas, aficiones, empleo del tiempo libre, etc., o a partir de datos obtenidos en anuarios, publicaciones estadísticas o Internet.
- *Fuerte apoyo en representaciones gráficas:* Una idea fundamental del análisis exploratorio de datos es que el uso de representaciones múltiples de los datos se convierte en un medio de desarrollar nuevos conocimientos y perspectivas. Por ejemplo, pasar de lista de números a representaciones como la del "tallo y hojas", puede facilitar la exploración de la estructura total, así como construyendo gráficos, como el de la "caja" que hace posible la comparación de varias muestras.
- *No necesita una teoría matemática compleja*: Como el análisis de datos no supone que estos se distribuyen según una ley de probabilidad clásica (frecuentemente la normal), no utiliza sino nociones matemáticas muy elementales y procedimientos gráficos fáciles de realizar.

2.6. Asociación

2.6.1. Importancia en estadistica

Un tema que aparece con frecuencia en el análisis datos, tanto exploratorio como confirmatorio, es la asociación o dependencia estadística entre variables, al que también están ligadas dificultades de tipo filosófico.

En general, las técnicas de regresión y correlación se preocupan de la interrelación entre dos o más variables continuas. De la extensión de la idea de correlación a variables cualesquiera, incluso no numéricas, surge el concepto general de asociación.

- **Regresión**

En el análisis de regresión estudiamos la relación entre una variable dependiente y otro conjunto de una o varias variables independientes. Esta relación se representa por un modelo matemático, dado por la ecuación de regresión, a la que se añaden un conjunto de hipótesis básicas. La ecuación de regresión incluye un conjunto de parámetros desconocidos.

En particular, si la ecuación de regresión es lineal respecto a estos parámetros (no necesariamente respecto a las variables independientes) hablamos de regresión lineal. Los modelos lineales para variables no numéricas incluyen el análisis de la varianza y los modelos logarítmico lineales. Entre los tipo de regresión no lineal podemos citar la regresión logística y la regresión polinómica. Hay dos posibles razones para efectuar un análisis de regresión:

- Se desea obtener una descripción de la relación entre las variables, como una indicación de una posible causalidad.
- Se quiere predecir la variable dependiente, a partir de los valores de las variables independientes, lo cual es muy útil si la variable dependiente es costosa o difícil de medir.

Los problemas estadísticos implicados en el análisis de regresión son:

- Obtener un buen estimador de los parámetros del modelo;
- Contrastar hipótesis sobre dichos parámetros;
- Determinar la bondad del modelo para los datos particulares
- Comprobar que se cumplen las hipótesis exigidas.

- **Asociación y Correlación**

Dos variables están asociadas en la extensión en que una de ellas es un buen predictor de la otra. La medida de la intensidad de la relación entre dos variables, mediante un coeficiente adecuado, constituye el problema de *correlación* (en el caso numérico) y *asociación,* en el caso general. Además del coeficiente de correlación lineal, que mide la intensidad de la relación lineal entre dos variables numéricas, existen otros: el coeficiente de correlación múltiple, parcial, así como las diferentes medidas de asociación para variables ordinales y nominales. Se plantean problemas estadísticos semejantes a los descritos para la regresión.

- **Papel de la asociación y regresión en los métodos estadísticos**

Las ideas de asociación y regresión aparecen en muchos temas estadísticos. Los modelos lineales de regresión univariantes se extienden a los modelos multivariantes, MANOVA, MANCOVA y regresión multivariante, en el caso de múltiples variables dependientes. Esta no es, sin embargo, la única forma de extensión posible. En los casos mencionados, las ecuaciones obtenidas expresan la variable dependiente (o un vector formado por las variables dependientes) como combinación lineal de las variables independientes. Otras posibilidades se presentan con ligeras variantes sobre estos modelos, que, sin embargo, tienen consecuencias metodológicas muy importantes en la investigación.

Supongamos que no es necesario distinguir entre variables dependientes e independientes, sino que lo que se trata es de reexpresar cada variable en función del resto. Ello puede interpretarse como obtener un cambio de base en un espacio vectorial en el que los ejes primitivos de coordenadas son las variables observadas y los nuevos ejes combinaciones lineales de las mismas.

Si se combina esta idea con la de que el nuevo sistema de ejes esté formado por un conjunto de variables incorrelacionadas y que, además, se obtengan en un orden decreciente en cuanto a la varianza explicada, se obtiene el modelo de componentes principales.

Este modelo guarda una gran relación con la distribución normal multivariante, ya que los ejes obtenidos con las hipótesis anteriores son los ejes de los elipsoides de inercia de dicha distribución. En el caso de sólo dos variables se obtiene la recta de regresión lineal, por lo que el análisis de componentes principales puede considerarse una extensión de la regresión.

La idea de componentes principales, obtenidos en orden decreciente de varianza, que pueden usarse como nuevas coordenadas para los datos, lleva a la idea de reducción de la dimensión y representación de los datos multivariantes en un número pequeño de factores.

De ahí surge primitivamente el análisis factorial y posteriormente las técnicas derivadas de él, hasta llegar al análisis de correspondencias, que extiende el análisis factorial al estudio de las tablas de contingencia de variables cualitativas dobles o múltiples.

Otra posibilidad de extensión es la de suponer que no existe una diferenciación estricta entre variables dependientes e independientes, sino que las diversas variables están ligadas entre sí formando redes causales donde unas variables dependen de otras formando un sistema de interrelaciones múltiples. De este modo se llega a los modelos de estructuras covariantes y otros modelos de desarrollo muy reciente.

Actividad 2.22.¿Dónde aparece la idea de asociación en el currículo de secundaria? ¿Cuáles son las actividades básicas en este nivel, relacionadas con el tema?

Actividad 2.23. Analizar en un libro de texto de secundaria el tema de correlación. ¿Se discuten las relaciones entre correlación y causalidad? ¿Cómo podríamos presentar a los alumnos estas relaciones?

2.6.2. Asociacion y causalidad

Con frecuencia, estamos interesados en buscar explicaciones para los valores observados de las variables dependientes, en términos de las variables explicativas, es decir en buscar explicaciones de tipo "causal"

Numerosos filósofos y científicos han mantenido que "causalidad" y asociación están esencialmente relacionadas. Por ejemplo, Kendall y Lazarsfeld, Campbell y Stanley, o Blalock, consideran que si una variable A es "causa" de otra variable B entonces A está correlacionada positivamente con B.

Otros autores tales como Asher, consideran que si A es "causa" de B entonces el coeficiente de correlación puede servir confirmar el grado de la influencia "causal" de A en B. Esto justifica la práctica de rechazar las hipótesis "causales" cuando A no está correlacionado positivamente con B.

No entraremos aquí en un análisis completo de la noción de causa a la que se ha atribuido diferentes significados en filosofía y metodología de investigación (algunos de estos significados son analizados por Estepa, 1994).

Lo que haremos es precisar el punto de vista "estadístico", base de los modelos causales-estructurales (Bollen, 1989). Estos modelos estructurales extienden las ideas de correlación y regresión en el sentido de admitir en el estudio redes complejas de variables interconectadas, en lugar de efectuar una separación estricta de las variables en dependientes e independientes.

Actividad 2.24. ¿Son siempre las relaciones de tipo funcional relaciones causales? Pon un ejemplo de relación causal de tipo aleatorio.¿En que parte del proceso de estudio de relaciones causales interviene la asociación estadística?

- **Definición de causalidad, desde el punto de vista estadístico**

Supongamos una variable y_1 aislada de toda influencia salvo de la de otra variable x_1. Si un cambio en x_1 puede producir otro cambio (no siempre lo produce) en y_1, entonces, se admite que x_1 es *causa débil* de y_1.

Ejemplo 2.2. El vacunarse contra algunas enfermedades generalmente (aunque hay algunas excepciones) confiere una inmunidad o al menos resistencia antes la enfermedad, por lo que podemos decir que la vacuna es causa de inmunidad. No se exige que la variable causa tenga que ser manipulada, por ejemplo, podemos admitir que la posición relativa de la luna causa las mareas.

Davis (1985) señala tres características de la definición de causación débil entre variables:

- Hablamos de la *tendencia* de x_1 a producir un cambio en y_1, por lo cual se esperan excepciones individuales.
- Decimos "*x_1 es una causa de y*" y no "x_1 es la causa de y_1". Puede haber más de una causa para y_1. En el ejemplo de las vacunas, la inmunidad puede también depender de características físicas del sujeto.
- *La existencia de correlación por sí sola no implica caus*ación. Hay variables correlacionadas, como la renta per cápita y la tasa de natalidad (negativamente). Pero esto no quiere decir que una sea causa de la otra, sino que la correlación es debida a la existencia de terceras variables correlacionadas con las anteriores.

La definición anterior tiene tres componentes: *aislamiento, asociación y dirección* de la influencia.

- **Modelo determinista lineal**

La forma matemática más simple en que podría representarse este tipo de causalidad, sería por una expresión del tipo: $y_1 = f(x_1)$ es decir, una dependencia de tipo funcional. Esta expresión es más sencilla aún si toma una forma lineal, como al estudiar la relación entre el precio que debo pagar y_1 por una cierta cantidad de azúcar x_1 (g_{11} es el coste unitario):

(1) $$y_1 = \gamma_{11}x_1$$

En este caso, se produciría una asociación entre x_1 e y_1, puesto que, por cada unidad de cambio en x_1, se produciría exactamente g_{11} unidades de cambio en y_1.

Esta expresión matemática corresponde a la noción determinista de causalidad, ya que aparece una sola causa y un sólo efecto. Por otro, se manifiesta la hipótesis de Hume de "conjunción constante": cada vez que ocurre la causa se produce el efecto (cada vez que compramos patatas pagamos un precio).

- **Modelo aleatorio lineal**

Un modelo más razonable, que incluiría el componente estocástico, sería el siguiente, que serviría para representar, por ejemplo, la relación entre el peso y_1, y la altura $x_1$de una persona; g_{11} sería el coeficiente de la ecuación de regresión.

(2) $$y_1 = \gamma_{11}x_1 + \xi_1$$

El introducir este nuevo componente presenta ya una seria implicación sobre el concepto de causalidad. Se supone que ξ_1 es un error aleatorio y no podemos controlarlo directamente, por lo que para cada caso particular no podemos evaluar su valor. El peso de una persona depende de su talla, pero hay personas de la misma altura con pesos diferentes. Suponemos que el peso depende de la altura más un factor aleatorio.

Al admitir un modelo del tipo anterior, es importante la hipótesis de *aislamiento*, que consiste en suponer que el error aleatorio no depende del resto de las variables. Esta hipótesis, usual en los modelos lineales (regresión y análisis de varianza), nos permite evaluar la influencia de x_1 sobre y_1 aisladamente de ξ_1, es decir, estimar el coeficiente de regresión.

La hipótesis de aislamiento se cumple cuando los datos se han tomado aleatoriamente, y han sido medidos todos con el mismo instrumento, pero puede no cumplirse en otros casos.

Ejemplo 2.3. Si al estudiar la relación entre peso y altura, los hombres jóvenes han sido pesados en una báscula incorrecta, el error de medición para los hombres jóvenes podría ser mayor que para el resto.

En el modelo aleatorio, por otro lado, desaparece la condición de Hume y el efecto ya no es perfectamente predecible a partir de la causa. Además, incluso el modelo (2) es demasiado simple. Es verosímil que en un fenómeno real haya muchas causas de y_1, además de x_1. Por ejemplo, si consideramos el peso de los sujetos este depende también de la alimentación, raza, sexo, práctica de deporte, constitución, etc. Un modelo del tipo (3) es mas realista, lo que supone que la correlación de cada variable explicativa con la independiente no será excesivamente alta y que más que de una sola causa habría que hablar de multiplicidad de causas o conjunción de causas.

(3) $$Y_1 = \gamma_{11}x_1 + \gamma_{12}x_2 + ... + \gamma_{1q}x_q + \xi_1$$

Actividad 2.25. Analiza qué tipos de relaciones entre variables pueden originar la existencia de correlación y cuáles de ellos son de naturaleza causal. Pon ejemplos de relaciones causales que den origen a un bajo coeficiente de correlación.

Actividad 2.26. En el Proyecto 3 (Capítulo 5) relaciona la esperanza de vida con otras variables. Observa los gráficos y analiza:
¿Cuál de las variables influyen en la esperanza de vida del hombre?
¿Cuáles son influídas por la esperanza de vida?
¿Cuáles hacen crecer / disminuir la esperanza de vida?
¿Cuál de ella sirve mejor para predecir la esperanza de vida?
¿En qué casos la relación es causal?¿Es en algún caso debida a otras variables?
¿Podrías en alguno de los casos hallar una función matemática para predecir, aproximadamente la esperanza de vida del hombre a partir de la otra variable?

- **Aislamiento**

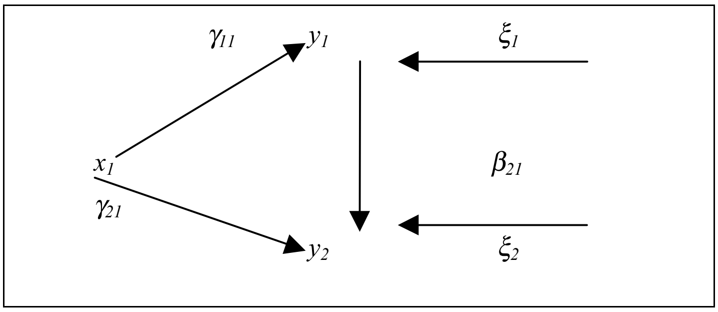
En los modelos (2) y (3) se supone que el término de error es incorrelacionado con las variables $x_1, ..., x_q$. Estudiemos el efecto de violar esta hipótesis. Una situación común viene dada cuando una variable interviniente no se controla y su efecto es, como consecuencia, parte del error aleatorio ξ. En este caso se produce una correlación entre x y ξ.

Ejemplo 2.3. Consideremos el modelo dado por las dos ecuaciones:

$y_1 = \gamma_{11}x_1 + \xi_1$

$y_2 = \gamma_{21}y_1 + \gamma_{21}x_1 + \xi_2$

en donde los errores ξ_1 y ξ_2 están incorrelacionados entre sí y con x_1 . El diagrama de la figura muestra la situación.



Si, por error, un investigador omite la variable interviniente en la segunda ecuación y usa el modelo:

$$y_2 = \gamma_{21}x_1 + \xi^*_2$$

donde,

$$\xi^*_2 = \beta_{21}y_1 + \xi_2$$

se puede demostrar que al estimar el efecto directo de x_1 sobre y_2, γ_{21}, obtendremos en su lugar una estimación del efecto total $\gamma^*_{21} = \gamma_{21} + \beta_{21}\gamma_{11}$.

Si los efectos directo e indirecto tienen magnitudes similares pero distinto signo, podremos obtener una estimación del efecto total mucho menor que el de cualquiera de los efectos mencionados. Otro problema surge cuando olvidamos incluir causas comunes de las variables explicativa y dependiente. Supongamos que en el modelo especificado, olvidamos por error la consideración de x_1.

$$y_2 = \beta_{21}y_1 + \xi^*_2 \text{ donde, } \xi^*_2 = \gamma_{21}x_1 + \xi_2$$

Puede mostrarse que al tratar de estimar β_{21} obtenemos en su lugar $\beta^*_{21} = \beta_{21} + \gamma_{21}b$ en donde b es el coeficiente de regresión de x_1 sobre y_1 .

Actividad 2.27. Analizar las distintas situaciones que pueden presentarse según los valores tomados por los diferentes coeficiente se resumen en la tabla 1.

Tabla 1 Relación entre los valores de $b^*_{21} = (b_{21} + \gamma_{21}b)$ y b_{21} cuando se omite la causa común

β_{21}	$\gamma_{21}b$	β^*_{21} **versus** β_{21}	***Relación entre y_1 e y_2***
0	*0*	*igual*	*Ninguna*
0	*0*	*mayor*	*Totalmente espúrea*
> 0	*> 0*	*mayor*	*Parcialmente espúrea*
< 0	*< 0*	*mayor*	*Parcialmente espúrea*
> 0	*< 0*	*menor*	*No se advierte*
< 0	*> 0*	*menor*	*No se advierte*

Actividad 2.28. Analizar las siguientes medidas de asociación entre variables cualitativas, desde el punto de vista de su equivalencia por la información que proporcionan sobre la existencia de una relación causal entre las variables intervinientes (Damos sólo la definición de asociación directa).

- *Dos variables A y B se correlacionan positivamente si y solamente si la probabilidad de que ocurran simultáneamente A y B es mayor que el producto de las probabilidades de A y B (Kendall, Lazarsfeld y Nagel):*

$$P(A \bullet B) - P(A)P(B) > 0$$

- *Dos variables están positivamente correlacionadas si y solamente si la probabilidad de B condicionada a A menos la probabilidad de B es mayor que cero (Reinchebach y Suppes). Esto es, A y B están correlacionadas positivamente si y solamente si :*

$$P(B/A) - P(B) > 0$$

- *Dos variables están positivamente correlacionadas si y solamente si la probabilidad de B condicionada a A menos la probabilidad de B condicionado a no A es mayor que cero (Salmon y Suppes). Esto es, A y B están correlacionadas positivamente si y solamente si:*

$$P(B/A) - P(B/A) > 0$$

2.7. Inferencia estadística

Pearson fue uno de los pioneros en establecer un puente entre la estadística descriptiva y la probabilidad. Usó los resúmenes de los datos (estadísticos como la media, varianza y el coeficiente de correlación), para extraer inferencias acerca de las distribuciones básicas, y creó el test de Chi cuadrado de adherencia de ajuste. Los trabajos de Fisher entre 1920 y 1930 aportan una gran cantidad de ideas nuevas, sobre pequeñas muestras, inspirando la teoría de los test de hipótesis. Asimismo creó los cimientos del diseño de experimentos y análisis de la varianza. La teoría de los test de hipótesis es perfeccionada por Neyman y Pearson, construyendo hacia 1928 una teoría sistemática de los mismos, que sería la base para la teoría de decisión creada por Wald hacia 1950.

Las ciencias empíricas, en general, y en particular la psicología y la educación, dependen en gran medida de la demostración de la existencia de efectos a partir del análisis estadístico de datos para el avance de su investigación. Sin embargo, debido a que la lógica de la inferencia estadística es difícil, su uso e interpretación no es siempre adecuado y han sido criticados en los últimos 50 años. Yates (1951), por ejemplo, sugirió que los científicos dedicaban demasiada atención al resultado de sus contrastes, olvidando la estimación de la magnitud de los efectos que investigaban. Una amplia revisión de estas críticas puede encontrarse en Morrison y Henkel (1979).

Esta controversia se ha intensificado en los últimos años en algunas instituciones profesionales (Menon, 1993; Thompson, 1996; Ellerton, 1996, Robinson y Levin, 1997, 1999 Levin, 1998 a y b, Wilkinson et al., 1999) que sugieren importantes cambios en sus políticas editoriales respecto al uso del contraste de hipótesis. Por ejemplo, la American Psychological Association resalta en su manual de publicación del año 1994 que los contrastes estadísticos no reflejan la importancia o la magnitud de los efectos y animan a los investigadores a proporcionar información sobre el tamaño de estos efectos (APA, 1994, pg. 18).

Más recientemente, la Task Force on Statistical Inference organizada por la APA ha publicado un artículo para iniciar la discusión en el campo, antes de revisar el manual de publicación de la APA (Wilkinson, 1999). Una decisión de este comité ha sido que la revisión cubra cuestiones metodológicas más generales y no sólo el contraste de hipótesis. Entre otras cuestiones, se recomienda publicar los *valores-p* exactos, las estimaciones de los efectos y los intervalos de confianza.

En la American Education Research Association, Thompson (1996) recomienda un uso más adecuado del lenguaje estadístico en los informes de investigación, enfatizando la interpretación del tamaño de los efectos y evaluando la replicabilidad de los resultados. Estas instituciones, así como la American Psychological Society han constituido comités específicos para estudiar el problema, los cuales sugieren no abandonar el contraste de hipótesis, sino complementarlo con otros análisis estadísticos (Levin, 1998 b, Wilkinson et al., 1999). Un resumen comprensivo de estos debates, así como de las alternativas sugeridas, se presenta en Harlow, Mulaik y Steiger (1997).

A pesar de estas recomendaciones, los investigadores experimentales persisten en apoyarse demasiado en la significación estadística, sin tener en cuenta los argumentos de que los tests estadísticos por si solos no justifican suficientemente el conocimiento científico. Algunas explicaciones de esta persistencia incluyen la inercia, confusión conceptual, falta de mejores instrumentos alternativos o mecanismos psicológicos, como la generalización inadecuada del razonamiento en lógica deductiva al razonamiento en la inferencia bajo incertidumbre (Falk y Greenbaum, 1995).

Podríamos pensar que esta situación se debe a una enseñanza insuficiente del tema, a pesar de que la estadística es una asignatura obligatoria en la mayor parte de las licenciaturas e ingenierías. Sin embargo, a veces estos errores se encuentran también en personas con fuerte preparación estadística (Morrison y Henkel, 1970). En esta sección estudiaremos algunas dificultades epistemológicas del tema que pueden explicar la dificultad de comprensión y aplicación de los métodos estadísticos inferenciales.

2.7.1. La justificación del razonamiento inductivo

La problemática filosófica de la inferencia estadística surge de la dificultad de justificar el razonamiento inductivo empírico y, por tanto, del carácter provisional de las teorías científicas obtenidas a partir del mismo. Resumiremos esta problemática siguiendo a Rivadulla (1991).

En cualquier ciencia experimental hay un conjunto de proposiciones o cuerpo científico admitido como "verdadero". Un nuevo conocimiento se produce mediante un proceso de inferencia, es decir mediante el paso de una verdad ya admitida (premisa) a un nuevo enunciado (conclusión). En este proceso se usan dos tipos de razonamientos básicos: *deductivo e inductivo.* En el razonamiento deductivo la conclusión está contenida en las premisas, por lo que, si partimos de un conocimiento verdadero, obtenemos otro conocimiento también cierto.

Ejemplo 2.4. Partiendo de que la suma de los ángulos de un triángulo es 180º (premisa) podemos demostrar fácilmente (concluir) que en un triángulo rectángulo la suma de los dos ángulos opuestos a los catetos es 90º. Puesto que el triángulo rectángulo es un caso particular de triángulo, la conclusión está contenida en la premisa y se trata de un razonamiento deductivo, como la mayor parte de las demostraciones de los teoremas matemáticos.

Por el contrario en el razonamiento *inductivo* la conclusión es más amplia que las premisas, siendo por tanto posible la falsedad de la conclusión, aunque las premisas sean verdaderas. Este es el tipo de razonamiento que se usa con más frecuencia en las ciencias empíricas, lo que explica que, a veces, una teoría científica se admite como cierta durante un cierto periodo histórico y luego se abandona, cuando se encuentra una evidencia en contra. Ejemplos de teorías que se han descartado posteriormente serían la de que el sol giraba alrededor de la tierra, o de que el sida se contagia por el aire.

A pesar de que a veces es difícil de justificar, el método inductivo es con frecuencia el único posible en una cierta ciencia, o bien es necesario usar un método inductivo para complementar el método deductivo en muchas situaciones. Por otro lado, muchos descubrimientos científicos o incluso teoremas matemáticos han sido formulados, primero, en forma inductiva (a partir de la experiencia) y luego se han podido demostrar en forma deductiva; baste pensar en la "demostración" del teorema central del límite la hizo Galton en forma experimental utilizando su "aparato de Galton". En otros casos muchas regularidades en la naturaleza y muchas leyes empíricas son aceptadas por la comunidad científica, por la fuerza de su evidencia.

La búsqueda de criterios para dar validez al razonamiento inductivo es muy antigua, incluso preocupó a Aristóteles, aunque Hume fue el primero que planteó con claridad el problema de *invalidez de la inducción:* la incompatibilidad entre el principio fundamental del empirismo (todas las teorías científicas debieran obtenerse a partir de la observación de la realidad) y el de invalidez de la inducción (las conclusiones obtenidas a partir de la inducción no tienen validez lógica). Muchos filósofos han tratado de resolver este problema, sin que hasta la fecha se haya logrado una solución.

Actividad 2.29.Analiza la naturaleza de las demostraciones matemáticas por inducción completa. ¿Se usa un razonamiento inductivo o deductivo? Ponerlo de manifiesto mediante un ejemplo.

Actividad 2.30. Elaborar una lista de teorías científicas que han sido admitidas, antes de su demostración, a partir de un razonamiento inductivo y por la fuerza de su evidencia. Busca algún ejemplo dentro de teorías científicas que hayan sido admitida en un principio a partir de un razonamiento inductivo y posteriormente se haya mostrado su invalidez.

2.7.2. Popper y la refutación de hipótesis

Una gran influencia sobre el desarrollo de la inferencia estadística y el debate sobre el método inductivo ha sido la de Popper. Este autor razona que, aunque muchos casos particulares hayan confirmado una teoría, es posible que más adelante se encuentre un caso que pruebe que la teoría es falsa. Por ejemplo, bastó un sólo ejemplar vivo del celacanto para echar abajo la teoría de que esta especie se había extinguido.

Por ello Popper propone que en lugar de intentar demostrar inductivamente las teorías científicas, debemos tratar de refutarlas (ponerlas a prueba) mediante experimentos u observaciones, donde se comparen los patrones deducidos de la teoría los con datos obtenidos.

Un conocimiento inductivo sería lógicamente posible para este autor y podríamos considerar cierta una teoría, si, a pesar de nuestros intentos, no conseguimos refutarla. Si los datos se ajustan a lo esperado, la teoría recibe un apoyo provisional (corroboración), al menos que otra teoría conocida pueda dar una explicación mejor de este patrón. Pero esta corroboración nunca prueba que la teoría sea cierta, porque futuros datos podrían ser discordantes, de modo que la aceptación de una teoría es siempre provisional.

Sin embargo, si los datos no se ajustan a los patrones esperados se puede falsar (mostrar que no es cierta) la teoría sometida a prueba (Rivadulla, 1991). Vemos que el rechazo o aceptación de las teorías no tienen el mismo estatuto.

Una objeción a la propuesta de Popper es que los datos nunca se ajustan exactamente a lo que las teorías predicen y por ello, nunca pueden "confirmar" con exactitud las teorías, ya que los modelos teóricos son siempre aproximaciones de la realidad y nunca copias exactas de la misma. La comunidad científica, será la que decida un criterio para dividir los resultados posibles de los experimentos en aquellos que confirman la hipótesis, los que la falsan y los ambiguos. Otro problema es que para establecer una teoría científica habría que compararla con todas las hipótesis rivales plausibles, pero éstas no se suelen conocer de antemano.

Finalmente, Kuhn y Feyerabend sostienen que las teorías no siempre pueden ser comparadas entre sí, puesto que los datos no son objetivos al están impregnados por la teoría o paradigma bajo la que son recogidos.

Actividad 2.31. El caso del aceite de colza fue prototípico de la dificultad con que, a veces, se tropieza el científico para deducir factores causales en la investigación. ¿Puedes preguntar a tus padres o buscar en alguna hemeroteca las primeras explicaciones que se barajaron para explicar la enfermedad causada por este aceite tóxico?

Actividad 2.32. Describe las etapas que, a tu juicio, debe recoger una investigación empírica completa y señala en cuáles de ellas sería necesaria la colaboración del estadístico.

- **La búsqueda de una solución probabilística al problema de la inducción**

Una de las soluciones posibles del problema planteado por Hume es tratar de debilitar la conclusión del razonamiento inductivo, recurriendo a la probabilidad. Ya que no podemos obtener con garantías las certeza de la conclusión, tratemos de calcular la probabilidad de que sea cierta, es decir, admitamos la "verdad probable"

Popper diferencia la probabilidad de un suceso y la probabilidad de una hipótesis:

- La probabilidad de un suceso hace referencia a su frecuencia, porque el suceso unas veces ocurre y otras no.
- Es más difícil dar sentido a la probabilidad de una hipótesis, porque la hipótesis es o cierta o falsa. Se cree que no es posible obtener la probabilidad objetiva de una hipótesis.

También diferencia dos usos del concepto de probabilidad de una hipótesis:

- Probabilidad a priori, que no depende de los datos.
- Probabilidad a posteriori, es decir, en relación con los test a los que ha sido sometida esta hipótesis. Esta sería una probabilidad condicional, obtenida a partir de los datos empíricos.

¿Es posible que la estadística conceda a las hipótesis científicas una probabilidad, obtenida en forma inductiva a partir de datos experimentales? Analizaremos en lo que sigue la lógica de los test estadisticos para ver si puedn dar respuesta a esta pregunta

Actividad 2.33. Busca en un diccionario diferentes significados del término hipótesis. Compara con los que puedes obtener en un texto de estadística y en un libro de metodología de *investigación. Compara con el significado de las hipótesis en otras ramas de las matemáticas (por ejemplo la geometría). ¿Piensas que los alumnos pueden tener confusión debido a estos usos diferenciados?*

Actividad 2.34. Analiza en algún libro de introducción a la estadística en ciencias sociales el uso que se hace del término "probabilidad" en el tema de inferencia. ¿Se hace referencia a las probabilidades de hipótesis o se restringe este uso a los sucesos de experimentos aleatorios? ¿Piensas que en algunos casos el lenguaje empleado puede inducir confusión?

Actividad 2.36. En el contexto del diagnóstico médico, a veces se dice "la probabilidad de que tengas esta enfermedad" es x. Analiza el empleo del término probabilidad en esta frase. ¿Cuáles son los experimentos y poblaciones implicados? Busca otros ejemplos similares en diferentes contextos profesionales.

2.7.3. La Lógica de los Tests Estadísticos

Para comenzar la discusión presentaremos una situación típica en la que un investigador recurre a la estadística para dar apoyo a una hipótesis referida a su campo de estudio y resumimos los pasos y lógica del contraste de hipótesis.

Ejemplo 2.5. De acuerdo con algunas teorías de aprendizaje, las representaciones contribuyen a la construcción del significado de los objetos matemáticos, de modo que un contexto rico en representaciones, que facilite el cambio de una representación a otra, favorecerá el aprendizaje. Un investigador que acepta esta teoría tiene buenas razones para esperar que el uso de los ordenadores refuerce el aprendizaje de la estadística, porque los ordenadores proporcionan potentes herramientas y sistemas de representación de los conceptos estadísticos. Para probar su conjetura, supongamos que el investigador selecciona una muestra aleatoria de 80 estudiantes entre todos los alumnos que ingresan a la universidad un curso dado. Aleatoriamente divide los 80 estudiantes en dos grupos de igual tamaño y encuentra que los dos grupos tienen un conocimiento inicial equivalente de la estadística. Organiza un experimento, donde el mismo profesor, con los mismos materiales, imparte un curso introductorio de estadística a los dos grupos durante un semestre. El grupo C (grupo de control) no tiene acceso a los ordenadores, mientras que la enseñanza en el grupo E (grupo experimental) se basa en un uso intensivo de los ordenadores.

Al final del periodo, el mismo cuestionario se pasa a los dos grupos. Si el aprendizaje con los dos métodos de enseñanza fuese igual de efectivo, no debiera haber diferencia entre la puntuación media en el cuestionario μ_e, de la población teórica de todos los estudiantes a los que se enseña con ayuda del ordenador y la puntuación media en el cuestionario μ_c, de la población teórica de estudiantes que no tienen acceso a un ordenador. Si consideramos que los grupos C y E son muestras representativas de estas poblaciones teóricas, y no hay diferencia de efectividad entre los dos métodos de enseñanza, la diferencia en las puntuaciones medias del cuestionario en los dos grupos debería ser próxima a cero. Si el investigador encuentra una diferencia positiva entre las puntuaciones medias en los dos grupos $\overline{x}_e - \overline{x}_c$, ¿podría deducir que su conjetura era cierta? Es importante resaltar que el interés del investigador va más allá de las muestras particulares de alumnos (grupos C y E). Está interesado en comprobar el efecto del aprendizaje con ordenadores en la población, en general, y a través de ello, encontrar un apoyo empírico para su hipótesis sobre el efecto del contexto en el aprendizaje.

El ejemplo anterior ilustra una situación típica de uso de la inferencia estadística para determinar si los datos experimentales (las puntuaciones de los alumnos de los grupos E y C) apoyan o no una *hipótesis substantiva* (el aprendizaje está influenciado por los sistemas de representación disponibles). Como no podemos probar directamente la hipótesis substantiva, organizamos un experimento para obtener datos y contrastar una *hipótesis de investigación* deducida de la anterior (los ordenadores refuerzan el aprendizaje de la estadística). Tampoco podemos confirmar directamente la hipótesis de investigación, porque el aprendizaje es un constructo inobservable que no podemos evaluar directamente.

En consecuencia, elegimos un instrumento (el cuestionario) directamente relacionado con el aprendizaje, y que produce resultados observables (las respuestas de los alumnos). Si la hipótesis de investigación es cierta, esperamos que las puntuaciones del grupo de alumnos enseñados con ayuda del ordenador sean más altas que las de los estudiantes que no tienen acceso al mismo (*hipótesis experimental).* De cualquier modo, ya que hay múltiples factores que afectan el aprendizaje, además del ordenador, algunos estudiantes del grupo control tendrán mejores resultados que los de otros estudiantes del grupo experimental. Por tanto, necesitamos un procedimiento para comparar la distribución global en las dos poblaciones de estudiantes, en vez de comparar los casos aislados.

Esta comparación se hace usualmente considerando las puntuaciones medias en las dos poblaciones teóricas y especulando sobre su posible diferencia. Si la hipótesis experimental es cierta, esperaremos que esta diferencia sea positiva (*hipótesis estadística alternativa)* aunque quizás no podamos precisar el valor de la diferencia. En este caso no podemos trabajar directamente con la hipótesis alternativa y razonamos en su lugar como si las dos poblaciones tuvieran el mismo rendimiento, es decir asumimos que la *hipótesis nula* es cierta (la diferencia de puntuaciones medias en las dos poblaciones es cero)[1].

Para probar este supuesto calculamos un estadístico de contraste relacionado con el parámetro de interés, a partir de los datos de nuestras muestras. Al aceptar que la hipótesis nula es cierta, determinamos la distribución de este estadístico (una distribución T con 78 grados de libertad) que servirá para calcular el valor crítico y tomar una decisión acerca de si debemos o no aceptar nuestra hipótesis de investigación inicial.

- **Los tests de significación de Fisher**

Hay dos concepciones sobre los contrastes estadísticos:

1. Las pruebas de significación, que fueron introducidas por Fisher;
2. Los contrastes como reglas de decisión entre dos hipótesis, que fue la concepción de Neyman y Pearson.

---

[1] Nota: En el ejemplo, usaremos un test unilateral, porque hemos especificado la dirección de la diferencia con la hipótesis nula. En este caso la hipótesis nula es, en realidad $\mu_e[\mu_{c,}$ (el complemento de la hipótesis alternativa). Desde el punto de vista matemático, sin embargo, solo necesitamos calcular el valor crítico para el test bilateral, ya que siempre que un resultado sea significativo para la hipótesis $\mu_e=\mu_{c,}$ también lo será para la hipótesis $\mu_e<\mu_{c,}$. Por tanto, y para simplificar la exposición, supondremos en lo que sigue que la hipótesis nula es $\mu_e=\mu_{c.}$

La diferencia no se debe a los cálculos, sino al razonamiento subyacente. Siguiendo a Moore (1995) describiremos, en primer lugar, el razonamiento típico de una prueba de significación, que sería adecuada para el ejemplo 1, y en la que habría que seguir los pasos siguientes:

a) Describir el efecto que estamos buscando en función de los parámetros de una o varias poblaciones (la puntuación medias en el cuestionario de los estudiantes enseñados con ordenador μ_e es mayor que la de los alumnos que no tienen acceso al ordenador μ_c). El efecto que sospechamos es el verdadero describe la hipótesis alternativa: $H_1 \equiv \mu_e > \mu_c$.
b) Establecer la hipótesis nula de que el efecto no se presenta: $H_0 \equiv \mu_e = \mu_c$, (que no hay diferencias entre la puntuación media μ_e, de los estudiantes enseñandos con ordenador y la puntuación media μ_c de los estudiantes sin acceso al ordenador). La prueba de significación se diseña para evaluar la fuerza de la evidencia en contra de la hipótesis nula.
c) Calcular un estadístico a partir de los resultados en la muestra (estadístico calculado). La distribución del estadístico queda especificada cuando asumimos que la hipótesis nula es cierta (en el caso del ejemplo sería una distribución *T* con 78 grados de libertad). Supongamos que en el Ejemplo 1 obtenemos los siguientes valores para las medias de las dos muestras: $\overline{x}_e = 115.1$; $\overline{x}_c = 101.78$ y que $s^2_e = 197.66$; $s^2_c = 215.19$ son los estimadores insesgados de las varianzas en las poblaciones; la diferencia media de la puntuación en los dos grupos para estos datos es $\overline{x}_e - \overline{x}_c = 13.32$, la estimación conjunta de la varianza $s^2 = 202.48$, y usando las formulas standard obtendríamos un valor $t = 4.16$. La cuestión que el contraste de significación trata de contestar es la siguiente: Supongamos que la hipótesis nula es cierta y que, en promedio, no hay diferencia entre la puntuación media de las muestras tomadas de las dos poblaciones, ¿es entonces el resultado muestral $t = 4.16$ demasiado grande? O ¿podríamos obtener fácilmente este valor simplemente por causa de las fluctuaciones aleatorias del muestreo?
d) La probabilidad de obtener un valor *t* tan extremo o más que el valor *t* calculado cuando la hipótesis nula es cierta se llama *valor-p*. En nuestro ejemplo, el valor-p es extremadamente pequeño (menor que .001). Si la hipótesis nula es cierta y el valor-p muy pequeño, los resultados son altamente improbables y se llaman estadísticamente significativos. En este caso, y si los datos coinciden con la dirección especificada por la hipótesis alternativa, asumimos que nuestros datos proporcionan evidencia en contra de la hipótesis nula (ello no implica que creamos que la hipótesis nula es imposible; la hipótesis se aceptará en la comunidad científica sólo a partir de un programa de experimentos repetidos en los que repliquemos nuestros resultados; la ciencia se construye a partir de hallazgos acumulativos).
e) Incluso cuando la hipótesis nula es cierta, esperaremos algunas discrepancias entre las puntuaciones medias en los grupos experimental y control en el Ejemplo 1, debido a las fluctuaciones aleatorias del muestreo. No hay una regla fija acerca de cuan pequeño debe ser el valor-p para que un resultado se considere estadísticamente significativo, aunque convencionalmente adoptamos un valor fijo con el que comparamos el valor-p para decidir sobre su significación estadística. Es el nivel de significación α, o máximo valor-p admisible para considerar los datos como significativos, que se usa para calcular los valores críticos. Supongamos que tomamos $\alpha = .05$ en el ejemplo 1. El

valor crítico es la máxima diferencia que esperaríamos entre las dos muestras (grupos E y C) con probabilidad .05 (α), en caso de que las dos poblaciones tengan el mismo rendimiento. Este valor crítico se obtiene de la distribución teórica del estadístico en caso de que la hipótesis nula sea cierta (distribución T con 78 g.l.).

Actividad 2.33. ¿Equivale la obtención de un resultado estadísticamente significativo a la refutación lógica de la hipótesis nula? ¿Por qué?

Actividad 2.34. ¿Por qué no son lógicamente equivalentes el rechazo y la aceptación de una hipótesis en un contraste estadístico? ¿Cuáles son las conclusiones cuando se obtiene un resultado que no es estadísticamente significativo?

Actividad 2.35. ¿De qué factores depende el que un resultado sea estadísticamente significativo? (Por ejemplo, la diferencia entre las medias de dos grupos en una misma población).

Actividad 2.36. Supongamos que obtenemos que un coeficiente de correlación es significativo al nivel p=0.001. ¿Quiere decir esto que hay una relación muy intensa entre las variables?

Actividad 2.37. Supongamos que, sobre una misma muestra de estudiantes estudiamos las calificaciones en 10 asignaturas diferentes y nos interesa analizar cuáles de estas calificaciones difieren significativamente. Si usamos el test T de diferencias de medias relacionadas, a un nivel de significación del 0.01. ¿Cuántas diferencias habría que esperar resultasen significativas, simplemente por las fluctuaciones del muestreo, en el caso de que no existiese ninguna diferencia real entre las calificaciones? ¿Cómo podríamos solucionar este problema?

- **El contraste de hipótesis como proceso de decisión (teoría de Neyman y Pearson)**

En el ejemplo anterior hemos usado un contraste de significación para evaluar la fuerza de la evidencia en contra de una hipótesis nula. Hay, sin embargo, otras situaciones donde la inferencia se usa para tomar una decisión entre dos acciones posibles.

Ejemplo 2.6. Supongamos que una escuela secundaria quiere evaluar la efectividad de un nuevo método de enseñanza sobre el aprendizaje de la estadística de sus estudiantes. La escuela selecciona aleatoriamente 80 alumnos entre todos los estudiantes del último curso y divide aleatoriamente los 80 estudiantes en dos grupos de igual tamaño. El mismo profesor enseña estadística a ambos grupos durante un semestre. En el grupo C (grupo de control) se usa el método usual en la escuela, mientras que la enseñanza el grupo E (grupo experimental) se basa en los nuevos materiales. Al final del periodo el mismo cuestionario se aplica a los dos grupos (que fueron juzgados como equivalentes en su conocimiento inicial mediante un pretest). La escuela quiere cambiar al nuevo sistema si la diferencia media de puntuaciones en las dos poblaciones de estudiantes es positiva. Aquí el interés también va más allá de las dos muestras particulares (los grupos E y C), ya que el método sería aplicado a otros estudiantes.

En el ejemplo 2.6 se usaría un contraste estadístico con un razonamiento diferente, como procedimiento para tomar una decisión. Se seguirían los pasos siguientes:

a) Establecer las hipótesis nula $H_0 \equiv \mu_e = \mu_c$ y alternativa $H_1 \equiv \mu_e > \mu_c$, como en el Ejemplo 1.

b) Calcular el estadístico a partir de los datos de las muestras, como en el Ejemplo 1.
c) Se tomaría una decisión: O bien rechazamos la hipótesis nula (y aceptaríamos H_1) o bien no rechazaríamos H_0.
d) La decisión se hace comparando el p-valor con el nivel de significación α, es decir, comparando el valor t calculado con el valor crítico. En el Ejemplo 2 (suponiendo los mismos datos numéricos) el valor calculado $t = 4.16$ es mayor que el valor crítico $t_c = 1.665$, y por tanto, la hipótesis nula sería rechazado.

Es importante resaltar que rechazar la hipótesis nula no implica necesariamente que sea falsa, ya que es posible cometer dos tipos de error al tomar una decisión a partir de los resultados del contraste. En primer lugar, sería posible que no hubiese diferencia real en las puntuaciones medias de los estudiantes enseñados con los dos métodos y que, debido a la variabilidad aleatoria del muestreo, hayamos obtenido en nuestros grupos particulares E y C un valor t que ocurre con baja probabilidad. Como sabemos, que la probabilidad de un suceso sea muy baja no implica que el suceso sea imposible. Cometeríamos un *Error Tipo I* si rechazásemos una hipótesis nula que de hecho sea verdadera y la *probabilidad de Error Tipo I* es numéricamente igual al nivel de significación α.

Por otro lado, si el resultado no es significativo, ello no implica que las dos poblaciones tengan resultados igualmente buenos en el cuestionario. Incluso cuando los estudiantes enseñados con el nuevo método tengan mejores resultados, podríamos no obtener un resultado significativo en nuestras muestras particulares si el efecto de la enseñanza es pequeño o si hay demasiada variabilidad en los datos. Ocurre un *Error Tipo II* cuando el investigador acepta la hipótesis nula cuando, de hecho, es falsa. Puesto que hay muchas posibilidades diferentes para la diferencia de media en la hipótesis alternativa, la probabilidad de error Tipo II, β es variable. Normalmente estaremos interesados en algunos valores particulares de esta probabilidad y la calcularemos sólo para los casos más desfavorables.

El complemento de β se llama potencia del contraste y es la probabilidad de rechazar la hipótesis nula en caso de que sea falsa. Es también variable, ya que depende del valor verdadero del parámetro (la diferencia de medias en las poblaciones en nuestro caso). Conviene enfatizar la naturaleza condicional de las probabilidades de los dos tipos de error, ya que es en la interpretación de estas probabilidades condicionales donde encontramos más errores y concepciones erróneas respecto al contraste de hipótesis.

Es importante resaltar que, incluso cuando fijemos el nivel de significación α, es decir, la probabilidad de rechazar la hipótesis (supuesto que es cierta) y podamos calcular la probabilidad de obtener un valor del estadístico de contraste menor que un valor particular (supuesta la hipótesis nula cierta), la probabilidad de que la hipótesis nula sea cierta una vez la hemos rechazado y la probabilidad de que la hipótesis nula sea cierta una vez que hemos obtenido el valor del estadístico de contraste no pueden conocerse a partir del contraste de hipótesis, en las acepciones de Fisher o Neyman-Pearson.

- **La inferencia bayesiana**

Un planteamiento diferente de la inferencia estadística viene dado por la escuela Beyasiana. El *teorema de Bayes*, que permite calcular las *probabilidades a posteriori*, a partir del conocimiento de las *probabilidades a priori* y de los datos obtenidos experimentalmente. En su forma más simple, este teorema lo podemos expresar en la forma siguiente:

Supongamos que tenemos un suceso B (los datos) y queremos saber si ha sido producido por una de las causas A_1 A_2 , ...A_n (una serie de hipótesis científicas rivales; son las posibles causas de B).

Suponemos también que conocemos las probabilidades $P(A_1)$, $P(A_2)$, ...$P(A_n)$, es decir la probabilidad de cada una de las hipótesis rivales, así como las probabilidades $P(B/A_1)$, $P(B/A_2)$, $P(B/A_n)$ es decir las probabilidades "a priori" de obtener los datos B dependiendo de cual de las hipótesis es cierta.

Entonces, la probabilidad $P(A_i /B)$ (probabilidad a posteriori de que la hipótesis Ai sea la verdadera, una vez que hemos obtenido los datos B viene dada por la siguiente expresión:

$$P(Ai / B) = \sum_{i=1}^{n} \frac{P(Ai / B) P(Ai)}{P(Ai / B) P(Ai)}$$

En inferencia Bayesiana los parámetros de las poblaciones no se consideran valores constantes, sino variables aleatorias con una cierta distribución "a priori" de probabilidad, que queda revisada y transformada en distribución "a posteriori", como consecuencia de un proceso de inferencia (o proceso de aprendizaje) que es una generalización del teorema de Bayes.

De Finetti cree que este teorema (y toda la inferencia bayesiana que se obtiene generalizando y desarrollando este teorema) es un instrumento adecuado para obtener un conocimiento inductivo, pues las probabilidades iniciales pueden ser transformadas en probabilidades finales a la luz de los sucesos observados.

El teorema de Bayes contiene la esencia del método inductivo, porque podría aplicarse sucesivamente en nuevos experimentos. En el segundo experimento se tomarían como probabilidades a priori las probabilidades a posteriori que se obtuvieron en el primer experimento y así sucesivamente.

Una crítica al método bayesiano es que, si no se tiene un conocimiento del fenómeno, no es posible asignar las probabilidades iniciales y por tanto no es posible aplicar el teorema de Bayes.

Este problema se suele soslayar, suponiendo que, en ausencia de información inicial, las probabilidades a priori tienen una distribución uniforme, lo cual es considerado por algunos autores como inaceptable, porque en muchas situaciones no conocemos la distribución inicial, pero sabemos claramente que no es uniforme.

Otra posibilidad sería asignar subjetivamente las probabilidades iniciales, pero esto llevaría a que dos investigadores con los mismos datos obtuviesen unas probabilidades finales diferentes.

Otra sugerencia es que la asignación inicial no importa demasiado, porque la aplicación sucesiva del teorema de Bayes llevaría a ir corrigiendo poco a poco, en los nuevos experimentos el posible sesgo introducido en la asignación inicial de probabilidades.

Sin embargo, esto lleva a otro problema filosófico, porque admitir esto es admitir la existencia de probabilidad "a priori" objetivas hacia las cuales se irían acercando los cálculos y esto es precisamente contrario a la filosofía bayesiana, donde todas la probabilidades son subjetivas.

En definitiva, el método de Bayes no permite calcular ni revisar "probabilidades a priori /objetivas de la hipótesis" sino las "probabilidades a priori subjetivas" que estableció cada investigador. Es decir su grado personal de creencia en la veracidad de la hipótesis.

En este sentido subjetivo la inferencia bayesiana apoya el método inductivo, modificado en el sentido de buscar no la veracidad de una hipótesis (probabilidad objetiva) sino nuestra creencia en la misma (probabilidad subjetiva) y traslada el problema al debate sobre el carácter objetivo/subjetivo de la probabilidad.

En resumen, la probabilidad a posteriori de la hipótesis nula dado un resultado significativo depende de la probabilidad a priori de la hipótesis nula, así como de las probabilidades de obtener un resultado significativo, dadas las hipótesis nula y alternativa. Desafortunadamente estas probabilidades no pueden determinarse. Más aún, una hipótesis es o cierta o falsa y, por tanto, no tiene mucho sentido calcular su probabilidad en un paradigma inferencial clásico (donde damos una interpretación frecuencial a las probabilidades objetivas). Sólo en la inferencia Bayesiana pueden calcularse las probabilidades a posteriori de la hipótesis, aunque son probabilidades subjetivas, Lo más que podemos hacer, y es usando inferencia Bayesiana, es revisar nuestro grado de creencia personal en la hipótesis en vista de los resultados.

Actividad 2.38. Elaborar una lista de los conceptos claves que intervienen en un contraste de hipótesis según el modelo de Fisher y según el modelo de Neynam-Pearson y detectar las diferencias entre ambos.

Actividad 2.39. Analizar el concepto de potencia de un contraste y su relación con las probabilidades de error. ¿En qué momentos de la investigación debiera hacerse un análisis de la potencia del contraste?

Actividad 2.40. Analizar el uso que se hace del nivel de significación en algunas publicaciones sobre resultados experimentales. ¿Es posible detectar algunos errores de aplicación de este concepto? ¿A qué crees pueden ser debidos?

2.7.4. Los Diferentes Niveles de Hipótesis en la Investigación

El nivel de significación no es el único concepto mal comprendido en las pruebas de hipótesis. Algunas investigaciones también muestran confusión entre los papeles de las hipótesis nulas y alternativas (Vallecillos, 1994, 1995), así como entre la hipótesis estadística alternativa y la hipótesis de investigación (Chow, 1996). Chow diferencia diversas hipótesis implicadas en los diversos niveles de abstracción de la investigación experimental orientada a la confirmación de teorías, como la descrita en el Ejemplo 1.

- *Hipótesis substantiva* (que, en el Ejemplo1, es que el aprendizaje se ve influenciado por las herramientas semióticas disponibles para tratar un concepto). Una hipótesis substantiva es una explicación especulativa de un fenómeno. Normalmente no podemos investigarla directamente porque se refiere a un constructo o a un mecanismo inobservable. Para poder investigar la hipótesis substantiva debemos deducir algunas implicaciones observables de la misma.
- *Hipótesis de investigación* (que los ordenadores mejoran el aprendizaje de la estadística). Es una implicación observable de la hipótesis alternativa. Si no obtenemos

apoyo para la hipótesis de investigación, la hipótesis substantiva no se verá apoyada.

- Con frecuencia la hipótesis de investigación no es lo suficientemente específica para dirigir una investigación empírica. Es necesario diseñar una variable dependiente bien definida (en el Ejemplo 1 la puntuación total) obtenida a partir de una tarea experimental (el cuestionario) propuesta a algunos sujetos (los grupos control y experimental al finalizar la enseñanza). Con esta base podemos construir una *hipótesis experimental* (que el rendimiento en el cuestionario será mejor en la población experimental).

Tabla 2.*Diferentes Niveles de Hipótesis en la Investigación Experimental*

Hipótesis implicada	Ejemplo
Hipótesis substantiva	El aprendizaje depende de las representaciones disponibles
Hipótesis de investigación	Los ordenadores favorecen el aprendizaje de la estadística
Hipótesis experimental	Las puntuaciones en el cuestionario son más altas en los estudiantes que usan el ordenador
Hipótesis estadística alternativa	$H_1 \equiv \mu_e > \mu_c$
Hipótesis nula	$H_0 \equiv \mu_e = \mu_c$

- Una implicación de la hipótesis experimental se usará para llevar a cabo el análisis estadístico (que la puntuación media en el cuestionario será más alta en los estudiantes enseñados con ayuda del ordenador que en los estudiantes que no tienen acceso al ordenador). Esta implicación es la *hipótesis estadística alternativa*, $H_1 \equiv \mu_e > \mu_c$, que no coincide con la hipótesis experimental, sino que es una consecuencia de ella, a nivel estadístico.
- Finalmente, el complemento lógico de la hipótesis estadística alternativa es que la puntuación media en los dos poblaciones de estudiantes es la misma, $H_o \equiv \mu_e = \mu_c$. El establecer la hipótesis nula sirve para especificar la distribución del estadístico de contraste en el muestreo y comenzar la cadena de razonamientos (Tabla 3) que nos llevarán a rechazar o no la serie de hipótesis que hemos descrito y que se muestran en la Tabla 2.

Siguiendo a Chow (1996) presentamos en la Tabla 3 la serie de implicaciones deductivas anidadas de las cuales sólo la más interna (Implicación 5) se relaciona con el proceso de contraste de significación.

Este es el núcleo central de la cadena completa de implicaciones que en su conjunto constituyen el proceso de inferencia científica para dar apoyo a una hipótesis substantiva. En consecuencia, la concepción del contraste de hipótesis como prueba de significación se ajusta en forma natural a este tipo de investigación, mientras que la concepción de los contrastes estadísticos como proceso de decisión serían preferibles en las situaciones prácticas donde debemos tomar una decisión, como en el Ejemplo 2 o en control de calidad.

Tabla 3. *Cadena de Razonamientos Implicados en la Obtención de Apoyo a una Hipótesis Substantiva*

Implicación 1	Si el aprendizaje depende de las representaciones disponibles, entonces los ordenadores favorecen el aprendizaje de la estadística
Implicación 2	Si los ordenadores favorecen el aprendizaje de la estadística, entonces las puntuaciones en el cuestionario son más altas en los estudiantes que usan el ordenador
Implicación 3	Si las puntuaciones en el cuestionario son más altas en los estudiantes que usan el ordenador, entonces $\mu_e > \mu_c$
Implicación 4	Si no ocurre que $\mu_e > \mu_c$, entonces $\mu_e = \mu_c$
Implicación 5	Si $\mu_e = \mu_c$, entonces un valor significativo de $\overline{x}_e - \overline{x}_c$ es altamente improbable
Observación	$\overline{x}_e - \overline{x}_c$ es significativo,
Conclusión 5	$\overline{x}_e - \overline{x}_c$ es significativo, por tanto rechazamos que $\mu_e = \mu_c$
Conclusión 4	Rechazamos que $\mu_e = \mu_c$, por tanto asumimos que $\mu_e > \mu_c$
Conclusión 3	$\mu_e > \mu_c$; así que, supuesto que hubo un control experimental adecuado, suponemos que las puntuaciones en el cuestionario son más altas en los estudiantes que usan el ordenador
Conclusión 2	Las puntuaciones en el cuestionario son más altas en los estudiantes que usan el ordenador; supuesto que el test es una medida válida y fiable del aprendizaje, entonces los ordenadores favorecen el aprendizaje de la estadística
Conclusión 1	Los ordenadores favorecen el aprendizaje de la estadística, supuesto que la única diferencia en los dos métodos de enseñanza es el sistema de representaciones disponible, entonces la hipótesis susbtantiva es apoyada por nuestros datos

Las implicaciones 1 a 4 en la Tabla 3 no están apoyadas por la teoría estadística, sino por consideraciones teóricas del campo bajo estudio y por un control experimental adecuado que asegure que todas las variables concomitantes relevantes se mantienen constantes y que el cuestionario dado a los estudiantes es una medida válida y fiable del constructo estudiado (aprendizaje). Según Chow (1996), muchas de las críticas en contra del contraste estadístico son injustas, pues se refieren no al procedimiento estadístico (implicación 5 en la Tabla 3), sino al resto de los componentes del proceso inferencial (implicaciones 1 a 4). Reemplazar o complementar los contrastes estadísticos con otros métodos tales como intervalos de confianza o análisis de potencia no resolverá los problemas de adecuado control experimental o falta de marco teórico pertinente en un campo particular de estudio.

2.7.5. Revisión de Algunas Críticas Comunes en Contra de los Contrastes de Hipótesis

Hemos identificado ya algunas de las razones que dificultan la comprensión de los tests estadísticos. Por un lado, el contraste estadístico implica una serie de conceptos como hipótesis nula y alternativa, errores Tipo I y II, probabilidad de los errores, resultados significativos y no significativos, población y muestra, parámetro y estadístico, distribución de la población y distribución muestral. Algunos de estos conceptos son mal interpretados o confundidos por los estudiantes e investigadores experimentales. Más aún, la estructura formal de los contrastes estadísticos es superficialmente parecida a la de la prueba por contradicción, pero hay diferencias fundamentales, no siempre bien comprendidas entre estos dos tipos de razonamientos.

En una prueba por contradicción razonamos en la forma siguiente:

Si A, entonces B no puede ocurrir.

B ocurre; entonces deducimos que A es falso.

En un contraste estadístico tratamos de aplicar un razonamiento similar en la siguiente forma:

Si A, entonces es muy poco verosímil que B ocurra;

Ocurre B, y rechazamos A

Sin embargo no sería una conclusión válida deducir que es muy improbable que A sea cierto y aquí es donde encontramos la confusión.

Estas dificultades han originado una gran controversia sobre el uso de la inferencia estadística en la investigación experimental. La ciencia se construye a partir de observaciones empíricas y no podemos tomar datos de las poblaciones completas, sino sólo de muestras de las mismas. Esperamos del contraste de hipótesis más de lo que nos puede dar y bajo estas expectativas subyace el problema filosófico de hallar criterios científicos para justificar el razonamiento inductivo, como estableció Hume.

Hasta ahora las contribuciones de la inferencia estadística en esta dirección no han dado una solución completa al problema (Black, 1979; Burks, 1977; Hacking, 1975; Seidenfeld, 1979). Sin embargo, los contrastes de hipótesis y toda la inferencia son una herramienta muy adecuada, la mejor que tenemos por el momento. Incluso aunque no resuelvan todos nuestros problemas, usadas de manera inteligente constituyen una guía en el proceso de descubrimiento científico. Revisaremos a continuación algunas críticas frecuentes en contra del contraste estadístico, argumentando que carecen de fundamento.

1. *Lo que establecemos en la hipótesis nula del ejemplo 1 es que no hay diferencia entre las medias de dos poblaciones. Es evidente para muchos críticos que la hipótesis nula nunca es cierta y por tanto los contrastes estadísticos no son válidos, al basarse en una premisa falsa (que la hipótesis nula es cierta).*

   Es fácil deducir que esta crítica no es pertinente, ya que el hecho de que la hipótesis nula no sea cierta no invalida la lógica del contraste estadístico. Esta lógica no se ve afectada por el hecho de que la hipótesis nula sea cierta o falsa, porque lo que afirmamos en un contraste es que un resultado significativo es improbable, en caso de que la hipótesis nula sea cierta. Esto es una propiedad matemática de la distribución muestral que no tiene nada que ver con la certeza o falsedad de la hipótesis nula.

2. *En la práctica identificamos la hipótesis de interés con la hipótesis estadística alternativa. Pero la hipótesis estadística alternativa no indica la magnitud exacta de la diferencia entre las medias de las poblaciones. La significación estadística no informa sobre la significación práctica de los datos.*

   Cuando aplicamos esta crítica a los contrastes de significación (Ejemplo 1) estaríamos cayendo en la confusión entre los diferentes niveles de hipótesis implicados en un procedimiento inferencial, que se muestran en las Tablas 2 y 3. El fin de la investigación experimental orientada a la confirmación de teorías es proporcionar apoyo a la hipótesis substantiva. Como vemos en el ejemplo 1, la magnitud de las diferencias entre las medias de las poblaciones no depende de la hipótesis substantiva. La diferencia se refiere a la distribución del estadístico en el muestreo, esto es, a la hipótesis estadística alternativa. No hay una única correspondencia entre la hipótesis substantiva y la hipótesis estadística alternativa, que se deduce de un experimento particular y de un instrumento particular. Las teorías deben evaluarse con un razonamiento cuidadoso y un profundo juicio (Harlow, 1977).

En el contexto de toma de decisión (Ejemplo 2), por el contrario, la magnitud del efecto podría ser relevante para la decisión. En estos casos, los contrastes estadísticos son todavía útiles para hacer la decisión, aunque deben complementarse con el análisis de la potencia y/o la estimación de la magnitud de los efectos, dependiendo de las preguntas de interés en la investigación (Levin, 1998a).

3. *La elección del nivel de significación es arbitraria; por tanto algunos datos podrían ser significativos a un nivel dado y no serlos a un nivel diferente.*

Es cierto que el investigador escoge el nivel de significación, pero esta arbitrariedad no implica, sin embargo, que el procedimiento sea inválido o inútil. Además es también posible, si se sigue el enfoque de Fisher, usar el p-valor exacto para rechazar las hipótesis nulas a diferentes niveles, aunque en la práctica actual del contraste estadístico se aconseja elegir el nivel de significación antes de recoger los datos para dar mayor objetividad a la decisión.

4. *La significación estadística no informa de la probabilidad de que la hipótesis sea cierta ni del verdadero valor del parámetro. Por ello muchos investigadores sugieren reemplazar los contrastes por intervalos de confianza.*

Es verdad que los contrastes no informan acerca de la probabilidad de que la hipótesis sea cierta, pero tampoco los intervalos de confianza informan sobre esta probabilidad. Los intervalos de confianza son intervalos en los que el verdadero valor del parámetro se encuentra en un porcentaje dado de muestras, aunque no aseguran en que intervalo estará el parámetro en nuestro experimento particular. Por tanto no substituyen sino más bien complementan los contrastes de hipótesis y están sujetos a las mismas controversias e interpretaciones erróneas que aquellos.

5. *Los errores Tipo I y Tipo II están inversamente relacionados. Para los críticos los investigadores parecen ignorar los errores Tipo II mientras prestan una atención indebida a los errores Tipo I.*

Aunque las probabilidades de los dos tipos de error están inversamente relacionadas, hay una diferencia fundamental entre ellas. mientras que la probabilidad de error Tipo I α es una constante que puede elegirse antes de realizar el experimento, la probabilidad de error Tipo II es función del verdadero valor desconocido del parámetro.

Para resolver este problema, el análisis de la potencia supone diferentes valores posibles del parámetro y calcula la probabilidad de error Tipo II para estos diferentes valores. Esta práctica es útil para algunas aplicaciones de la inferencia, como en la toma de decisiones (ejemplo 2) y control de calidad. Sin embargo podemos aplicar aquí las mismas objeciones que en el punto 2 respecto al experimento orientado a apoyar una hipótesis substantiva teórica dada (ejemplo 1), donde no hay indicación sobre el valor particular de los parámetros de las variables experimentales dependientes. Es decir, los dos tipos de error no juegan el mismo papel en la corroboración de teorías científicas, aunque pueden ser igualmente importantes en otras aplicaciones de la inferencia, como el control de calidad.

6. *El significado de un resultado no significativo no es claro. Para algunos críticos esto se debe a que el contraste que se usa no tiene suficiente potencia.*

Podemos aplicar aquí el mismo razonamiento seguido en el punto 4. Es claro que las hipótesis nulas y alternativa, el rechazo y aceptación de la hipótesis nula, los resultados significativos y no significativos no juegan un papel simétrico en los contrastes de significación. Mientras que un resultado significativo contradice la hipótesis nula por su baja probabilidad, un resultado no significativo es altamente probable cuando la hipótesis nula es cierta, pero también podría deberse a otros factores. Esto también puede ocurrir en la prueba por contradicción, donde hacemos el siguiente razonamiento:

Si A es cierta, B es falsa

Entonces si B es cierta, A es falsa

Si B ocurre, podemos concluir que A no es posible, pero cuando ocurre B no podemos deducir que A sea necesariamente cierta; hay aquí también una asimetría entre las consecuencias de B y no B.

2.7.8. Analisis multivariante de datos

Las técnicas multivariantes de análisis de los datos constituyen una ruptura epistemológica con la estadística clásica, desde el punto de vista de los objetivos, el tipo de datos tratados (número, naturaleza, variedad), los sujetos de análisis (variables o individuos), el proceso (de los datos al modelo y no a la inversa), los métodos matemáticos empleados y los conceptos implícitos en los mismos. Supongamos que nuestros datos vienen dados por la siguiente matriz:

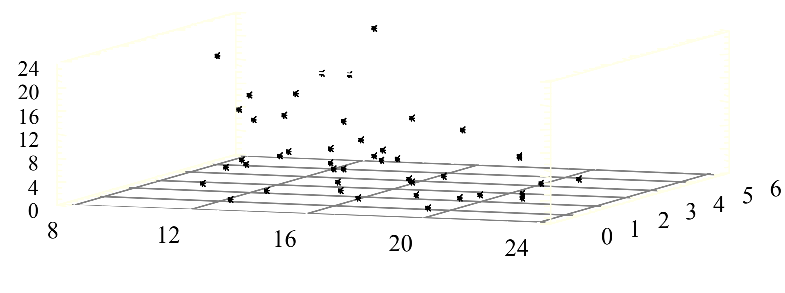
$$\begin{array}{c} x_{11}\ ,\ x_{12}\ ,\ldots\ldots x_{1p}\ , \\ \ldots\ldots\ldots\ldots\ldots\ldots, \\ x_{21}\ ,\ x_{22}\ ,\ldots\ldots x_{2p}\ , \\ x_{n1}\ ,\ x_{n2}\ ,\ldots\ldots x_{np}\ , \end{array}$$

En esta matriz, cada fila consta de los valores observados de *p* variables diferentes sobre un individuo y cada columna el conjunto de *n* valores diferentes de una misma variable aleatoria sobre n individuos distintos.

El conjunto de respuestas de los sujetos a los ítems se toman como coordenadas de dos sistemas de puntos en un doble espacio vectorial multidimensional empírico.

El primero de estos espacios es el de los individuos, en el cual cada fila de la matriz (individuo) es un vector (o punto), cuyas coordenadas son los valores correspondientes de las p variables. De esta forma obtenemos n puntos en un espacio vectorial de dimensión p. En este espacio podemos usar la distancia euclídea para medir la "proximidad" entre dos individuos.

El segundo de estos espacios es el de las variables, que se obtiene tomando como



vectores las columnas de la matriz. Este espacio es de dimensión *n* y consta de *p* puntos. Si usamos la distancia euclídea para medir la "proximidad" entre variables, puede demostrarse la siguiente relación

$$d(v_i,v_j)= 2(1-r_{i,j})$$

siendo $r_{i,}$ el coeficiente de correlación muestral de las variables v_i , v_j. Por tanto, *dos variables estarán a distancia máxima si $r_{i,j}$ =-1, a distancia cero si $r_{i,j}$ =1 y a distancia media si están incorreladas.*

Puesto que en este espacio el coseno del ángulo que forman dos vectores o variables es el coeficiente de correlación entre ellos, *dos variables incorreladas son ortogonales.*

Además de las distancia anterior (distancia euclídea) pueden utilizarse otras diferentes, como la distancia Chi-cuadrado, o la suma de los valores absolutos de las diferencias de coordenadas. La elección de una u otra dependerá del tipo de datos y el propósito del análisis.

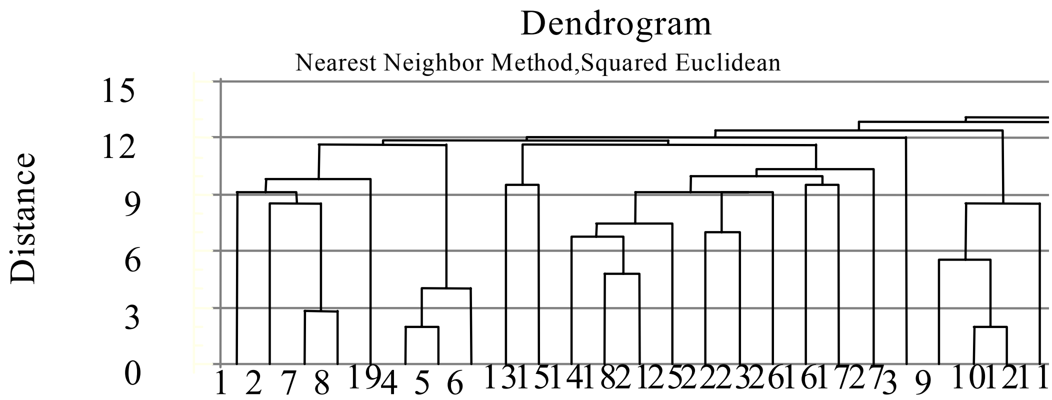
Actividad 2.38. Propón una actividad para alumnos de secundaria para representar un conjunto sencillo (3 o 4 variables) de tipo multivariante.

Actividad 2.39. Propon un ejemplo en que, para interpretar los valores de una variable, sea necesario conocer los de otras relacionadas con la anterior.

Actividad 2.40. Analiza la interpretación geométrica del coeficiente de correlación como producto escalar entre dos vectores. ¿Qué implicaciones tiene para la importancia que en análisis factorial tiene el buscar factores no correlacionados?

- **Las técnicas de clasificación**

Un primer grupo de técnicas de clasificación permiten agrupar individuos y variables mediante algoritmos de tipo muy diverso (análisis "cluster" de sujetos o variables; jerárquico o no jerárquico; ascendente o descendente).



La proximidad entre ítems señalará la semejanza de respuestas respecto a estos ítems por parte de los sujetos, o lo que es lo mismo, la presencia de similitudes en componentes del conocimiento relacionadas con la resolución de estos ítems (análoga interpretación tiene la proximidad entre sujetos).

Actividad 2.42. La agrupación en intervalos es una primera técnica de clasificación. Analizar las semejanzas entre el histograma de frecuencias y el dendograma empleado en análisis cluster, como gráficos que permiten producir tipologías de individuos. ¿Tienen estas tipologías un carácter objetivo o subjetivo? ¿Por qué?

Actividad 2.43. Analizar hasta qué punto la clasificación es el primer paso en la conceptualización y cómo los métodos multivariantes nos pueden ayudar en este proceso.

- **Las técnicas de reducción de dimensión**

Son técnicas geométricas que tratan de visualizar y simplificar la estructura de los datos. Se pretende proyectar los puntos dados en un espacio de dimensión inferior, conservando ciertas propiedades del conjunto de datos inicial, especialmente las referidas a las distancias entre individuos. Mediante el estudio de las proyecciones obtenidas podemos deducir ciertas propiedades de la estructura global, al igual que mediante una fotografía, podemos reconocer los objetos tridimensionales. Estas técnicas,

> "*permiten utilizar las facultades de percepción que usamos a diario: sobre los gráficos del análisis factorial, se ven, en el propio sentido del término (con los ojos y el misterioso análisis que nuestro cerebro hace de una imagen) reagrupamientos, oposiciones, tendencias imposibles de discernir directamente sobre una gran tabla de números, incluso después de un examen prolongado*" (Escofier y Pagés, 1988, p.1).

Al reducir la dimensión en que se representan los datos, en general tendremos el problema de la no unicidad de solución. Pero no todas las soluciones posibles nos proporcionan la misma información sobre el objeto total, al igual que no todos los puntos desde los que se fotografía un objeto son igualmente buenos para lograr su reconocimiento. De este modo se necesitarán ciertos criterios, para lograr estos óptimos. Así, al representar gráficamente, por ejemplo, la relación entre peso y altura, vemos que este conjunto de puntos bidimensional tiene una estructura aproximadamente lineal, aunque para elegir una recta que represente los datos, se añaden ciertas condiciones, como la de minimización de la varianza residual en el método de mínimos cuadrados. En general, se emplea una extensión de este principio -minimización de la varianza residual o maximización de la varianza explicada cuando se realiza un reducción de la dimensión en un conjunto de datos.

Podemos considerar también que estamos haciendo un cambio de base. Las observaciones originales de los individuos están expresadas en la base formada por las variables, que pueden o no estar correlacionadas, es decir en eneral no es una base ortogonal. Al transformar el conjunto inicial de variables en otro incorrelado, hacemos un cambio de base a una nueva ortogonal, en la que la interpretación de las interrelaciones existentes resulta más clara, ya que cada elemento de la base se interpreta independientemente del resto.

Entre los métodos que podemos incluir en este apartado se hallan el análisis factorial, análisis de componentes principales, análisis factorial de correspondencias y las técnicas de producción de escalogramas

Actividad 2.44. Analiza la regresión bivariante como técnica de reducción de la dimensión. ¿Podríamos *decir que es precursora del análisis factorial?*

Actividad 2.45. ¿Qué ideas del diseño experimental (debidas a Fisher) están latentes en el análisis factorial?

3

INVESTIGACIONES SOBRE COMPRENSIÓN Y APRENDIZAJE DE CONCEPTOS ESTADÍSTICOS

3.1. Introducción

En este capítulo analizaremos las investigaciones sobre razonamiento estocástico, así como las relacionadas con los errores y dificultades que persisten después de la enseñanza de la estadística. Esta investigación proviene, principalmente de dos fuentes: la psicología y la propia educación estadística, que trabajan con distintos fines sobre una misma problemática:

Para las personas que se interesan por la educación estadística la preocupación fundamental es identificar los puntos difíciles y los errores que continúan al finalizar la enseñanza, para poder diseñar actividades didácticas adecuadas para superar estas dificultades e informar al profesor sobre las mismas.

En Psicología lo fundamental es analizar el razonamiento humano en situaciones de incertidumbre y la preocupación por la enseñanza es secundaria. Libros como los de Nisbett y Ross (1980) y Kahneman, Slovic y Tversky (1982) describen los errores sistemáticos que cometemos al tomar decisiones en situaciones de incertidumbre, situaciones que son frecuentes en campos tan diversos como la abogacía, educación, economía, medicina o política.

Estos errores tienen una base psicológica, y la comprensión de las leyes teóricas de la probabilidad no siempre es suficiente para superarlos, si el sujeto no llega a tomar conciencia de ellos, por lo que es importante que el profesor reconozca estos sesgos cuando sus alumnos los manifiesten.

Kilpatrick (1994) hace un resumen de la historia de la educación matemática, analizando la influencia que en su desarrollo han tenido la propia matemática y otras disciplinas, como la psicología y pedagogía. Aunque la educación estadística puede considerarse una rama de la educación matemática, tiene sin embargo un desarrollo mucho más reciente, pues la investigación sobre la enseñanza de la estadística no ha interesado a los educadores matemáticos, sino muy recientemente. Otra característica es que la preocupación por la enseñanza a nivel universitario es mucho mayor en el caso de la estadística y participan en ella profesores que, teniendo una formación básica de otras especialidades, como educación, economía, ingeniería, empresariales, etc. se han hecho cargo de la enseñanza de la estadística en sus propias especialidades.

En todo caso la influencia de la psicología ha sido también muy grande en la Educación Matemática, ya que como indica Kilpatrick aportó un conjunto de marcos teóricos y métodos potencialmente útiles para sus investigaciones. Un resumen de la historia de la educación matemática en España se presenta en Rico (1991) y Rico y Sierra (1994).

En lo que sigue analizamos los resultados de estas investigaciones, comenzando por describir sus diferentes marcos teóricos, así como por explicitar nuestra propia concepción del conocimiento matemático que servirá para fundamentar el posterior análisis, así como nuestras sugerencias para la enseñanza de la estadística.

3.2. Investigaciones sobre desarrollo cognitivo de Piaget y Fischbein

Cuando se sugiere introducir un nuevo tema en el currículo es muy importante estudiar el razonamiento de los niños respecto al mismo, para poder valorar hasta qué punto son asequibles para ellos los nuevos conocimientos que tratamos de enseñar. Los niños aprenden no sólo en la escuela, sino en su entorno familiar y social, y su razonamiento se modifica gradualmente, a partir de sus experiencias y de la interacción con los objetos y el mundo que les rodea. En el caso de la probabilidad y la estadística es especialmente importante analizar los razonamientos de los niños, puesto que en dichas materias tratamos con ideas bastante abstractas y no tan ligadas a la experiencia directa del niño como pudieran ser los conceptos geométricos o numéricos.

Desde muy pequeño el niño debe aprender a estimar, discriminar y diferenciar formas, distancias y cantidades. Las operaciones aritméticas básicas se pueden también concretizar en operaciones con objetos físicos (juntar o separar colecciones, etc.) que tienen la propiedad de ser reversibles (volver a los operandos primitivos al deshacer la operación). Por el contrario, no existe una experiencia concreta similar de lo aleatorio, ya que no podemos manipular estos fenómenos para producir un resultado específico, ni devolver los objetos a su estado inicial deshaciendo la operación.

Ejemplo 3.1. Si hacemos girar la aguja en una ruleta, desde una posición inicial, impulsándola hacia la derecha, es muy poco probable que un impulso a la izquierda devuelva la aguja a su posición inicial.

Esta falta de reversibilidad de los experimentos aleatorios sin duda influye en el desarrollo más tardío de las nociones de probabilidad. Respecto a la estadística, son prácticamente inexistentes los estudios sobre el desarrollo evolutivo de los conceptos y los pocos que se han llevado a cabo, están relacionados con la instrucción, por lo que los analizaremos dentro de los estudios sobre enseñanza y aprendizaje.

En nuestro libro anterior sobre Azar y Probabilidad (1987) hemos analizado con detalle las investigaciones sobre el desarrollo del razonamiento probabilístico en los niños llevadas a cabo por Piaget e Inhelder (1951) y Fischbein (1975). Sin embargo, creemos necesario hacer un breve resumen aquí y, sobre todo, completar los aspectos no tratados en nuestro libro anterior.

Piaget se centró en dar criterios para determinar en qué nivel de desarrollo intelectual se encuentra el niño a diversas edades, y analiza la comprensión formal de los conceptos. Piaget postula que la experiencia, la actividad y el conocimiento previo son las bases que determinan el aprendizaje. El conocimiento es construido activamente por el sujeto y no recibido pasivamente del entorno. El niño trata de adaptarse al mundo que le rodea. Cuando una idea nueva se le presenta, se crea un *"conflicto cognitivo"* o "*desequilibrio*" en su estado mental si esta idea choca con las ya existentes.

Ejemplo 3.2. Cuando se comienza el estudio de los números decimales, puede crearse un conflicto cognitivo respecto a la idea asumida previamente de que todo número tiene un número siguiente (y *que era válida para los números naturales).*

Para reaccionar a este desequilibrio se requiere un proceso de *"equilibración"* que consiste en los pasos de *asimilación* y *acomodación*. La asimilación es la incorporación (aceptación) por parte del sujeto de los datos nuevos La acomodación es el cambio o reestructuración de los ya existentes. El aprendizaje se concibe como un proceso que progresa lentamente con puntos conflictivos que el alumno debe superar mediante el proceso descrito. Conocer es un proceso de adaptación que organiza el propio mundo experiencial.

La posibilidad de aprender depende del conocimiento previamente adquirido y del desarrollo intelectual del alumno, que sigue una serie de etapas. Las etapas son particiones en fases, de modo que los sujetos que están en una misma fase tienen un modo de razonamiento similar y la progresión de una etapa a otra siempre sigue un cierto patrón. Para el estudio del desarrollo del razonamiento probabilístico son relevantes tres etapas:

- *Preoperatoria,* caracterizada por la necesidad de manipular objetos reales para el aprendizaje de un cierto concepto, pues el niño se apoya en sus experiencias empíricas para comprender los conceptos.
- *Operaciones concretas*: Se comienza a comprender la conservación de la masa, peso, número y volumen. Aparecen conceptos secundarios, que no necesitan ser abstraídos de la experiencia concreta.
- *Operaciones abstractas*: Se pueden manipular relaciones entre representaciones simbólicas, se formulan hipótesis y se establecen conclusiones. Se comprende el significado de abstracciones verbalmente, sin referirse a objetos particulares.

Fischbein, por su parte, se preocupó por demostrar que los niños tienen ideas correctas parcialmente formadas sobre los conceptos probabilísticos y analizó el efecto de la instrucción para la mejora de estas intuiciones. También concede una gran importancia a la intuición como componente de la inteligencia.

Las intuiciones son, según Fischbein, procesos cognitivos que intervienen directamente en las acciones prácticas o mentales, y tienen las siguientes características: inmediatez, globalidad, capacidad extrapolatoria, estructurabilidad y auto-evidencia. La inmediatez significa que las intuiciones no son reflexivas, sino que surgen con frecuencia en forma espontánea. El carácter global se opone al analítico o descomposición en partes. Las intuiciones van más allá de un caso particular, en cierto modo tienen un carácter teórico y por eso sirven para extrapolar o hacer predicciones. Parecen autoevidentes para el sujeto, quien no necesita demostración. Diversas intuiciones de relacionan entre sí, formando estructuras de razonamiento. Fischbein diferencia entre intuiciones primarias y secundarias:

- Las *intuiciones primarias* se adquieren directamente con la experiencia, sin necesidad de ninguna instrucción sistemática. Ejemplo de ellas son las intuiciones espaciales elementales, como el cálculo de distancia y localización de objetos, o el admitir que al lanzar un dado todas las caras tienen la misma probabilidad de salir.
- Por el contrario, las *intuiciones secundarias* se forman como consecuencia de la educación, principalmente en la escuela.

Ejemplo3.3. La idea de que un móvil conserva su estado de movimiento o de reposo mientras no intervenga una fuerza exterior es una intuición primaria. Una intuición secundaria (errónea) es la llamada "falacia del jugador", por la cual, después de lanzar una moneda una serie de veces y haber obtenido caras, el sujeto tiende a predecir que la próxima vez es más probable que salga cruz. Esto se debe a una mala interpretación de la ley de los grandes números.

Una intuición secundaria no se reduce a una simple fórmula aceptada o utilizada automáticamente, sino que se transforma en convicción, en creencia, en un sentimiento de evidencia. Pero una intuición no se forma a partir de la información obtenida de una lectura o de una explicación teórica, sino de una información que el alumno utiliza en sus propias acciones y predicciones a lo largo de gran parte de su desarrollo intelectual.

Actividad 3.1. Analiza las características que Fischbein atribuye a las intuiciones, comprobando si pueden aplicarse a la *"falacia del jugador"*.¿Podrías dar una lista de otras intuiciones primarias y secundarias en el campo de la probabilidad? ¿Es alguna de ellas incorrecta?

- **La Intuición del Azar**

El primer paso para comenzar a enseñar probabilidad es asegurarnos que los niños son capaces de diferenciar las situaciones aleatorias y deterministas, es decir de apreciar algunas características básicas de la aleatoriedad.

Piaget e Inhelder (1951) defienden que la comprensión del azar por parte del niño es complementaria a la de la relación causa-efecto. Los niños concebirían el azar como resultado de la interferencia y combinación de una serie de causas, que actuando independientemente producirían un resultado inesperado. En consecuencia, hasta que el niño no comprende la idea de causa, no tiene un marco de referencia para identificar los fenómenos aleatorios. El azar habría que considerarlo asimismo como complementario a la composición lógica de operaciones reversibles y requere un razonamiento combinatorio, para poder concebir las distintas posibilidades existentes en estas situaciones.

Un experimento piagetiano clásico utiliza una bandeja con dos compartimentos en los extremos de la bandeja, que se unen en el centro de la misma (Figura 1). En uno de estos dos compartimentos se colocan cuatro bolas blancas y en el otro cuatro rojas, de modo que al bascular la bandeja se produce la mezcla progresiva de las dos clases de bolas. Antes de mover la bandeja, Piaget pide a los niños que hagan una predicción sobre la colocación final de las bolas.

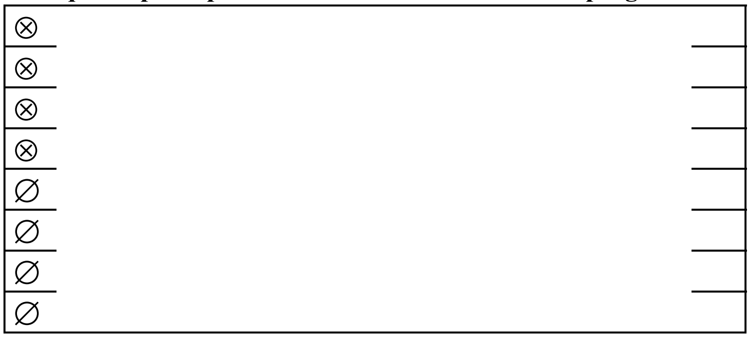
En el *periodo preoperacional* los niños piensan que, después de mover varias veces la bandeja, las bolas vuelven nuevamente a su lugar original, o bien que el conjunto completo de blancas acabarán en el lugar ocupado originalmente por las rojas, y viceversa.

Piaget interpreta esta reacción típica de los niños, antes de los 7 años, indicando que el niño no comprende la naturaleza irreversible de la mezcla aleatoria por tener un pensamiento reversible. Además el niño de esta edad no comprende bien la relación entre causa y efecto, ni tiene razonamiento combinatorio completo, por lo que Piaget piensa no hay una intuición del azar innata en el niño, como no existía tampoco en el hombre primitivo, que atribuía los sucesos aleatorios a causas ocultas o a la "voluntad de los dioses".

Esta opinión de Piaget es rechazada por Fischbein para quien la *intuición primaria* del azar, esto es, la distinción entre fenómeno aleatorio y determinista aparece antes de los

7 años. Fischbein se basa en la conducta de los niños al practicar juegos de azar, ya que en juegos sencillos, los niños son capaces de elegir la opción de mayor probabilidad.

Figura 1: El experimento de la bandeja: Al mover la bandeja, las bolas, que en principio estaban ordenadas se mezclan progresivamente



En el *periodo de las operaciones concretas,* con la adquisición de esquemas operacionales espacio-temporales y lógico-matemáticos, el niño alcanza la capacidad de distinguir entre el azar y lo deducible, aunque esta comprensión no es completa, puesto que el pensamiento está todavía muy ligado al nivel concreto. No obstante, el niño comienza a comprender la interacción de cadenas causales que conducen a sucesos impredecibles, y la irreversibilidad de los fenómenos aleatorios. Pero, bien porque comprende la interferencia de las causas, sin reconocer su independencia, bien porque comprende la independencia y no la interferencia, no llega a construir la idea de azar.

En los fenómenos aleatorios los resultados aislados son imprevisibles pero el conjunto de posibilidades puede determinarse mediante un razonamiento de tipo combinatorio, con lo que se vuelve previsible. Cuando se comprende ésto aparece la idea de probabilidad expresada por la razón entre las posibilidades de un caso particular y del conjunto de posibilidades. Por tanto, la idea de azar, para Piaget, lo mismo que la de probabilidad, no puede ser totalmente adquirida hasta que se desarrolle el razonamiento combinatorio, en la etapa de las *operaciones formales.*

Fischbein sostiene que la distinción entre el azar y lo deducible no se realiza espontánea y completamente al nivel de las operaciones formales. En experimentos donde se requiere al sujeto calcular probabilidades el adolescente a veces busca dependencias causales que reduzcan lo incierto, incluso en situaciones donde no existen tales dependencias, siendo influenciado por las tradiciones culturales y educativas de la sociedad moderna, que orientan el pensamiento hacia explicaciones deterministas.

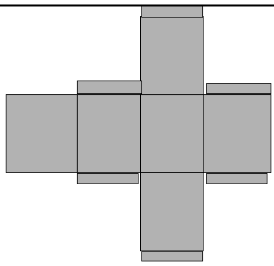
Actividad 3.2. ¿Cómo podríamos aportar datos a favor de una de las dos teorías de Piaget o Fischbein sobre que la intuición del azar aparece en los niños pequeños?

Actividad 3.3. Analizar algunos segmentos de las entrevistas contenidas en el libro de Piaget e Inhelder (1951) sobre el experimento de la bandeja, con los niños de preescolar. ¿Crees que hay evidencias suficientes para apoyar la teoría de Piaget?

Actividad 3.4. Discutir las implicaciones educativas de la teoría de los estadios de Piaget. ¿Son los estadios de desarrollo simultáneos para los diversos tipos de razonamiento (por ejemplo, espacial, aritmético, probabilístico)?

- **La estimación de la frecuencia relativa**

Supuesto que un niño sea capaz de diferenciar los fenómenos aleatorios y deterministas, el segundo paso es que pueda estimar en una serie de experimentos cuáles son los sucesos que aparecen con mayor o menor frecuencia. Muchos psicólogos han llevado a cabo experimentos de aprendizaje probabilístico, en los cuales se trata de estudiar las predicciones de los sujetos ante situaciones en que un suceso se repite con una determinada frecuencia relativa.



Ejemplo 3.4. Se presenta al alumno dos luces de color diferente (pueden ser rojo y verde) que se irán encendiendo intermitente y aleatoriamente con una determinada frecuencia, por ejemplo, el 70 y el 30%, respectivamente. El sujeto debe predecir cuál de las dos luces se encenderá la próxima vez.

Los resultados obtenidos en este tipo de experimentos apoyan fuertemente la conclusión de que *en el perido preoperacional* el niño adapta sus predicciones a las probabilidades de los sucesos que se le presentan como estímulo. Ello nos indica que los niños son capaces de apreciar las diferentes frecuencias relativas con que aparecen los resultados de los fenómenos aleatorios.

La estimación de la frecuencia relativa de sucesos mejora en el *periodo de operaciones concretas*. Como resultado de experiencias acumuladas, la intuición de la frecuencia relativa se desarrolla de un modo natural como consecuencia de las experiencias del niño con situaciones que implican sucesos aleatorios, en las que sea necesaria una estimación de las frecuencias relativas de los fenómenos. Es fácil pensar en tales experiencias en la vida diaria, por ejemplo, cuando estimamos el tiempo que transcurre entre un autobús y otro, o hacemos un a previsión sobre si va a llover o no, según el día se presente más o menos nublado, etc.

En el *periodo de las operaciones formales* el adolescente ha hecho progresos en comparación a los niños más pequeños en lo que se refiere a la intuición de la frecuencia relativa, particularmente en casos donde las predicciones tienen algún resultado práctico. La estrategia óptima ante decisiones en condiciones aleatorias muestra los efectos favorables del desarrollo de la inteligencia sobre las predicciones de ocurrencia de los fenómenos aleatorios.

Actividad 3.5. Podemos replicar las investigaciones sobre intuición de la frecuencia relativa en forma sencilla, tratando de ver si los niños diferencian, a partir de la experimentación entre dados sesgados y no sesgados. Un dado ordinario se puede construir recortando en cartulina el siguiente perfil. Se puede construir un dado sesgado recortando en cartulina este perfil, pero numerando dos caras con el número 5 y ninguna con el 1, o bien pegando un pequeño peso en la cara del 1, por ejemplo, un botón

a) Construye uno de estos dados y realiza un experimento con algunos niños viendo si estos adaptan sus apuestas a los valores que aparecen con frecuencia máxima. b) Analiza el tipo de situaciones didácticas en las que podría ser de utilidad este tipo de material.

- **Estimación de posibilidades y noción de probabilidad**

Piaget e Inhelder pensaron que el niño en el *periodo pre-operatorio* es incapaz de estimar correctamente las posibilidades a favor y en contra de los sucesos aleatorios, basándose en que el niño de esta edad no posee los recursos necesarios:

- la habilidad de distinguir entre el azar y lo deducible;
- el concepto de proporción;
- los procedimientos combinatorios.

Fischbein piensa que, a pesar de ello, el niño puede hacer juicios probabilísticos, en situaciones sencillas. Una de estas situaciones es pedir al niño que elija, entre dos urnas o cajas con diferente número de bolas blancas y negras, aquella que ofrezca más posibilidades de obtener una bola blanca.

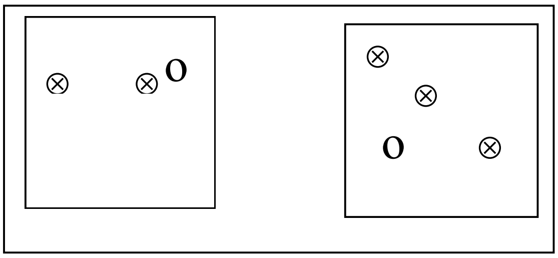
En el *periodo de las operaciones concretas* los niños pueden resolver problemas que impliquen comparación de probabilidades de un mismo suceso A en dos experimentos diferentes sólo en situaciones donde, bien el número de casos favorables o el número de casos no favorables a A son iguales en ambos experimentos (sus estimaciones se basan en comparaciones binarias). Posteriormente pasan a resolver problemas en que los casos se pueden poner en correspondencia mediante una proporción.

Los adolescentes progresan rápidamente en el cálculo de probabilidades, incluso cuando las fracciones a comparar tienen diferente denominador. Esto se observa con niños a partir de 12-13 años en incluso a partir de 10 años con la ayuda de la instrucción.

Actividad 4.6. El siguiente Ítem es un ejemplo de las tareas utilizadas en las investigaciones de Yost y otros autores sobre la capacidad de los niños para comparar probabilidades:

Item 1. En la caja A se han metido 3 fichas negras y 1 ficha blanca. En la caja B se han metido 2 fichas negras y 1 ficha blanca. (Mira el dibujo)

Si tienes que sacar una ficha negra para ganar un premio, sin mirar dentro de la caja, ¿cuál elegirías para hacer la extracción? Señala la respuesta correcta:



(A) La caja A da mayores posibilidades de obtener una ficha negra

(B) La caja B da mayores posibilidades de obtener una ficha negra

(C) Las dos cajas dan la misma posiblidad

(D) No lo se

¿Crees que puedes variar la dificultad de este problema para los niños, cambiando el número de bolas blancas y negras en cada urna? ¿Cómo?

Elabora una clasificación de estos problemas, atendiendo a la dificultad de comparación de las fracciones implicadas. Si un niño no es capaz de comparar fracciones, ¿podría resolver el problema con éxito en algunos casos? ¿Demostraría esto que el niño tiene una Intuición correcta sobre la probabilidad?

Muchos otros autores han analizado las estrategias que siguen los niños al comparar probabilidades. Reproducimos la tabla en que Cañizares (1997) hace una síntesis de las mismas:

A) Comparación del número de casos posibles: Consiste en elegir la caja que contenga mayor número de bolas	Es propia del período preoperacional. Los niños sólo tienen en cuenta el número de casos posibles de ambas cajas, sin comparar las proporciones de bolas blancas y negras.
B) Comparación del número de casos favorables: Consiste en elegir la caja que contenga más bolas del color favorable.	De los cuatro datos proporcionados en el problema, sólo se comparan dos y se ignoran los demás. Corresponde, al final del nivel preoperacional, en que el alumno no posee aún la capacidad para establecer relaciones entre el todo y sus partes.
*C) Comparación del número de casos desfavorables:*Se elige la caja con menor número de bolas del color desfavorables.	Los niños en el nivel preoperacional utilizan esta estrategia cuando existe igualdad de casos favorables, y centran su atención, entonces, sobre el número de casos desfavorables.
D) Estrategias aditivas: Se tienen en cuenta los casos favorables, los desfavorables y los posibles, pero comparan los datos por medio de alguna operación aditiva.	Es característicos del período de operaciones concretas
E) Estrategia de correspondencia: Se establece un criterio de proporcionalidad en una fracción y se aplica a la otra fracción.	A falta de un cálculo de fracciones, el sujeto determina un sistema de correspondencias cuando las proporciones o desproporciones son comparables en forma inmediata. Aparece durante el periodo de operaciones concretas, se desarrolla en el periodo de operaciones formales, para ir trasform·ndose en una estrategia puramente multiplicativa.
F) Estrategias multiplicativas: Es la más elaborada y requiere del dominio del cálculo con fracciones. Consiste en la aplicación de la regla de Laplace.	Es necesario establecer las fracciones formadas por los números de casos favorables y desfavorables para después comparar las fracciones así obtenidas. Es propia del periodo de operaciones formales
G) Otros tipos,	Hacer referencia a la suerte o elegir el color favorito.

- **Distribución y convergencia**

Piaget e Inhelder (1951) investigaron la comprensión de los niños sobre lo que ellos llamaron "distribuciones uniformes", que en realidad eran distribuciones de Poisson en el plano.

El niño tiene experiencia de observar la distribución de las gotas de lluvia sobre un embaldosado. Basándose en esta experiencia Piaget e Inhelder usan la siguiente técnica experimental: una hoja de papel blanco es dividida en cuadrados de 2 o 3 cm, y algunas fichas se lanzan sobre la hoja de papel al azar, simulando gotas de lluvia (o bien se les dan fichas para colocar sobre el embalsosado del patio o de una habitación). Se pide al niño que prevea donde caerán las gotas de lluvia sucesivas y cómo se efectuará la distribución, cuando aumentamos el número de gotas.

Los niños de *preescolar* saben que, cuando cae la lluvia, habrá gotas por todas partes, aunque ello no implica que comprendan que la distribución es, a la vez, aleatoria y cada vez más regular. En el primer estadio, el niño está convencido de la distribución regular de la lluvia. Cuando trata de reproducirla, distribuye las gotas sistemáticamente, de modo que van rellenando uno a uno todos los cuadrados, antes de repetir uno de ellos. Si la retícula tiene todos los cuadros con alguna gota, excepto un cuadro vacío, los niños colocan la gota en el cuadro vacío, de modo que se lograse un patrón uniforme. El deseo de regularidad domina las predicciones de los niños.

Al proponer a los niños del *periodo de las operaciones concretas* el problema, aceptan la irregularidad de la distribución, aunque si todos los cuadrados, menos uno tienen al menos un punto, el cuadrado "seco" se considera todavía como el más probable para recibir la siguiente gota. Es más difícil ya encontrar niños que colocan las gotas en una posición fija (por ejemplo el centro) todos los cuadrados. La comprensión de la ley de los grandes números es sólo intuitiva y empírica,

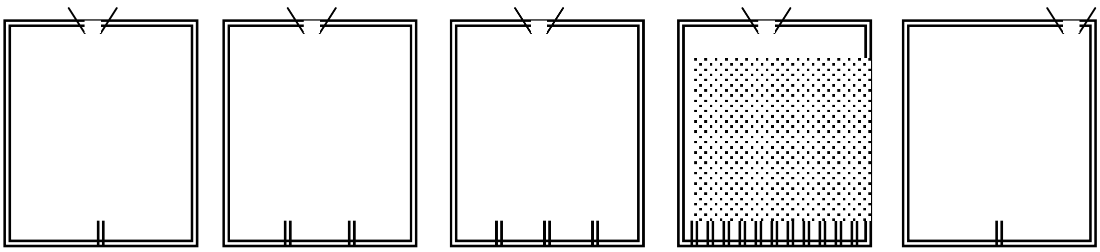
Con el *periodo de las operaciones formales* se comprende, finalmente, el mecanismo de la convergencia progresiva. En función del número cada vez mayor de gotas, la diferencia en el número de gotas en las baldosas cada vez disminuye más, no en forma absoluta, sino en forma relativa. La ley de los grandes números se comprende por su doble aspecto combinatorio y proporcional. En este estadio (12 años o más) aparece el razonamiento proporcional y Piaget e Inhelder creen que los niños comprenden la ley de los grandes números.

- **La distribución normal**

Otro problema estudiado por Piaget e Inhelder es la comprensión de la idea de distribución normal que se produce, por ejemplo cuando los granos de arena caen a través de un pequeño orificio (en un aparato de Galton o en un reloj de arena). Para Piaget, para comprender el mecanismo de esta distribución es preciso captar la simetría de las trayectorias posibles de los granos al caer por el orificio, porque hay las mismas posibilidades para cada grano de orientarse a derecha o izquierda.

Piaget indica que la representación matemática de dicha distribución corresponde a la curva de Gauss. Se emplea a menudo para hacer comprender su significado el aparato construido por Galton, el Quincux, que tiene forma de plano inclinado de madera donde se han colocado unos clavos regularmente espaciados. Al dejar caer algunas bolas por el orificio superior, se dispersan al azar, recogiéndose en unas casillas colocadas al final del plano inclinado. Las casillas centrales reciben más bolas que las extremas y la disposición es simétrica. Piaget utiliza un dispositivo semejante, pero simplificado en sus experimentos. Utiliza 5 aparatos, de los cuales cuatro tienen la abertura superior en la parte central y uno la tiene en el extremo superior derecho. Los aparatos se diferencian por el número de bifurcaciones (con 2, 3, 4 y un gran número de casillas finales en los aparatos 1 a 4 y 2 casillas finales en el aparato 5). (Ver figura 3.1)

Figura 3.1. Esquema de los aparatos de Galton utilizados en la experiencia



En cada una de las cajas se comienza introduciendo una bola, después una segunda y tercera, preguntando al niño donde cree que va a caer y por qué. Una vez comprendida la experiencia se pide al niño que explique la forma que tomaría la distribución cuando se dejasen caer un gran número de bolas. Finalmente se dejan caer las bolas y se pide al niño que interprete la distribución obtenida.

El primer estadio se caracteriza por la ausencia de la idea de distribución. En la caja I generalmente los niños hacen apuestas simples a favor de uno de los dos casilleros, pero sin la idea de igualación al aumentar el número de bolas. En la caja 2 el niño prevé bien una distribución igual en los tres casilleros o bien que todas las bolas irán a parar a uno de los casilleros. Con la caja III el niño apuesta por un de los casilleros centrales o por una distribución irregular. En la caja IV se espera en general una distribución irregular.

La estapa de operaciones concretas se caracteriza por un principio de distribución de conjunto generalizable, reconocible. El miño prevé la desigualdad entre las frecuencias laterales y centrales. Pero esta distribución permanece insuficientemente cuantificada, falta de comprensión de la ley de los grandes números. Por ello, aunque hay una simetría global, no hay aún equivalencia entre los sectores correspondientes. En la caja I el sujeto prevé una igualdad aproximada entre los dos casilleros, pero sin que esta igualdad se consiga progresivamente con el número de bolas. En la caja II se prevé un máximo central, pero sin igualdad en los casilleros laterales. La caja III da lugar a la previsión correcta de la ventaja de las casillas centrales, pero sin equivalencia entre ellas ni entre las laterales. La caja IV no provoca la previsión de una distribución simétrica regular, pero comienza a haber una generalización de las experiencias anteriores sobre la configuración de conjunto.

El tercer estadio (a partir de 12 años) está marcado por la cuantificación de la distribución de conjunto, es decir, por la previsión de una equivalencia entre las partes simétricas correspondientes de la dispersión. Esto es claro para las cajas 1 a II; la caja IV da lugar a ensayos de graduación hasta descubrir la distribución en forma de campana, el progreso más notable es la comprensión del papel de los grandes números en la regularidad de la distribución.

3.3. Investigaciones psicológicas sobre el razonamiento estocástico de sujetos adultos: heurísticas y sesgos

Sobre la década de los años setenta autores como Kahneman y Tversky contribuyen a un cambio en la forma de concebir el razonamiento no determinista, y su trabajo ha sido utilizado en muchos campos en los que la toma de decisiones es fundamental. La investigación psicológica en el campo del razonamiento parte de los estudios sobre la racionalidad humana.

En estas investigaciones se considera que una acción es racional si está de acuerdo con los valores y creencias del individuo, es decir, si es consistente con un conjunto de axiomas (modelos normativos). Por ejemplo, en un problema simple de decisión, donde de diversas acciones se deducen diferente consecuencias, se consideraría normativo si la decisión maximiza la utilidad esperada.

En la psicología de la decisión, la investigación se ha centrado en la descripción y explicación de las discrepancias entre la conducta que debieran tener los sujetos, si razonasen de acuerdo con los modelos normativos y los juicios y decisiones tomadas en la práctica. En los últimos años ha habido un giro: Se pasa de concebir el rasonamiento humano como estrictamente normativo al punto de vista de que nuestro razonamiento no está siempre basado en normas lógicas. Pueden distinguirse dos líneas en este debate:

1. *Los investigadores "pesimistas"* defienden que las decisiones y juicios bajo incertidumbre muestran a menudo errores serios y sistemáticos, debido a las características intrínsecas al sistema cognitivo humano. Por ejemplo, Simon (1955) habló de "*racionalidad limitada"*. Estos investigadores dan tres tipos de explicación para los sesgos que se encuentran con frecuencia al resolver problemas de probabilidad:

- El uso de "heurísticas" o estrategias inconscientes que suprimen parte de la información del problema;

- Los errores representacionales, que son debidos a ilusiones en la percepción, y causan deficiencias al percibir los problemas de decisión;
- La falta de motivación al buscar y seleccionar la información.

2. *Los investigadores "optimistas",* por su lado, piensan que en general los juicios y decisiones son bastante eficientes y funcionales, incluso en situaciones complejas. Para ello se basan en las siguientes explicaciones:

- El argumento de *metarracionalidad:* La decisión que, aparentemente, viola los principios de racionalidad, puede ser perfectamente racional, si se tiene en cuenta el coste cognitivo de la toma racional de decisiones, frente a los posibles beneficios al aplicar un razonamiento más intuitivo;
- La decisión puede ser totalmente racional, pero el individuo hace una interpretación del problema que no coincide con la que hace el investigador.

Entre todas estas teorías, la que tiene más implicaciones educativas es la que se refiere al empleo de *heurísticas* (Kahneman *et al.* , 1982), que se han basado en las *teorías cognitivas* y del *procesamiento de la información*. En estos estudios las *heurísticas* son procesos mentales que reducen la complejidad de un problema, de modo que sea accesible al resolutor. Aunque nos llevan a dar una solución inmediata del problema, no garantizan que la solución sea correcta. Pérez Echeverría (1990) las describe como :

> *"mecanismos por los que reducimos la incertidumbre que produce nuestra limitación para enfrentarnos con la complejidad de estímulos ambientales"* (pág. 51).

Se diferencian de los *algoritmos* en que son generalmente automáticas y se aplican *inconscientemente* sin pensar si es adecuada o no al juicio a realizar. Por el contrario, el algoritmo se aplica paso a paso en una forma concreta, establecida por criterios dados y produce la solución para cualquier problema, dentro de una clase dada.

En otros campos de investigación (didáctica, Matemáticas, Psicología, Inteligencia Artificial y Toma de Decisiones) el término "heurística/o" se aplica en forma diferente. En Matemáticas se encuentra el término en los comentarios a la obra de Euclides por Pappus de Alejandría y en Descartes, Leibnitz o Bolzano. Polya (1982) lo usa para referirse a la comprensión del método que conduce a la solución de problemas y más concretamente a las operaciones mentales típicamente útiles en este proceso. Para él la heurística se construye sobre una experiencia resultante, al mismo tiempo, de la solución de problemas y de la observación de los métodos de otras personas.

Tversky y Kahneman tomaron la idea de heurística para analizar los razonamientos probabilísticos, estudiando los errores y sesgos típicos al aplicar las heurísticas. En cuanto a la clasificación del tipo de heurística se describen tres principales: *representatividad, disponibilidad y ajuste y anclaje* que estudiaremos en la sección dedicada al muestreo.

Actividad 3.7. Elabora una lista de profesiones en las cuales sea precisa la toma de decisiones en ambiente de incertidumbre. Analizar las posibles consecuencias de estas decisiones y el impacto local y global de las mismas en su esfera de acción.

Actividad 3.8. Analizar el enunciado del ítem siguiente y estudio las posibles interpretaciones erróneas del enunciado que pudieran llevar a un niño a dar como correcta la respuesta A
La probabilidad de que un niño nazca varón es aproximadamente 1/2. ¿Cuál de las siguientes secuencias de sexos es más probable que ocurra en seis nacimientos?

a) VHHVHV; b) VHHHHV; c) las dos son igual de probables.

Actividad 3.9. Se aplica una prueba de tuberculina en una población, conociéndose que el test da resultado positivo en 1 de cada 10.000 personas sanas en 99 de cada 100 personas enfermas. La incidencia de enfermos en la población es 3 de cada 100.000 personas. ¿Qué pensarías si te haces la prueba y resulta positiva? Analiza en este ejemplo la importancia del razonamiento estocástico correcto en la toma de decisiones.

3.4. Investigaciones didácticas: errores, obstáculos y concepciones

Gran parte de la investigación en Didáctica de la Matemática, se interesa por explicar por qué el alumno se equivoca cuando se le pide realizar ciertas tareas. Los profesores y los mismos alumnos son conscientes de que a veces obtienen respuestas erróneas, o no son capaces de dar ninguna respuesta.

En los casos en que no se trata de una simple distracción u olvido, se dice que tal tarea resulta demasiado *difícil* para el alumno en cuestión. Pero las dificultades no se presentan de un modo aleatorio, imprevisible. Con frecuencia encontramos *errores* que se repiten, o que se producen regularidades, y asociaciones con variables propias de las tareas propuestas, de los sujetos o de las circunstancias presentes o pasadas. La investigación didáctica trata de identificar y explicar estos errores, que frecuentemente son debidos a las creencias de los estudiantes.

El término usado en Didáctica para referirse a estas creencias de los estudiantes, que pueden no coincidir con lo que les tratamos de enseñar es el de *"concepción"*. Con esta palabra se hace alusión a los diferentes puntos de vista (correctos o incorrectos) que es posible mantener sobre un mismo concepto matemático. Un principio ampliamente asumido en psicología educativa es el enunciado por Ausubel y cols. (1983):

> *"el factor más importante que influye en el aprendizaje es lo que el alumno ya sabe. Averigüese esto y enséñese consecuentemente*".

El interés de los estudios de didáctica por las concepciones de los estudiantes (Confrey, 1990) sería una consecuencia del mencionado principio psicológico. Un problema didáctico es que algunas de estas concepciones, que permiten resolver correctamente algunas tareas, son inapropiadas para otros problemas más generales, y que el sujeto muestra una resistencia a cambiar esta concepción. En estas circunstancias se habla de la existencia de un *obstáculo*. Brousseau (1983) describe las siguientes características de los obstáculos:

- *Un obstáculo es un conocimiento, no una falta de conocimiento*. El alumno utiliza este conocimiento para producir respuestas adaptadas a un cierto contexto que encuentra con frecuencia. Cuando se usa este conocimiento fuera de este contexto genera respuestas incorrectas. Una respuesta más general exigirá un punto de vista diferente.
- *El alumno no es consciente del obstáculo* y no logra establecer un conocimiento mejor.

Es indispensable que el alumno se consciencie del obstáculo, y que lo rechace, para adquirir un nuevo conocimiento más amplio

- Después de que el alumno se ha dado cuenta de que existía un obstáculo*, continúa manifestándolo, de forma esporádica.*

Brousseau ha identificado tres tipos de obstáculos:

- *Obstáculos ontogénicos* (a veces llamados obstáculos psicogenéticos): son debidos a las características del desarrollo del niño. Por ejemplo, para comprender la idea de probabilidad se requiere un cierto razonamiento proporcional, por lo que un niño muy pequeño no puede comprender la probabilidad.
- *Obstáculos didácticos:* resultan de alguna forma inadecuada de enseñar un concepto. Por ejemplo, la introducción de un nuevo simbolismo tal como:
$$(\Sigma_{i=1}{}^{n}\ x_i\ )/n$$
para la media, en lugar de facilitar la comprensión del significado de la media, puede oscurecerlo, en caso de alumnos con poca base matemática. Sería mejor con estos alumnos trabajar con ejemplos concretos o comparaciones, tal como comparar la media con el punto de equilibrio o centro de gravedad.
- *Obstáculos epistemológicos*: Relacionados intrínsecamente con el propio concepto y conteniendo parte del significado del mismo. Por ejemplo, las circularidades que se presentan en las diferentes definiciones del significado de la probabilidad (clásica, frecuencial, subjetiva) que mostraron en su día la necesidad de una definición axiomática.

Encontrar estos obstáculos, y ayudar al alumno a superarlos parece ser una condición necesaria para la construcción de una concepción adecuada. Otras dificultades experimentadas por los estudiantes se deben a una falta del conocimiento básico necesario para una comprensión correcta de un concepto o procedimiento dado.

Actividad 3.10. Una creencia frecuente por parte de los estudiantes es que después de una racha de obtención de caras al lanzar una moneda, la probabilidad de obtener una cruz aumenta. Analiza los posibles obstáculos implicados en este razonamiento y su tipología.

Actividad 3.11. Busca ejemplos de obstáculos epistemológicos ligados al desarrollo histórico de algún concepto estocástico fundamental. ¿Cómo se resolvieron?

Actividad 3.12. Da algunos ejemplos de obstáculos de tipo cognitivo para el estudio de la probabilidad por parte de niños en el periodo 8-14 años.

Actividad 3.13. Analiza el desarrollo del tema de estadística en un libro de texto de ESO. ¿Es el vocabulario usado adecuado? ¿Podría algún término o definición empleada constituir un posible obstáculo didáctico?

3.5. Significado y comprensión

Cuando queremos reflexionar sobre la dificultad que el aprendizaje de ciertos conceptos tiene para los alumnos, es necesario comenzar por hacer un análisis epistemológico de su significado. Como indica Godino (1996),

> "*el problema de la comprensión está intimamente ligado a cómo se concibe el propio conocimiento matemático. Los términos y expresiones matemáticas denotan entidades abstractas cuya naturaleza y origen tenemos que explicitar para poder elaborar una teoría útil y efectiva sobre qué entendemos por comprender tales objetos. Esta explicitación requiere responder a preguntas tales como: ¿Cuál es la estructura del objeto a comprender? ¿Qué formas o modos posibles de comprensión existen para cada concepto? ¿Qué aspectos o componentes de los conceptos matemáticos es posible y deseable que aprendan los estudiantes en un momento y circunstancias dadas? ¿Cómo se desarrollan estos componentes?* (pg. 418).

En lo que sigue describimos un modelo de la actividad matemática propuesto por Godino y Batanero (1994, 1997), en el que se problematiza la naturaleza de un objeto matemático, partiendo del supuesto de que un mismo término o expresión matemática, puede tener distinto significado para diferentes personas o instituciones. Usaremos el concepto de media para contextualizar la discusión, partiendo del trabajo de Batanero (2000).

Se parte de la *situación-problema* como noción primitiva, considerandola como cualquier circunstancia en la que se debe realizar actividades de matematización, definidas por Freudenthal (1991) en la forma siguiente:

- Construir o buscar soluciones de un problema que no son inmediatamente accesibles;
- Inventar una simbolización adecuada para representar la situación problemática y las soluciones encontradas, y para comunicar estas soluciones a otras personas;
- Justificar las soluciones propuestas (validar o argumentar);
- Generalizar la solución a otros contextos, situaciones-problemas y procedimientos.

Cuando una clase de situaciones-problemas comparten soluciones, procesos, etc, hablamos de un *campo de problemas*. En el caso de la media, podemos considerar el siguiente problema:

> *Ejemplo 3.5. Un objeto pequeño se pesa con un mismo instrumento por ocho estudiantes de una clase, obteniéndose los siguientes valores en gramos: 6'2, 6'0, 6'0, 6'3, 6'1, 6'23, 6'15, 6'2 ¿Cuál sería la mejor estimación del peso real del objeto?*

Si planteamos este problema a los alumnos de secundaria, la mayoría sumará los valores y dividirá por ocho para obtener el valor 6'1475. Este es un ejemplo particular de una clase de problemas: *estimación de una cantidad desconocida, en presencia de errores de medida*. En muchas situaciones necesitamos medir una cantidad X desconocida de una cierta magnitud. Pero debido a la imperfección de nuestros instrumentos, en mediciones sucesivas obtenemos distintos números como medidas de X. No tenemos ninguna razón para pensar que el verdadero valor esté más cercano a uno u otro de los datos obtenidos.

El problema consiste en determinar, a partir de un conjunto de medidas $x_1, x_2, \ldots, x_n$ la mejor estimación posible del verdadero valor X desconocido. Según Plackett

(1970), los astrónomos de Babilonia resolvieron el problema calculando la suma total de las observaciones y dividiendo por el número de datos y esta práctica se ha conservado hasta nuestros días, dando origen a lo que hoy conocemos por media aritmética.

Para enunciar y resolver el problema enunciado los sujetos utilizan representaciones simbólicas de los objetos matemáticos abstractos (números, operaciones, ...). Por ejemplo, es una práctica habitual usar la expresión (3.1) para expresar la solución del problema en su enunciado general. En esta expresión los símbolos representan el número de datos, los valores obtenidos en las distintas mediciones, su suma, la división y el resultado obtenido:

(3.1) $$\overline{x} = (x_1 + x_2 + \ldots + x_n)/n$$

Una característica de la actividad matemática es que, una vez hallada la solución de un problema, se trata de extender esta solución a otros casos diferentes de la situación concreta particular. En relación al problema se puede generalizar la expresión (3.1) para un valor arbitrario *n*, posteriormente a un número infinito de valores, en variables discretas o continuas, obteniéndose las expresiones (3.2) y (3.3).

(3.2) $$E(X) = \sum x_i p_i = \mu$$

(3.3) $$E(X) = \int x f(x) dx$$

Sumar un conjunto dado de valores y dividir por el número de valores, sumar una serie, hallar una integral o escribir las expresiones anteriores son ejemplos de *prácticas* matemáticas, es decir de acciones llevadas a cabo para resolver problemas matemáticos, comunicar a otros la solución, mostrar que la solución es correcta y generalizarla a otros contextos y problemas.

Aunque en cada problema concreto de estimación de la magnitud de interés, el instrumento de medición, el número de medidas tomadas y los valores concretos obtenidos varían, las expresiones (3.1) a (3.3) son aplicables de forma general para el cálculo de la mejor estimación del valor desconocido. Estas prácticas han dado lugar poco a poco al concepto que hoy conocemos como "media aritmética", primeramente como útil implícito en la solución de problemas prácticos, más tarde como objeto de estudio en sí mismo. El estudio y caracterización de sus propiedades llevó progresivamente a la aplicación del concepto en la solución de otras situaciones problemáticas como las siguientes:

Ejemplo 3.6. Unos niños llevan a clase caramelos. Andrés lleva 5, María 8, José 6, Carmen 1 y Daniel no lleva ninguno. ¿Cómo repartir los caramelos de forma equitativa?

En el ejemplo 3.2 y otros problemas similares se necesita obtener una *cantidad equitativa* a repartir para conseguir una distribución uniforme, como en el ejemplo, se toma la media aritmética. Este tipo de problemas surgen con frecuencia al obtener la "renta per cápita", la velocidad media durante un viaje o la calificación final en un examen compuesto de varios exámenes parciales.

Ejemplo 3.7. Al medir la altura en cm. que pueden saltar un grupo de escolares, antes y después *de haber efectuado un cierto entrenamiento deportivo, se obtuvieron los valores siguientes. ¿Piensas que el entrenamiento es efectivo?*

Altura saltada en cm.										
Alumno	*Ana*	*Bea*	*Carol*	*Diana*	*Elena*	*Fanny*	*Gia*	*Hilda*	*Ines*	*Juana*
Antes del entrenamiento	*115*	*112*	*107*	*119*	*115*	*138*	*126*	*105*	*104*	*115*
Después del entrenamiento	*128*	*115*	*106*	*128*	*122*	*145*	*132*	*109*	*102*	*117*

El ejemplo 3.7 muestra otra aplicación típica de la media, que consiste en servir de *elemento representativo* de un conjunto de valores dados x_i, cuya distribución es aproximadamente simétrica. En el ejemplo 3.7. usaríamos la altura media saltada antes y después del entrenamiento para ver si éste ha producido algún efecto.

Para representar un conjunto de datos se toma la media por sus propiedades de localización central, por ser "centro de gravedad" del espacio de valores muestrales o poblacionales. Si la distribución es muy asimétrica, el valor más frecuente (Moda) o el valor central en el conjunto de datos ordenados (Mediana) podría ser más representativo. Vemos que cuando añadimos condiciones a un campo de problemas surgen conceptos relacionados con el de interés con el cual guardan diferencias y semejanzas, que es necesario investigar. De los primitivos problemas extramatemáticos, pasamos posteriormente a problemas internos a la misma matemática, como estudiar las diferentes propiedades de las medidas de posición central.

Ejemplo 3.8. *La altura media de los alumnos de un colegio es 1'40. Si extraemos una muestra aleatoria de 5 estudiantes y resulta que la altura de los 4 primeros es de 1'38, 1'42, 1'60, 1'40. ¿Cuál sería la altura más probable del quinto estudiante?*

En otras ocasiones se necesita conocer el valor que se obtendrá con mayor probabilidad al tomar un elemento al azar de una población. Por ejemplo, al predecir la esperanza de vida o el beneficio esperado en una inversión en bolsa, se toma la media de la variable en la población como predicción, como valor esperado, por sus propiedades muestrales derivadas del teorema central del límite. Del concepto de valor esperado se derivan muchos modelos de predicción, como los distintos tipos de regresión. Así, cuando predecimos el peso de una persona en función de su altura, usamos el peso promedio de todas las personas que en la población tienen la altura dada.

Problemas como los planteados y otros problemas, primero prácticos más tarde teóricos, han llevado a la definición del concepto de media, a la identificación de sus propiedades, más tarde a la definición de otras medidas de posición central, como la mediana o moda, que son preferibles a la media en algunas situaciones concretas. Además, ha sido necesario "probar" o "demostrar" la validez de estas soluciones y propiedades, para aceptarlas como parte del conocimiento matemático.

- **Elementos de significado**

Por tanto, cuando nos preguntamos por el *significado* de la media o de las medidas de posición central, observamos que este significado tiene un carácter complejo. En el trabajo matemático se pueden identificar los siguientes tipos de entidades:

- *Problemas y situaciones* que inducen actividades matemáticas y definen el campo de problemas de donde surge el objeto. En nuestro caso los problemas presentados como ejemplos y sus generalizaciones formarían parte del campo de problemas asociados a las medidas de posición central;

- *Procedimientos, algoritmos, operaciones*. Cuando un sujeto se enfrenta a un problema y trata de resolverlo, realiza distintos tipos de *prácticas*, que llega a convertir en rutinas con el tiempo. Prácticas características en la solución de problemas de promedios serían sumar una serie de valores y dividir por el número de sumandos, encontrar el valor más frecuente en una tabla de frecuencias, calcular las frecuencias acumuladas y hallar el valor al que corresponde la mitad del número total de datos, o integrar el producto de la variable por la función de densidad en un cierto dominio;
- *Representaciones* materiales utilizadas en la actividad matemática (términos, expresiones, símbolos, tablas, gráficos). Las notaciones, gráficos, palabras y en general todas las representaciones del objeto abstracto; como los términos "media", "valor medio", "promedio", $E(X)$, $\sum x_i p_i$ μ, $\int xf(x)dx$, que podemos usar para referirnos al concepto;
- *Abstracciones* (conceptos, proposiciones). Las definiciones y propiedades características y sus relaciones con otros conceptos. Por ejemplo, en la investigación de Strauss y Bichler (1988) encuentra una proporción importante de niños de entre 8 y 12 años eran capaces de comprender y aplicar adecuadamente las propiedades a) c) y d) siguientes de la media, mientras que el resto de ella resultaron demasiado abstractas:

a) La media es un valor comprendido entre los extremos de la distribución;
b) La suma de las desviaciones de cada valor a la media es igual a cero;
c) El valor medio es influenciado por los valores de cada uno de los datos;
d) La media no tiene por qué ser igual a uno de los valores de los datos;
e) El valor obtenido de la media de números enteros puede ser una fracción, que no tenga sentido en el contexto de los datos;
f) Hay que tener en cuenta los valores nulos en el cálculo de la media;
g) El valor medio es representativo de los valores promediados.

- *Demostraciones* que empleamos para probar las propiedades del concepto y que llegan a formar parte de su significado y los argumentos que empleamos para mostrar a otras personas la solución de los problemas.

- **Dimensiones institucional y personal del conocimiento**

En general los problemas no aparecen de forma aislada, sino que los mismos problemas son compartidos dentro de cada institución, y las soluciones encontradas dependen de los instrumentos y prácticas sociales disponibles.

Así, problemas de estimación de una cantidad desconocida son compartidos en instituciones de investigación experimental, como la física, astronomía o la agronomía y también en las instituciones escolares, pero los instrumentos disponibles son muy diferentes en uno y otro caso, de modo que el significado de un concepto matemático varía según la institución considerada.

Los matemáticos y estadísticos profesionales constituyen una institución interesada en resolver problemas de promedios, pero existen otras instituciones diferentes que también podrían estar interesadas en la media, aunque podría atribuirle un significado más restringido al que recibe dentro de la matemática, por ejemplo:

(I1) En la escuela primaria los curriculos proponen que se enseñe a los alumnos:

- la definición de la media, mediana y moda en el caso más simple, empleando una notación sencilla (se evita el sumatorio y la ponderación);
- algunos ejemplos de aplicación, limitando el cálculo de las medidas de tendencia central a conjuntos sencillos de datos, y haciéndolo manualmente o con calculadora.
- discriminación respecto de otras medidas de tendencia central (mediana, moda).

(I2) En la escuela secundaria (y en la universidad) se amplia la definición de la media, trabajándose primero con medias ponderadas y luego con medias de variables aleatorias discretas y continuas. Se enuncian y demuestran algunas propiedades de los promedios y se presentan aplicaciones a situaciones problemáticas más realistas y complejas.

En la universidad se introduce la noción de media o esperanza matemática de una distribución de probabilidad y se muestra que la media es un parámetro que define algunas distribuciones de probabilidad, como la normal; al iniciar el estudio de la inferencia, distinguimos varias medias: media de la muestra, media de la población, media de la media muestral en todas las muestras de tamaño dado.

(I3) En la "vida diaria" encontramos la media en los medios de comunicación y el trabajo profesional, por ejemplo, cuando analizamos los números índices de la evolución de la bolsa, precios, producción, empleo y otros indicadores económicos.

Además, al considerar una cierta institución escolar, como la escuela primaria, el significado construido por un alumno particular, en un momento del proceso de aprendizaje puede no corresponder exactamente al significado del objeto en la institución dada, por lo que conviene distinguir entre *significado institucional* y *significado personal* de un objeto matemático.

Por otro lado, el significado de un objeto matemático se configura y evoluciona a lo largo del tiempo, según se va ampliando el campo de problemas asociado. Por ejemplo, el conocimiento sobre la media no ha sido siempre igual al actual, sino que se ha desarrollado lentamente a lo largo del tiempo, ya que a medida que se han ido resolviendo problemas progresivamente diferentes y más complejos.

Ejemplo 3.9. Hay 10 personas en un ascensor, 4 mujeres y 6 hombres. El peso medio de las mujeres es de 60 kilos y el de los hombres de 80. ¿Cuál es el peso medio de las 10 personas del ascensor?

No podemos ahora resolver este problema por medio de la media aritmética simple (60+80)/2=70, sino que necesitaríamos ampliar el concepto al de media ponderada: (60x4+70x6)=72. Como cualquier otro concepto, la media y otras medidas de tendencia central han tenido un lento desarrollo dentro de la matemática hasta el momento en que fueron reconocidos como conceptos matemáticos e incluidos en la enseñanza.

Este carácter progresivo de la construcción de los objetos en la ciencia tiene su paralelismo en el aprendizaje del sujeto que es un proceso lento y costoso. Para las medidas de posición central no hay todavía un estudio comprehensivo del desarrollo a diversas edades, aunque el trabajo de Watson y Moritz (2000) es un primer paso en este estudio. Como veremos en nuestra exposición, en realidad estos conceptos son bastante elaborados, de modo que el conocimiento que un sujeto puede adquirir fuera del ámbito escolar es necesariamente muy limitado y restringido. Ello posiblemente haya influido en la falta de interés por el desarrollo de los conceptos estadísticos por parte de la psicología.

- **Significado y comprensión**

Paralela a la noción de significado que hemos descrito se deduce una teoría de la comprensión (Godino, 1996). La comprensión personal de un concepto es, en este modelo, la "captación" del significado de dicho concepto. Ahora bien, puesto que el significado de un objeto no se concibe como una entidad absoluta y unitaria sino compuesta y relativa a los contextos institucionales, la comprensión de un concepto por un sujeto, en un momento y circunstancias dadas, implicará la adquisición de los distintos elementos que componen los significados institucionales correspondientes.

En la clase de matemáticas el profesor sigue las directrices curriculares, los libros de texto y materiales didácticos -que marcan un significado particular restringido para la media y las medidas de posición central. Al realizar la evaluación, el profesor considera que el alumno "conoce" o "comprende" las medidas de tendencia central si hay un ajuste entre el significado institucional y el personal construido por el sujeto.

Godino (1996) indica que la comprensión deja de ser meramente un proceso mental y se convierte en un proceso social ya que podemos considerar que un alumno "comprende" suficientemente los promedios desde el punto de vista de la enseñanza secundaria y que no lo comprende desde el punto de vista de unos estudios universitarios.

3.6. Significado subjetivo de la aleatoriedad

Recientemente, numerosos investigadores, tanto en psicología como en educación matemática han tratado de evaluar el significado atribuido a la aleatoriedad por niños y adultos, planteándoles problemas y centrándose en la determinación de las propiedades que se atribuyen a este concepto. Podemos clasificar en dos tipos los tipos de problemas planteados:

- En el primero (*problemas de generación*) se pide a los sujetos simular una secuencia de resultados aleatorios. Por ejemplo, se pide escribir puntos al azar en un folio o escribir una sucesión de dígitos aleatorios.
- En las *problemas de reconocimiento* se pregunta a los participantes si unas ciertas situaciones, secuencias o patrones espaciales son o no aleatorios.

Estas investigaciones muestran que los sujetos tienden a encontrar patrones deterministas en las situaciones aleatorias, es decir, tratan de encontrar asociaciones inexistentes, con objeto de reducir la incertidumbre. También tendemos a inferir aleatoriedad en situaciones en la que no está presente. En general se observa la tendencia a generar rachas cortas de dos o tres símbolos adyacentes en algún sentido, por ejemplo números consecutivos o letras sucesivas del alfabeto. También se produce un exceso de alternancias o *"recencia negativa"* que consiste en reproducir la frecuencia esperada del suceso con demasiada exactitud, incluso en rachas cortas.

En el caso de emplear experimentos con más de dos resultados, rara vez se obtiene una distribución uniforme. Por ejemplo, al pedir a los sujetos secuencias de dígitos aleatorios, éstos emplean los dígitos 0 y 1 con menor frecuencia que el resto. También los números centrales (como 4, 5 o 6) aparecen con mayor frecuencia que los valores extremos. Falk (1981) hizo un estudio con alumnos de secundaria, usando secuencias de cartas verdes y amarillas, preguntándoles si eran o no aleatorias. Los estudiantes tendían a considerar aleatorias las secuencias con más alternancias entre los dos colores, aunque de hecho es más probable que una secuencia aleatoria contenga algunas rachas largas de un mismo color. Obtuvo el mismo resultado al pedir a los estudiantes que escribiesen una secuencia aleatoria de longitud dada.

Actividad 3.14. Un problema clásico de probabilidad es el siguiente:¿Cuál es el mínimo número de personas en una habitación para que la probabilidad de que dos de ellos al menos celebren el cumpleaños el mismo día sea mayor que 1/2?
Parece sorprendente que la respuesta sea n=23. Analiza esta situación desde el punto de vista de sus implicaciones sobre la percepción subjetiva de la probabilidad y de la confusión entre el experimento aleatorio implicado y otros posibles experimentos aleatorios.

Actividad 3.15. Recopila en la prensa noticias referentes a los juegos de azar y su incidencia social, en particular, sobre el fenómeno de las ludopatías. ¿Crees que estos datos apoyan las investigaciones psicológicas sobre la creencia de los sujetos en la posibilidad de control de lo aleatorio?

Dos sucesiones aleatorias de igual longitud obtenidas a partir del mismo experimento, cuyos sucesos elementales son equiprobables, son también equiprobables. Sobre esta base, han de ser consideradas igualmente aleatorias. Sin embargo, al considerar diferentes atributos de la sucesión. como, por ejemplo, el número de rachas, una de las sucesiones podría ser consideradas como más característica de un proceso estocástico que otra por los sujetos.

Consideremos las secuencias (3.4) y (3.5):

(3.4) O X O X O X O X O O O X X X X O X O X O O
(3.5) O X O X X X X X O X O O O O X O O O X X O

Matemáticamente tienen igual probabilidad $(½)^{21}$, como cualquier otra secuencia ordenada de la misma longitud aleatoriamente producida de lanzamientos de una moneda no sesgada. Sobre esta base, se las podrían juzgar igualmente aleatorias. Por otro lado, si consideramos algunos atributos específicos de las secuencias, hay un mayor número de posibles secuencias semejantes a (3.5) que semejantes a (3.4), por lo que, en este sentido, se podría considerar más característica de un proceso aleatorio la segundaa

Uno de estos atributos es la *probabilidad de alternancia* entre los dos tipos de sucesos simbolizados. La probabilidad de cambio en una secuencia particular P(A), se obtiene al dividir el número de cambios actuales del símbolo tipo por n-1. Cuando las probabilidades de los dos símbolos son iguales, el valor de P(A) en muestras aleatorias tiende a 0,5. Este resultado es consecuencia del principio de independencia; independientemente de lo que ya haya ocurrido en la secuencia, la probabilidad de que el carácter siguiente a uno dado sea diferente del anterior es 0,5. Las secuencias con valores de probabilidad distinta que 0,5 ocurren con menos frecuencia. Los valores de P(A) de las sucesiones dadas son aproximadamente 0,7 y 0,5 respectivamente. Por tanto, la segunda sucesión se acercaría más a lo que esperamos en una secuencia aleatoria.

Por otro lado, la primera secuencia podría codificarse como 4 OX 3 O 4 X 2 OX 2 O, dando una medida de complejidad 12/21 = 0,57. Hay 18 caracteres en el código de la segunda, solo algo menos de los 21 caracteres originales; esta medida de complejidad 18/21 = 0,86 es muy cercana al valor máximo de 1.

Pese a estas razones matemáticas, las investigaciones psicológicas han mostrado que la mayoría de los sujetos tienen una concepción opuesta de las secuencias aleatorias. y prefiere secuencias que incluyan más alternancias de las que ocurren típicamente, como la secuencia primera. La conocida falacia del jugador, según la cual las cruces son consideradas más probables que las caras después de una serie de caras consecutivas, puede estar basada en esta creencia.

Konold y sus colaboradores (1991 a) exploran las diferencias entre la concepción de experimento aleatorio en personas sin instrucción en probabilidad y expertos y argumentan que, de hecho, el término "aleatoriedad" comprende una familia de conceptos. Clasifican las concepciones de los sujetos sobre los fenómenos aleatorios en las siguientes categorías:

- Sujetos para los que un fenómeno es aleatorio sólo si los posibles resultados son igualmente probables; esta concepción es errónea porque hay experimentos aleatorios con sucesos no equiprobables. En esta concepción no se consideraría aleatorio
- Sujetos para los que un fenómeno aleatorio es aquél con múltiples posibilidades;
- Aleatoriedad como falta de causalidad; ya hemos discutido que esta acepción, en general no es correcta.
- Aleatoriedad como incertidumbre;
- Aleatoriedad como modelo matemático para representar ciertos fenómenos. Esta sería la concepción del estadístico

Actividad 3.16. Pide a un amigo que escriba una lista de 5 fenómenos que consideres aleatorios y 5 que no considere aleatorios y, a partir de ellos, describe las creencias de tu amigo sobre la aleatoriedad. ¿En cuáles de las categorías de Konold podrías colocar a tu amigo?

Actividad 3.17. Pide a cinco amigos que te escriban cada uno 2 secuencias de resultados que esperarían obtener al lanzar una moneda equilibrada 10 veces. Estudia la tasa de complejidad de las secuencias obtenidas. ¿Estás de acuerdo con la teoría de Konold y sus colaboradores que la gente considera aleatorias las secuencias sólo si su grado de complejidad es alto?

- **Generación de resultados aleatorios**

Green (1991) pidió a niños de 7 a 11 años escribir una sucesión de caras y cruces que represente lo que ellos creen que se obtendría al lanzar 50 veces una moneda equilibrada. Identifica tres aspectos básicos en las sucesiones generadas: la frecuencia relativa de cada uno de los sucesos, la independencia de los elementos de la secuencia y la consistencia entre las dos mitades de la secuencia generada.

El estudio probó que son muy exactos al reflejar la equiprobabilidad. Además producen sucesiones cuya primera y última mitad son altamente consistentes; quizás demasiado consistentes para reflejar la variabilidad aleatoria.

Para analizar la independencia, Green estudió la racha de mayor longitud en las secuencias escritas por los niños. Observó que había casos en que la racha de mayor longitud es de longitud 1 o bien 25 o más signos iguales consecutivos. Asimismo se basé en el número total de rachas en cada una de las sucesiones generadas por los alumnos (relacionado con el contraste anterior, aunque no idéntico).

Green observo que la longitud media de la racha más larga en su muestra era inferior al valor teórico, en niños de todas las capacidades y de todas las edades. También comparó el número de rachas de las distribuciones teóricas y de las generadas por los alumnos, encontrando que, en general, los alumnos producen demasiadas rachas de longitud corta.

- **Reconocimiento de resultados aleatorios**

Las investigaciones de Green (1983) y Toohey (1995) incluyen una pregunta en la cual los niños deben diferenciar entre una sucesión generada por un mecanismo aleatorio y otra que no cumple esta condición. Como resultado, descubre que la mayor parte de los niños elige precisamente la secuencia no aleatoria y que no mejora la apreciación de la aleatoriedad con la edad. Entre las razones aducidas por los niños para justificar su decisión, encontró las siguientes:

- *Razones correctas*: El patrón de la sucesión es demasiado regular para ser aleatorio; Demasiadas rachas o rachas demasiado cortas
- *Razones incorrectas*: no se apreciaba suficientemente la irregularidad de una sucesión aleatoria; se espera que la frecuencia relativa fuera exactamente el 50% de caras y cruces o un valor muy próximo; rachas demasiado largas; no se admite la posibilidad de este tipo de rachas en una sucesión aleatoria.

Una sorpresa es que son mejores las respuestas en los niños más jóvenes, de los que casi la mitad seleccionó el patrón aleatorio. En general encontró que los niños más inteligentes seleccionaban con mayor frecuencia que sus compañeros el patrón regular. Estos dos hechos le plantearon serias dudas sobre la relevancia de los estadios de desarrollo cognitivo sugeridos por la teoría de Piaget.

Todos estos resultados son replicados y completados por Batanero y Serrano (1999), quienes sugieren que los alumnos atribuyen diferentes significados a la aleatoriedad y algunos de ellos coinciden con los admitidos en diferentes periodos históricos dentro de la estadística, por ejemplo:

- Aleatoriedad como inexistencia de causas o causa desconocida;
- Aleatoriedad como equiprobabilidad;
- Aleatoriedad como estabilidad de las frecuencias relativas;
- Aleatoriedad como impredecibilidad.

Cada una de estas concepciones recoge propiedades parciales del concepto y por ello puede ser válida en unas situaciones e incompleta en otras más complejas. Es importante que en la clase el profesor presente a los alumnos ejemplos variados de situaciones aleatorias, como las que hemos mostrado a lo largo de esta sección para ayudar a los alumnos a una construcción progresiva del concepto.

- **El sesgo de equiprobabilidad**

Lecoutre (por ejemplo, Lecoutre, 1992) describe la creencia de los sujetos en la equiprobabilidad de todos los sucesos asociados a cualquier experimento aleatorio. Como ejemplo, usan un problema en el que se pregunta si al lanzar dos dados hay la misma probabilidad de obtener un 5 y un 6 que la de obtener dos veces un 5. A pesar de variar el contexto y el formato de la pregunta, los resultados siempre coinciden y demuestran la estabilidad de la creencia en que los dos resultados son equiprobables. Lecoutre y sus colaboradores defienden que ello no es debido a la falta de razonamiento combinatorio, sino a que los modelos combinatorios no se asocian fácilmente con las situaciones en que interviene "el azar". Los alumnos a los que se les pasó la prueba consideran que el resultado del experimento "depende del azar" y en consecuencia todos los posibles resultados son equiprobables.

Actividad 3.18. Analiza los razonamientos subyacentes en las diferentes respuestas al ítem siguiente: Cuándo lanzamos tres dados simultáneamente, ¿Cuál de los siguientes resultados en más verosímil que ocurra?

a) Un 5, un 3 y un 6
b) Un 5 tres veces
c) Un 5 dos veces y un 3
d) Los tres resultados son equiprobables
e) Es imposible saberlo

¿Qué respuestas esperarías de los alumnos si preguntamos en el ítem anterior cuál de los resultados es menos verosímil?¿Cómo podríamos cambiar el contexto del ítem anteriores para hacer variar la respuesta?Diseña una situación de aprendizaje en que se haga ver a los alumnos que la regla de Laplace no se puede aplicar en todas las situaciones aleatorias.

- **Independencia**

Otro concepto implicado en la comprensión de la aleatoriedad es la independencia. Ojeda, 1995, 1996 y Sánchez, 1996 indican que, aunque este concepto parece sencillos de comprender desde el punto de vista teórico, en la práctica, es difícil estar seguros en una situación particular si hay o no independencia. Algunos puntos conflictivos sobre la independencia y probabilidad condicional son los siguientes:

- Apreciar la independencia en ensayos sucesivos de un mismo experimento; por ejemplo, algunas personas a las que nunca ha tocado la lotería, piensan que sus probabilidades de éxito son mayores que las de otra persona a la que ya ha tocado un premio.
- Comprender que la probabilidad de un suceso pueda condicionarse por otro que ocurra después que él. Hay personas que no le ven sentido a preguntas sobre la probabilidad de que la madre de una persona de ojos azules tenga ojos azules. En algunas

investigaciones se sugiere que esto puede ser debido a una confusión entre condicionamiento y causación.

- Confundir las probabilidades P(B/A) y P(A/B), por ejemplo confundir la probabilidad de toser si se fuma con la de fumar si se tose, o creer que estas probabilidades de refieren a un "suceso condicional".
- Confundir sucesos independientes con sucesos mutuamente excluyentes. Por ejemplo, creer que, al sacar una carta de una baraja, los sucesos "que sea rey" y "que sea copas" no son independientes porque existe el rey de copas.
- No identificar correctamente cual es el suceso que hay que poner como condición en una probabilidad condicional. Trata, por ejemplo, de resolver el siguiente problema, y verás como encuentras paradójica la solución: En una familia de dos hijos, uno es varón, ¿Cuál es la probabilidad de que el otro también sea varón?

Actividad 3.19. Un científico elige siempre usar el 0.05 como nivel de significación en los contrastes de hipótesis que realiza en sus experimentos. Esto significa que, a la larga:

a) El 5% de las veces rechazará la hipótesis nula
b) El 5% de veces que rechace una hipótesis nula se habrá equivocado
c) El 5% de las hipótesis nulas ciertas serán rechazadas

Actividad 3.20. Uno de los hijos de una familia con dos hijos es varón. ¿Cuál es la probabilidad de que el otro hijo también sea varón?

Actividad 3.21. Una urna contiene dos bolas rojas y dos blancas. Sacamos dos bolas de la urna sin reemplazamiento. ¿Cuál es la probabilidad de que la primera bola sea blanca, si la segunda también es blanca?

Actividad 3.22. Algunos autores atribuyen la dificultad con la probabilidad condicional a la ambigüedad de la notación y el lenguaje empleados en algunos libros de texto, que pueden dar la idea de que en la probabilidad sólo interviene un suceso (erróneamente denominado "suceso condicionado"). Busca un caso de un libro de texto en que aparezca esta imprecisión y razona el posible obstáculo didáctico que puede suponer para el alumno que lo usa.

- **Comprensión de la probabilidad desde un punto de vista frecuencial**

La investigación de Konold (1991) sugiere que con frecuencia tenemos dificultad de interpretar un experimento como parte de una serie de experimentos repetidos. Nos resulta difícil considerar un accidente que nos ocurre a nosotros como parte de la proporción normal de accidentes en esas circunstancias.

La interpretación frecuencial de la probabilidad o "probabilidad empírica" se restringe a fenómenos en los que es posible repetir indefinidamente ensayos "idénticos". En estos casos, la probabilidad se estima a partir de la frecuencia relativa del suceso en una serie larga de experimentos. No podemos aplicar esta perspectiva a un experimento del que sólo hay un ensayo aislado y único, a menos que imaginariamente pueda repetirse el experimento.

Aunque esta interpretación de la probabilidad se considera dentro de las corrientes objetivas, no significa que esté libre de consideraciones de tipo subjetivo. Por el contrario, esta interpretación requiere que un sujeto considere que los resultados de una larga serie de experimentos puedan ser considerados idénticos, para calcular la frecuencia de aparición de cada suceso particular. Por ejemplo, el hecho de que todos los lanzamientos que hacemos con una misma moneda puedan ser considerados idénticos, es, hasta cierto punto, subjetivo, ya que la persona que las lanza puede introducir sesgos en algunos de los lanzamientos.

Algunas personas tienen una especial dificultad para comprender la probabilidad frecuencial (se ha denominado su conducta "enfoque en el resultado aislado"), y consideran cada una de las repeticiones de un experimento como si estuviese aislada; sin guardar relación con las anteriores o posteriores. Para ellas, el fin de las situaciones de incertidumbre no es llegar a la probabilidad de ocurrencia, sino predecir con éxito el resultado de un ensayo simple. Las preguntas sobre la probabilidad se interpretan de forma no probabilística. Ante una pregunta en la que se pide explícitamente la probabilidad de un suceso, tratan de predecir si el suceso en cuestión ocurrirá o no en el siguiente experimento.

Ejemplo 3.10. Si en una predicción meteorológica se indica que la probabilidad de lluvia está en torno al 70%, algunas personas pensarán que va a llover. Si no llueve, pensarán que el meteorólogo se equivocó en sus predicciones. Si llueve un 70% de los días en que se pronosticó un 70% de probabilidades de lluvia, pensarán que el meteorólogo es poco fiable.

Estas personas evalúan las probabilidades comparándolas con los valores 0%, 50% y 100%. Si una probabilidad se acerca a los extremos 0% o 100%, el suceso se considerará como imposible o seguro, respectivamente. Sólo si se acerca al 50% se considerará verdaderamente aleatorio. Los estudiantes que muestran este tipo de comportamiento, tiende a buscar explicaciones causales en lugar de aleatorias a la ocurrencia de resultados inesperados y a la variabilidad de los fenómenos aleatorios. Por ejemplo, la frase "70% de posibilidades de lluvia" se interpreta como "70% de superficie cubierta por las nubes" o "70% de humedad relativa". Asimismo se ignora la información de tipo frecuencial, prefiriendo basar los juicios en consideraciones subjetivas sobre el fenómeno.

Actividad 3.23. Al comienzo de un camino se coloca un hámster y se le deja que circule libremente hacia su alimento situado al final de un camino que se bifurca hacia dos puntos A y B. En el punto A ponemos cacahuetes y en el B pipas. Un amigo criador de hámsters nos informa que el 70 % de los hámsters prefiere las pipas a los cacahuetes. ¿Pensarías que tu amigo está bien informado si el hámster prefiere los cacahuetes? ¿Y si de 3 hámsters, los 3 prefieren los cacahuetes?

Actividad 3.24. Supón que quieres comprar un coche nuevo y quieres decidir entre la marca A y B. En una revista de automóviles encuentras una estudio estadístico sobre reparaciones efectuadas el último años que muestra que la marca A tiene menos averías que la B. Sin embargo, te encuentras un amigo tuyo que te dice que compró el año pasado un coche B y no ha tenido más que problemas: primero se le estropeó la inyección de gasolina y gastó 25.000 pts, luego tuvo que cambiar el eje trasero y al final, ha vendido el coche porque se le fue la transmisión. ¿Que decisión tomarías, comprar un coche A o B?

3.7. Comprensión de tablas y gráficos estadísticos

Los profesores suponen, a veces, que la elaboración de tablas y gráficos es muy sencilla y dedican poco tiempo a su enseñanza. Sin embargo, elaborar una tabla de frecuencias o un gráfico supone una primera reducción estadística, pues se pierden los valores originales de cada uno de los datos individuales pasándose a la distribución de frecuencias. Este concepto es ya complejo, al referirse al agregado (población o muestra) y no a los datos particulares, lo que es una característica esencial de muchos conceptos estadísticos. Mientras que los alumnos comprenden bien propiedades que se refieren a individuos, como el color de pelo de una persona o su altura, les resulta más problemático comprender la idea de distribución de alturas de un grupo.

Hay que tener también en cuenta que en una tabla o gráfico aparecen (o pueden aparecer) distintos tipos de frecuencias: absolutas, relativas, porcentajes y frecuencias acumuladas. No hemos encontrado investigaciones sobre la comprensión de estos tipos de frecuencias por parte de los alumnos, pero sí sobre la comprensión de los gráficos. Algunos de estos resultados podrían también extenderse a las tablas estadísticas.

La destreza en la lectura crítica de datos es una componente de la alfabetización cuantitativa y una necesidad en nuestra sociedad tecnológica, ya que encontramos tablas y gráficos en la prensa, comercio, así como en distintas asignaturas del currículo. Además, las nuevas tecnologías posibilitan realizar gráficos estadísticos de modo rápido y eficaz. Curcio (1989) describe cuatro niveles distintos de comprensión de los gráficos, que pueden aplicarse a las tablas y gráficos estadísticos:

3. "Leer los datos": este nivel de comprensión requiere una lectura literal del gráfico; no se realiza interpretación de la información contenida en el mismo.
4. "Leer dentro de los datos": incluye la interpretación e integración de los datos en el gráfico; requiere la habilidad para comparar cantidades y el uso de otros conceptos y destrezas matemáticas.
5. "Leer más allá de los datos": requiere que el lector realice predicciones e inferencias a partir de los datos sobre informaciones que no se reflejan directamente en el gráfico.
6. "Leer detrás de los datos" supone valorar la fiabilidad y completitud de los datos.

Ejemplo 3.11. Si analizamos las tareas que se requieren en la interpretación de una serie de tiempo, "leer los datos" se refiere a cuestiones sobre la lectura de las escalas o encontrar el valor de una de las coordenadas de uno de los puntos, dado el valor de la otra coordenada. "Leer dentro de los datos" se refiere, por ejemplo, a cuestiones sobre la tendencia, sobre si podría ser representada o no mediante una función lineal o sobre si se observan ciclos. La predicción del comportamiento de la serie para los próximos meses, requeriría el trabajo en el nivel de "leer más allá de los datos". "Leer detrás de los datos" se refiere a valorar si los datos son completos, analizar la forma en que fueron recogidos y detectar posibles sesgos.

Una teoría relacionada con estos tres niveles de comprensión de gráficos es la de Wainer (1992). Este autor clasifica el tipo de preguntas que se pueden plantear a partir de un gráfico, en tres niveles:

- *Nivel elemental*. Preguntas relacionadas únicamente con la extracción de datos directamente del gráfico.
- *Nivel intermedio*. Preguntas relacionadas con la evaluación de tendencias basándose en una parte de los datos.
- *Nivel superior*. Preguntas acerca de la estructura profunda de los datos presentados en su totalidad, usualmente comparando tendencias y viendo agrupaciones.

Gerber, R. y cols. (1995) distinguen siete categorías sobre la comprensión de gráficos, que describen las diferencias en las habilidades de los estudiantes para interpretarlas.

Categoría 1. Los estudiantes no se centran en los datos, sino más bien en características idiosincrásicas de los mismos, que relacionan con su comprensión limitada del mundo de forma bastante imprecisa. No sólo tienen dificultades en interpretar el contenido de los gráficos, sino que son incapaces de procesar la información contenida en ellos de forma coherente.

Categorías 2 y 3. Se centran en los datos representados pero de forma incompleta. Se diferencian entre ellas en el foco de atención y en cómo se interrogan los datos. En la categoría 2 se centran en aspectos parciales de los datos, mientras que en la 3 se fijan en todo el conjunto, si bien en ambas aparecen dificultades para comprender el significado del gráfico.

En la categoría 2 no aprecian el propósito de cada gráfico. Por ejemplo en una pirámide de población interpretan las edades como distintos países en lugar de como propiedades de la población. En la categoría 3 aprecian el propósito del gráfico pero no comprenden aspectos específicos que son clave para entender la representación. Los estudiantes describen porciones discretas de los datos, más que patrones y regularidades. No hacen una interpretación global.

Categorías 4, 5 y 6. Representan vistas estáticas de los gráficos, aunque aumenta la precisión de la información cualitativa extraída de ellos. Se diferencian en el proceso de obtención de la información.

En la categoría 4 se reflejan patrones que generan los gráficos. Si un gráfico representa varias variables, los estudiantes son capaces de analizarlas una a una, pero no en su conjunto. Si tienen varios gráficos los analizan de uno en uno, pero no son capaces de utilizarlos todos simultáneamente para obtener más información.

En la categoría 5 los gráficos representan relaciones entre varias variables y los estudiantes pueden hacer comparaciones centrándose en todas ellas y no en una sola.

En la categoría 6 los estudiantes usan los gráficos para apoyar o refutar sus teorías. Van más allá de buscar similitudes y diferencias y pueden usar distintos tipos de representaciones para apoyar las informaciones.

Curcio (1989) estudió, el efecto que, sobre la comprensión de las relaciones matemáticas expresadas en los gráficos, tienen los siguientes factores:

- conocimiento previo del tema al que se refiere el gráfico; si el alumno es o no familiar con el contexto.
- conocimiento previo del contenido matemático del gráfico, esto es, los conceptos numéricos, relaciones y operaciones contenidas en el mismo;

- conocimiento previo del tipo de gráfico empleado (gráfico de barras, pictograma, etc.).

Reading y Pegg (1996), en una investigación centrada en la clarificación de lo que se entiende por "pensamiento estadístico", hacen un análisis de las respuestas de un grupo de estudiantes a dos tareas de respuesta abierta. En una de las tareas los datos se presentan sin agrupar, mientras que en la otra se dan mediante un gráfico. Los autores encuentran que hay dos pautas de razonamiento distintas y que la forma en que se presentan los datos tiene influencia sobre la elección del método de manejo de datos que se va a emplear. Parece que la comprensión de los datos es mejor cuando éstos se presentan sin agrupar, en lugar de hacerlo gráficamente.

Todos estos puntos intervienen en la dificultad que tienen los alumnos para interpretar los gráficos. Encontró que las principales dificultades aparecen en los niveles superiores ("leer dentro de los datos", "leer más allá de los datos", "leer detrás de los datos"), pero que las dificultades disminuyen con la edad.

Brigth, Curcio y Friel (en prensa) consideran los siguientes componentes en la comprensión de los gráficos:

- Traducción de un gráfico a otro o de gráfico a tabla o viceversa; lo que requiere un cambio en la forma de comunicar la información, e interpretar el gráfico a nivel descriptivo;
- Interpretación, que implica reorganizar el material y separar los factores más y menos importantes, búsqueda de relaciones entre los elementos específicos del gráfico o entre los elementos y las escalas en los ejes;
- Interpolación/ extrapolación implica la extensión de la interpretación, identificando tendencias o convenios implícitos.

Para cada uno de ellos deben diseñarse actividades de aprendizaje y tener en cuenta que los gráficos tienen diferentes componentes: ejes, escalas, y elementos específicos, cuyos convenios hay que conocer, y en cada uno de los cuales pueden cometerse errores.

En el caso más sencillo, los gráficos tienen un eje de coordenadas cartesianas en dos o tres dimensiones, aunque cada eje puede representar valores de las variables, frecuencias, promedios u otros resúmenes estadísticos, dependiendo del gráfico, y en algunos casos los ejes están implícitos, como en el gráfico del tallo y hojas o están basados en cordeenadas polares como el gráfico de sectores.

Li y Shen (1992) muestran ejemplos de elección incorrecta del tipo de gráfico en los proyectos estadísticos realizados por los estudiantes de secundaria. Algunos alumnos utilizaron un polígono de frecuencias con variables cualitativas, o un diagrama de barras horizontal para representar la evolución del índice de producción industrial a lo largo de una serie de años. Con frecuencia la elección de las escalas de representación son poco adecuadas para el objetivo pretendido. Los autores incluyen, además, una lista de errores de carácter técnico entre los cuales destacamos los siguientes:

- Omitir las escalas en alguno de los ejes horizontal o vertical, o en ambos;
- No especificar el origen de coordenadas;

- No proporcionar suficientes divisiones en las escalas de los ejes.

Este problema se agrava por la disponibilidad de "software" para la representación gráfica y el desconocimiento del modo correcto en que debe ser empleado por parte de los alumnos. Otras veces, el empleo inadecuado del "software" gráfico se debe a las concepciones incorrectas del estudiante, como al obtener un diagrama de sectores en los que éstos no son proporcionales a las frecuencias de las categorías, o comparar cantidades heterogéneas en un mismo gráfico.

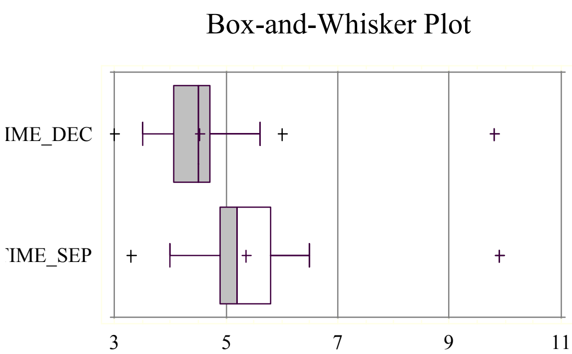
Actividad 3.20. Busca ejemplos en la prensa de tablas estadísticas o gráficos que presenten errores de construcción o que induzcan a obtener conclusiones equivocadas. Elaborar una lista de los principales tipos de errores detectados.

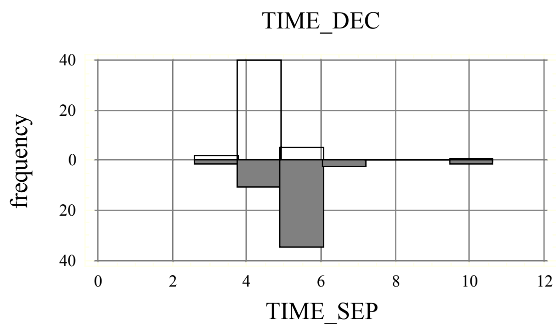
Actividad 3.21. Analiza los errores posibles en la interpretación y elaboración de los siguientes gráficos estadísticos: Diagrama de sectores, Histograma, Polígono acumulativo de frecuencias, Gráfico del tronco.

Actividad 3.22. Busca ejemplos de gráficos incorrectos o instrucciones inapropiadas para la realización de gráficos estadísticos en los libros de texto de enseñanza primaria o secundaria. ¿Qué obstáculos didácticos se deducirían para los estudiantes?

Actividad 3.23. Sobre un mismo conjunto de datos produce un gráfico que los represente adecuadamente y otro en que los datos queden distorsionados.

Actividad 3.24. Supongamos que queremos estudiar si existe o no diferencia en el tiempo que tardan unos alumnos en correr 30 metros en Septiembre y en Diciembre (después de 3 meses de entrenamiento). Analiza los dos gráficos que reproducimos a continuación. ¿Qué conocimientos estadísticos y sobre el gráfico necesitan los alumnos en cada caso para resolver el problema a partir del gráfico?





3.8. Investigaciones sobre la comprensión de las medidas de posición central

Además de ser uno de los principales conceptos estadísticos, y base en la construcción de otros, como momentos, etc., la media tiene muchas aplicaciones en cuestiones prácticas de la vida diaria, especialmente la media ponderada. Basta recordar conceptos como los índices de precios, la esperanza de vida o la renta per cápita. Las investigaciones muestran, sin embargo que no es un concepto tan simple como aparenta.

- ***Capacidad de cálculo y la comprensión de algoritmos***

El cálculo de la media ponderada parece sencillo. Sin embargo Pollatsek, Lima y Well (1981) encontraron que incluso alumnos universitarios no ponderan adecuadamente los valores al resolver el problema P5 y en ocasiones usan la media simple, en lugar de la media ponderada.

Las situaciones en las cuales se debe calcular una media ponderada y la selección de los correspondientes pesos no son fácilmente identificados por los estudiantes. Li y Shen (1992) indican que cuando los datos se agrupan en intervalos, los estudiantes olvidan con frecuencia que cada uno de estos grupos debería ponderarse de modo distinto al calcular la media y se limitan a calcular la media de todas las marcas de clase.

En otros casos el algoritmo se aplica de forma mecánica sin comprender su significado. Cai (1995) encontró que mientras la mayoría de alumnos de 12-13 años en su investigación eran capaces de aplicar adecuadamente el algoritmo para calcular la media, sólo algunos alumnos eran capaces de determinar un valor desconocido en un conjunto pequeño de datos para obtener un valor medio dado. Incluso encontrando el valor desconocido, fueron pocos los que lo hicieron a partir de un uso comprensivo del algoritmo, multiplicando el valor medio por el número de valores para hallar la suma total y de ahí el valor faltante, sino que la mayoría simplemente usó el ensayo y error.

Otros errores de cálculo en media, mediana y moda descritos por Carvalho (1998) al analizar las producciones escritas de los alumnos al resolver tareas estadísticas son los siguientes:

- Moda: Tomar la mayor frecuencia absoluta;
- Mediana: No ordenar los datos, para calcular la mediana; calcular el dato central de las frecuencias absolutas ordenadas de forma creciente; calcular la moda en vez de la mediana; equivocarse al calcular el valor central;
- Media: Hallar la media de los valores de las frecuencias; no tener en cuenta la frecuencia absoluta de cada valor en el cálculo de la media.

En realidad, el cálculo de la mediana es complejo, porque el algoritmo de cálculo es diferente, según tengamos un número par o impar de datos, y según los datos se presenten en tablas de valores agrupados o sin agrupar (Cobo y Batanero, 2000) y también el valor obtenido es diferente, según se aplique uno u otro algoritmo. Esto puede resultar difícil para los alumnos que están acostumbrados a un único método de cálculo y una única solución para los problemas matemáticos.

Gattuso y Mary (1998) analizan la evolución de la comprensión del algoritmo de cálculo de la media ponderada de los alumnos durante la enseñanza secundaria y universitarios, usando problemas con diferentes contextos y forma de representación. Las tareas presentadas fueron: cálculo de medias ponderadas, efecto que el cambio de un dato produce sobre la media y hallar un valor faltante en un conjunto de datos para obtener un promedio dado. Identifican las siguientes variables didácticas que afectan a la dificultad de las tareas: formato (tabla, serie de números, gráfico), si los valores de las variables son o no mucho mayores que los de las frecuencias (lo que influye en que el niño discrimine los dos conceptos); si una de las frecuencias es mucho mayor que las otras (de modo que se fuerce al niño a tener en cuenta las frecuencias). Observaron el efecto de estas variables y también la mejora con la instrucción, aunque no fue muy persistente en el tiempo.

Para Gattuso y Mary (1996) la comprensión conceptual no va paralela al número de años de instrucción en la materia en cuestión. Analizando las respuestas de estudiantes de secundaria y universitarios, con o sin instrucción previa sobre la media, esta autora ha observado que las estrategias preferidas por los de mayor edad eran más algebraicas y además obtenían mejores resultados cuando calculaban medias de conjuntos de datos agrupados, mientras que los más jóvenes preferían usar el conjunto de datos sin agrupar, aunque mostraron un nivel de éxito superior en los problemas de cálculo "invesos", es decir, aquéllos en los que se conoce la media y se deben averiguar algunos de los datos iniciales.

Un estudio sobre las dificultades de comprensión de los promedios, realizado por Batanero y otros (1997), muestra que la población estudiada, profesores de primaria en formación, encuentran dificultades en el tratamiento de los ceros y valores atípicos en el cálculo de promedios, posiciones relativas de media, mediana y moda en distribuciones asimétricas, elección de la medida de tendencia central más adecuada en una determinada situación y el uso de los promedios en la comparación de distribuciones. La conclusión a la que llegan estos autores es que la aproximación al estudio de los estadísticos de posición central basada en la definición algorítmica y el cálculo en colecciones de datos descontextualizados, no permite que los alumnos lleguen a una comprensión integral del concepto de promedio.

En nuestra opinión, los resultados de las investigaciones que hemos descrito sobre la media muestran también que el conocimiento de las reglas de cálculo por parte de los estudiantes no implica necesariamente una comprensión real de los conceptos subyacentes. Si los alumnos adquieren sólo el conocimiento de tipo computacional es probable que cometan errores predecibles, salvo en los problemas más sencillos.

Además, el proponer el algoritmo de cálculo prematuramente puede influir negativamente en la comprensión del concepto. Observan que los estudiantes desarrollan nociones intuitivas sobre el promedio en cuanto a medida de localización. Por ello se debería trabajar sobre las ideas intuitivas que tienen los alumnos para ayudarles a desarrollar caminos nuevos que les permitan enriquecer los conceptos que ya tienen asimilados. A la misma conclusión llega Tormo (1993) en un estudio realizado con alumnos de 12 a 15 años.

- ***Investigaciones sobre la comprensión de propiedades***

Cuando los alumnos comienzan a estudiar la media, mediana y moda por primera vez ya conocen ciertas operaciones aritméticas como la suma y multiplicación, e inconscientemente aplican a la operación de "promediar" algunas propiedades de las anteriores operaciones que no se cumplen en el caso de los promedios.

Mevarech (1983) observa que una explicación posible de los errores descritos por Pollasek y cols. (1981) es que los estudiantes suelen creer que un conjunto de números, junto con la operación media aritmética constituye un grupo algebraico, satisfaciendo los cuatro axiomas de clausura, asociatividad, elemento neutro y elemento inverso. En su investigación, llevada a cabo con 103 estudiantes de primer curso de universidad, encuentra un alto porcentaje de alumnos que atribuyen alguna de estas propiedades a la media aritmética cuando, en realidad, no las posee.

Mevarech (1983) observa que incluso los estudiantes universitarios piensan que la media tiene la propiedad asociativa y cuando tienen que hallar la media de un conjunto grande de números, lo dividen en partes hallando primero la media de cada parte y luego promediando el resultado obtenido. Podemos comprobar que esta propiedad, en general, no es cierta, si hallamos primero la media de tres números diferentes y luego promediamos los dos primeros y hacemos la media del valor obtenido con el último elemento. En otros casos no se tiene en cuenta el cero en el cálculo de la media, como si fuese un elemento neutro o bien se piensa que la media debe ser un elemento del mismo conjunto numérico del que se toman los datos.

Strauss y Bichler (1988) investigaron el desarrollo evolutivo de la comprensión de esta noción en alumnos de 8 a 12 años, distinguiendo las siguientes propiedades:

- La media es un valor comprendido entre los extremos de la distribución.
- La suma de las desviaciones de los datos respecto de la media es cero, lo que hace que sea un estimador insesgado.
- El valor medio está influenciado por los valores de cada uno de los datos. Por ello, la media no tiene elemento neutro.
- La media no tiene por qué ser igual a uno de los valores de los datos.
- El valor obtenido de la media puede ser una fracción (ello puede no tener sentido para la variable considerada), como cuando decimos que el número medio de hijos es 1,1.
- Hay que tener en cuenta los valores nulos en el cálculo de la media.
- La media es un "representante" de los datos a partir de los que ha sido calculada.

Para cada una de estas propiedades, los autores citados emplearon diversas tareas con las cuales evaluar su comprensión intuitiva, variando el tipo de datos (continuos, discretos) y el medio de presentación (verbal, numérico y concreto). No encontraron efectos significativos respecto al tipo de datos o medio de presentación empleado. Sus resultados sugieren una mejora de la comprensión con la edad, y diferencias de dificultad en la comprensión de las propiedades, siendo más fáciles las a), c) y d) que las b) f) y g).

Aunque una proporción importante de niños parecieron usar espontáneamente estas propiedades, algunos niños no tenían en cuenta el cero para calcular la media, o bien suponían que la media podría estar fuera del rango de variación de la variable, o que debería coincidir con uno de los valores de los datos.

Roth y Zawojewski (1991) realizaron una investigación sobre el efecto de la edad en la comprensión de las siete propiedades identificadas por Strauss y Bichler (1988), el efecto de las propiedades estadísticas y el de los diferentes formatos presentados a los sujetos. El trabajo se centró sólo en cuatro de estas siete propiedades (las identificadas como a, b, f y g), ya que la mayoría de los alumnos a partir de los 12 años habían superado las restantes. Como resultado de esta investigación se abrieron nuevas líneas de trabajo, entre ellas: analizar el tipo de explicaciones escritas dadas por los alumnos como respuesta a los ítems presentados y realizar una clasificación de las mismas, lo que podría aportar información adicional a la clasificación de Strauss y Bichler obtenida a partir de entrevistas clínicas.

León y Zawojeski (1991) realizan entrevistas a niños entre 8 y 14 años y analizan el efecto de la edad sobre la comprensión de estas propiedades. Además de encontrar una importante influencia de la edad sobre la comprensión de la media, también observaron que la contextualización de las tareas facilita mucho su resolución. Sin embargo, propiedades tales como que la suma de desviaciones respecto a la media es cero, que la media es un valor representativo de los valores promediados o que hay que tener en cuenta los valores nulos en el cálculo de la media continuaron siendo demasiado abstractas para una proporción importante de alumnos de 14 años.

La idea de representante de un conjunto de datos es importante en las aplicaciones prácticas, por ejemplo, al comparar dos conjuntos de datos respecto a una misma variable de interés. Como indican Mokros y Russell (1995) hasta que los niños no conciben el conjunto de datos como un todo, y no como un agregado de valores, no podrán comprender las ideas de resumen de los datos o representante de los datos, que se refiere al conjunto global y no a ninguno de sus valores aislados. Si la media artimética es un parámetro estadístico utilizado para resumir información de un conjunto de datos, entender el concepto de media aritmética implica reconocerle el papel representante del conjunto de datos. La idea de *representatividad* no es inmediata, antes de llegar a ella los alumnos deben captar la idea del conjunto de datos como una unidad.

Como se sabe la media es un valor "típico" o "representativo" de los datos. Campbell (1974) observa que, debido a ello, se tiende a situar la media en el centro del recorrido de la distribución, propiedad que es cierta para distribuciones simétricas. Pero cuando la distribución es muy asimétrica la media se desplaza hacia uno de los extremos y la moda o la mediana serían un valor más representativo del conjunto de datos. Esto no es siempre comprendido por algunos alumnos quienes invariablemente eligen la media como mejor representante de los datos sin tener en cuenta la simetría de la distribución o la existencia de valores atípicos, como hemos observado en nuestra propia experiencia.

Respecto a la comprensión de la mediana Barr, (1980) indica que los alumnos entienden que la mediana es el centro de "algo" pero no siempre comprenden a que se refiere ese "algo" porque no comprenden realmente que una tabla de frecuencia es sólo un resumen de los datos y no son capaces de pasar de la tabla a la lista de valores que es una representación alternativa de los datos. Incluso si se les da los datos en forma de lista no entienden por qué hay que ordenarlos para calcular la mediana, porque no entienden que la mediana es un estadístico que se refiere al conjunto ordenado de datos.

- ***Identificación de los campos de problemas***

No sirve de nada conocer las definiciones de las medidas de posición central y saber calcularlas si luego no se reconocen los problemas relacionados con estos conceptos. Pollasek y cols. (1981) propusieron a sus alumnos el siguiente problema, que es semejante al P4 descrito anteriormente, al tratar de determinar las concepciones de los alumnos universitarios sobre el valor esperado de una observación de una variable aleatoria, de la que se conoce su esperanza matemática.

En el ejemplo 3.8. la respuesta correcta a este ítem es 1,40, el valor esperado en la población. Sin embargo, son pocos los alumnos que dieron una respuesta correcta al problema, en la investigación citada, sino que, generalmente se busca un valor de la puntuación del quinto sujeto tal que, sumada a las cuatro anteriores, dé una media de 1,40. Pollatsek y cols. observan que los estudiantes esperan que la media de la muestra sea idéntica a la de la población, es decir, no aprecian la variabilidad aleatoria de la misma.

La comprensión de la idea de "valor típico" implica, según Russel y Mokros (1991), tres tipos diferentes de capacidades:

- Dado un conjunto de datos, comprender la necesidad de emplear un valor central, y elegir el más adecuado.
- Construir un conjunto de datos que tenga un promedio dado.
- Comprender el efecto que, sobre los promedios (media, mediana o moda), tiene un cambio en todos los datos o parte de ellos.

Russell y Mokros estudiaron las concepciones que los alumnos de 4º a 8º de enseñanza primaria tienen sobre los valores de tendencia central, empleando para ello las tareas anteriores, de las cuales la más difícil fue la segunda. Respecto al problema P3 (usar la media como representante de un conjunto de datos), resulta aún más difícil para los alumnos construir un conjunto de datos que tenga un promedio dado. Goodchild (1988) proporcionó a los estudiantes cajas de cerillas en las que se había impreso la frase "contenido medio 35 cerillas" y pidió a sus alumnos construir una distribución hipotética del contenido de 100 cajas. Lo que más le sorprendió fue que las distribuciones construidas por los alumnos, no tenían forma acampanada como la distribución normal. Goodchild sugirió que ello se debe a la falta de comprensión de la media como medida de posición central de la distribución.

En la investigación de Estepa y Batanero (1994) con alumnos del curso preuniversitario hemos observado casos de alumnos que basan la comparación de dos conjuntos de datos en valores aislados, por ejemplo, en la comparación de los máximos o los mínimos, o bien en la comparación de totales, o la inspección visual de la distribución global.

- ***Comprensión de representaciones y capacidad de argumentación***

Los términos matemáticos con que designamos los conceptos tiene un significado preciso, pero éste no siempre coincide con el asignado al término en el lenguaje coloquial. Russell y Mokros (1991) clasificaron en cuatro categorías los significados incorrectos atribuidos por los estudiantes a la palabra "media": valor más frecuente (en realidad esto sería una confusión son la palabra "moda"), "valor razonable" (significado coloquial del término), "punto medio" (confusión con la mediana) y "algoritmo" (es un significado restringido, donde la media se ve sólo como el algoritmo de cálculo). Cada uno de estos aspectos puede ser cierto en un caso dado, pero puede ser inapropiado en otro. Acaban el artículo señalando la necesidad de usar diferentes contextos y representaciones en la enseñanza de un concepto matemático.

Watson y Moritz (2000), analizan el significado intuitivo dado por los niños al término "promedio" y hallan un gran número de niños para los cuales el promedio es simplemente un valor en el centro de la distribución (es una idea próxima al concepto de mediana). Pocas veces se relaciona la palabra "promedio" con la moda y menos aún con la media aritmética. Las siguientes definiciones de "promedio" fueron obtenidas en entrevistas a niños realizadas por Watson y Moritz (2000): "*Significa igual", "que es normal", "no eres realmente bueno, pero tampoco malo"*. Al preguntar que quiere decir que el número medio de niños por familia es 2'3, obtienen respuestas correctas y otras como las siguientes:

> *"Que tienen dos niños grandes y otro que no ha crecido todavía", "que en las familias australianas el número más frecuente de niños es 2'3", "el '3 es un niño que tiene que crecer para hacerse mayor. Por ejemplo, tiene 3 años ahora y cuando cumpla 10, contará como 1 y entonces el número promedio de niños será 3".*

Para el profesor los enunciados sobre los promedios pueden parecer muy claros, pero estas respuestas indican la necesidad de poner atención al significado que las palabras y valores numéricos tienen para los estudiantes en relación a contextos específicos.

Eisenbach (1994) plantea a estudiantes universitarios en un curso introductorio de estadística el significado de la frase: "*¿Qué quiere decir que el salario medio de un empleado es 3.600 dólares?"* obteniendo respuestas como "*que la mayoría de los empleados gana alrededor de 3.600 dólares",* o que *"es el salario central; los otros trabajadores ganan más o menos de 3600 dólares",* que muestran la confusión terminológica entre las palabras "media", "mediana" y "moda".

Ya hemos indicado que la idea de promedio no se puede comprender hasta tanto se visualice el conjunto de datos como un todo y también que la forma de presentación de los datos (tabla, gráfico, datos sin tabular) incide en la dificultad de las tareas. Reading y Pegg (1996) estudiando la forma en que los niños de grados 7 a 12 reducen los conjuntos de datos observaron que algunos alumnos que eran capaces de dar un resumen de datos presentado en forma numérica, fracasaron en la tarea cuando los datos se presentaban por medio de un gráfico estadístico. También observaron que los niños mostraban dificultad a la hora de dar un argumento o justificar su respuesta de por qué se elegía un cierto promedio, al plantearles el siguiente problema, que es, esencialmente una versión simplificada del problema P3.

Ejemplo 3.12. Como parte de un proyecto los estudiantes de una clase miden cada uno su número de calzado, obteniéndose los siguientes datos:

26 26 26 27 27 27 27 28 28 28 28 28 28 29
29 29 29 29 30 30 30 30 30 30 30 31 32 32
33

Si te preguntan cuál sería el mejor número para representar este conjunto de datos, ¿Qué número o números elegirías? Explícanos por qué has elegido ese(esos) número(s).

En su investigación clasifica a los alumnos en 8 niveles diferentes de respuesta, pero, incluso los estudiantes de nivel 6 y 7 que son capaces de proporcionar un resumen como la media, son incapaces de argumentar el por qué de su decisión, más allá de dar la definición del concepto. Solo una pequeña parte de los estudiantes de su investigación (nivel 8) fueron capaces de justificar la elección de las medidas de valor central y dispersión relacionándolas con características del conjunto de datos.

Por nuestra parte en Estepa y Batanero (1999) hemos encontrado algunos alumnos del curso preuniversitario que basan sus argumentos en sus teorías previas, en lugar de en los datos al plantearles problemas similares al P3.

Actividad 3.24. Un objeto pequeño se pesa con un mismo instrumento por nueve estudiantes de una clase, obteniéndose los siguientes valores en gramos:

6.2, 6.0, 6.0, 15.3, 6.3, 6.1, 6.23, 6.15, 6.2

¿Cuál sería la mejor estimación del peso del objeto?

Actividad 3.25. El ayuntamiento de un pueblo quiere estimar el número promedio de niños por familia. Dividen el número total de niños de la ciudad por 50 (que es el número total de familias) y obtienen 2.2. ¿Cuáles de las siguientes afirmaciones son ciertas?

a) La mitad de las familias de la ciudad tienen más de 2 niños
b) En la ciudad hay más familias con 3 niños que familias con 2 niños
c) Hay un total de 110 niños en la ciudad
d) Hay 2,2 niños por adulto en la ciudad
e) El número más común de niños por familia es 2

Actividad 3.26. Construye una distribución simétrica y otra simétrica cuyo promedio sea 10 para una tamaño de muestra de 100 elemento. ¿Qué distribución probabilidad podría servir para representar tus datos?

Actividad 3.27. Un error frecuente en los estudiantes al calcular la media a partir de una tabla de frecuencias es sumar todas las marcas de clase y dividir por el número de intervalos. ¿Qué concepción errónea sobre la media se manifiesta en esta respuesta?

3.9. Otros resúmenes estadísticos

- **Características de dispersión**

El estudio de una distribución de frecuencias no puede reducirse al de sus promedios, ya que distribuciones con medias o medianas iguales pueden tener distintos grados de variabilidad. Un error frecuente es ignorar la dispersión de los datos cuando se efectúan comparaciones entre dos o más muestras o poblaciones.

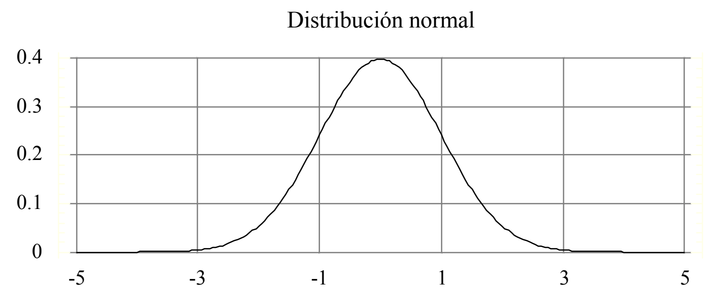
La desviación típica mide la intensidad con que los datos se desvían respecto de la media. Loosen y cols. (1985) hicieron notar que muchos libros de texto ponen mayor énfasis en la heterogeneidad entre las observaciones que en su desviación respecto de la posición central. Como señalan Loosen y cols., las palabras empleadas: variación, dispersión, diversidad, fluctuación, etc. están abierta a diferentes interpretaciones. Es claro para el profesor, pero no para el estudiante, cuándo estas palabras se refieren a una diversidad relativa a la media o en términos absolutos.

En un experimento, estos autores tomaron 154 estudiantes de primer curso de psicología, que no habían recibido ese curso una instrucción específica sobre la dispersión, mostrándoles dos conjuntos diferentes de bloques A y B. Las longitudes de los bloques en el conjunto A fueron 10, 20, 30, 40, 50 y 60 cm. y las longitudes de los bloques en el conjunto B fueron 10, 10, 10, 60, 60 y 60 cm. Al preguntar a los sujetos cual de los dos conjuntos presentaba mayor variabilidad, se obtuvieron las siguientes respuestas: el 50 % pensó que el conjunto A era más variable, el 36% que era más variable el conjunto B y el 14% que los dos conjuntos presentaban igual variabilidad. Loosen y cols. interpretaron estas respuestas como prueba de que el concepto intuitivo de variabilidad se equipara al de "no semejanza", es decir, cuánto varían unos valores respecto a otros, más que cuánto varían los valores respecto a un punto fijo. En este sentido el conjunto A ciertamente debe ser considerado mas variable que el B, aunque la desviación típica es mayor en el conjunto B.

Actividad 3.29. Analiza las propiedades de la desviación típica y de la desviación media y comparar con las propiedades descritas en el apartado anterior para la media. ¿Qué propiedades se conservan y cuáles desaparecen?

- **Puntuaciones tipificadas**

Uno de los usos más comunes de la media y desviación típica es el cálculo de puntuaciones Z (o puntuaciones tipificadas). La mayoría de los estudiantes no tienen dificultad en comprender este concepto ni en calcular las puntuaciones Z para un conjunto de datos particular. Sin embargo hay dos concepciones erróneas ampliamente extendidas entre los estudiantes, referentes al rango de variación de las puntuaciones Z, cuando se calculan a partir de una muestra finita o una distribución uniforme.



Por un lado, algunos alumnos creen que todas las puntuaciones Z han de tomar un valor comprendido entre -3 y +3. Otros estudiantes piensan que no hay límite para los valores máximo y mínimo de las puntuaciones Z. Cada una de estas creencias está ligada a una concepción errónea sobre la distribución normal. Los alumnos que piensan que las puntuaciones Z siempre varían de -3 a +3, han usado frecuentemente una tabla o gráfico de la curva normal N(0,1) con este rango de variación. De igual modo, los estudiantes que creen que las puntuaciones Z no tienen límite superior ni inferior, han aprendido que las colas de la curva normal son asintóticas a la abscisa y hacen una generalización incorrecta.

Por ejemplo, si consideramos el número de niñas entre diez recién nacidos, obtenemos una variable aleatoria X que sigue la distribución binomial con n=10 y p=0.5. La media de esta variable es np=5 y la varianza npq=2.5. Por ello, la puntuación Z máxima que puede obtenerse en esta distribución es Z =3.16 que es un límite finito pero mayor que 3.

- **Estadísticos de orden**

Llamamos estadísticos de orden a la mediana, percentiles, cuartiles, y rangos de cuartiles, porque nos indican la posición particular de un elemento deentro del conjunto de datos. Por ejemplo, la mediana nos indica el centro de la distribución. El estudio de los estadísticos de orden toma una gran importancia por dos motivos:

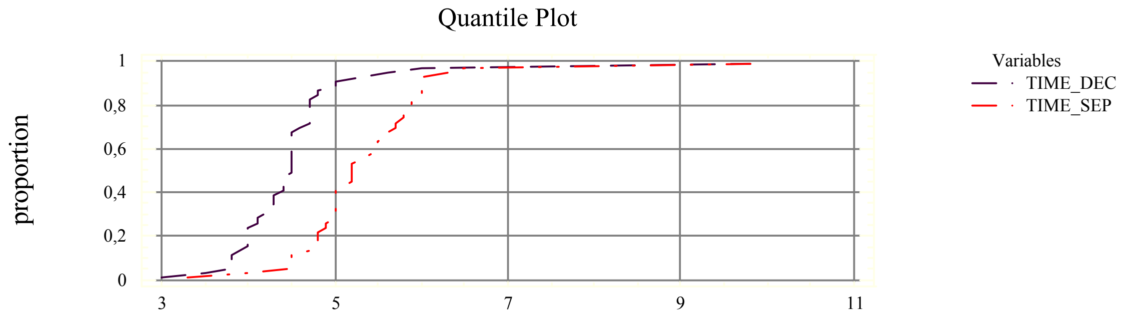
- El análisis exploratorio de datos se basa en estos estadísticos, porque son "robustos", esto es, menos sensibles a pequeños cambios en los datos y a los valores atípicos.
- Son la base de los métodos no paramétricos, que requieren para su aplicación un menor número de hipótesis que la estadística paramétrica y pueden ser aplicados con mayor generalidad, aunque son menos potentes.

El estudio de los estadísticos de orden presenta dificultades a nivel procedimental y conceptual. El cálculo de la mediana, percentiles y rango de percentiles se enseña empleando un algoritmo diferente para el caso de variables estadísticas agrupadas en intervalos o no agrupadas. Como sabemos, la opción de agrupar o no en intervalos se toma a juicio del que analiza los datos. Incluso los alumnos universitarios encuentran difícil aceptar que se pueda emplear dos algoritmos diferentes de cálculo para el mismo promedio y que puedan obtenerse valores distintos para el mismo parámetro, al variar la amplitud de los intervalos de clase.

Hay también diferencia entre el conocimiento conceptual de la mediana y el método de cálculo de la misma. Desde la definición de la mediana como "valor de la variable estadística que divide en dos efectivos iguales a los individuos de la población supuestos ordenados por el valor creciente del carácter", hasta su cálculo basado en la gráfica de frecuencias acumuladas intervienen una serie de pasos no siempre suficientemente comprendidos.

Esto puede observarse si pedimos a los alumnos que comparen las medianas de las dos variables que aparecen en el siguiente gráfico, si no les hemos enseñado explícitamente el método de cálculo a partir de la representación gráfica.

Estepa (1990) observa las dificultades de los alumnos al interpretar la gráfica de



frecuencias acumuladas de variables discretas, debido a que presenta discontinuidades de salto y su inversa no es una aplicación: en esta correspondencia un punto puede tener más de una imagen, o varios puntos pueden tener la misma imagen.

Barr (1980) llama la atención sobre la falta de comprensión de los estudiantes sobre la mediana en un estudio llevado a cabo con estudiantes de edades entre 17 y 21 años. El 49% dio una respuesta incorrecta a la cuestión siguiente:

La mediana del siguiente conjunto de números.

1, 5, 1, 6, 1, 6, 8 es

a) 1; b) 4; c) 5; d) 6; e) (otro valor); f) no se;

La mayoría de los alumnos entiende la idea de mediana como valor central, pero no tienen claro a que secuencia numérica se refiere ese valor central. Los estudiantes pueden interpretar la mediana como el valor central de los valores de la variable, de las frecuencias o incluso de la serie de datos antes de ser ordenada.

Actividad 3.30. Analiza los diferentes métodos de cálculo de la mediana, según el número de datos y si éstos están o no agrupados. Elabora una lista de errores posibles de los alumnos el cálculo de la mediana. Analiza los tipos de conocimientos implícitos en los diferentes métodos de cálculo de la mediana.

Actividad 3.31. Elabora una lista de propiedades de la mediana y compara con las correspondientes propiedades de la media. ¿Cuál de las propiedades de la mediana hace que sea preferida sobre la media en análisis exploratorio de datos? ¿Y en estadística no paramétrica?

3.10. Asociacion estadistica

El concepto de asociación tiene una gran relevancia en la educación matemática porque extiende la dependencia funcional y es fundamental para muchos métodos estadíticos, permitiendo modelizar numerosos fenómenos en diversas ciencias.

Este tema tiene conexiones importantes con el pensamiento funcional y otras áreas de la educación matemática, como la probabilidad y el razonamiento proprocional. El fin principal en muchas de estas aplicaciones es encontrar explicaciones causales que nos ayuden a comprender nuestro entorno. Sin embargo, la asociación no implica necesariamente relación causal. A veces existe "correlación espúrea" entre variables debida a la influencia de factores concurrentes sin que haya un vínculo causal.

Aparte de esta dificultad epistemológica, la investigación psicológica ha mostrado que la habilidad para emitir juicios de asociación no se desarrolla intuitivamente. Las personas adultas a veces basan sus juicios en creencias previas sobre el tipo de asociación que debería existir entre las variables, en lugar de en las contingencias empíricas presentadas en los datos.

La existencia de tales preconcepciones en las situaciones aplicadas es otra de las dificultades de la enseñanza de la asociación. A pesar de estos problemas epistemológicos y psicológicos, los investigadores en educación matemática han realizado pocas investigaciones sobre este tema, y la mayoría de las investigaciones psicológicas se han concentrado en las tablas de contingencia 2x2.

- ***Tablas de contingencia***

Hay tres tipos generales de problemas simples de asociación estadística: tablas de contingencia, correlación numérica y comparación de una variable numérica en dos o más muestras (que pueden ser independientes o relacionadas). A partir de aquí se definen problemas más complejos implicando a tres o más variables.

Una *tabla de contingencia* o clasificación cruzada de dos variables sirve para presentar en forma resumida la distribución de frecuencias de una población o muestra, clasificada respecto a dos variables estadísticas. En su forma más simple, cuando las variables poseen sólo dos categorías, toma la forma de la Tabla 1.

Tabla 1: Formato típico de la tabla de contingencia 2x2

	A	no A	Total
B	a	b	a+b
No B	c	d	c+d
Total	a+c	b+d	a+b+c+d

Podríamos proponer a los estudiantes diferentes problemas respecto a este tipo de tabla. Incluso la interpretación de las frecuencias reviste dificultad, ya que, a partir de la frecuencia absoluta de una celda, por ejemplo, la celda *a*, podemos obtener tres frecuencias relativas diferentes: la frecuencia relativa doble [a/(a+b+c+d)], la frecuencia relativa condicional respecto a su fila [a/(a+b)] y ña frecuencia relativa condicional respecto a su columna [a/(a+c)].

La investigación sobre los juicios de asociación ha sido objeto de gran interés en psicología y ha estado ligada a los estudios sobre toma de decisiones en ambiente de incertidumbre, ya que la toma de decisiones precisa, generalmente, un juicio previo sobre la asociación entre variables. La mayor parte de estas investigaciones han empleado tablas 2x2, como la mostrada en el ejemplo siguiente:

Ejemplo 3.13. Se quiere estudiar si un cierto medicamento produce trastornos digestivos en los ancianos. Para ello se han observado durante un periodo suficiente de tiempo a 25 ancianos obteniendo los siguientes resultados:

	Molestias digestivas	No tiene molestias	Total
Toma la medicina	9	8	17
No la toma	7	1	8
Total	16	9	25

Utilizando los datos de la tabla, razona si en estos ancianos, el padecer trastornos digestivos depende o no del medicamento.

El estudio del razonamiento sobre la asociación estadística fué iniciado por Piaget e Inhelder (1951), quienes consideraron que la comprensión de la idea de asociación implica las de proporción y probabilidad. Por esta razón, sólo estudiaron este problema con chicos en la etapa de operaciones formales (13-14 años).

El contexto empleado fue estudiar la asociación entre el color de los ojos y el de los cabellos. Para ello emplean cartas con dibujos de rostros en los que los ojos y el cabello están coloreados, preguntando al sujeto si existe o no una relación entre el color de los ojos y el del cabello, no en forma general, sino cuando se consideran los únicos datos presentados.

Si analizamos con detalle la tarea presentada podemos observar que, a pesar de su aparente simplicidad, es para el alumno un problema complejo y su dificultad depende de ciertos datos del enunciado. En el ejemplo dado, aparece una asociación de tipo inverso, puesto que el consumo del medicamento ha disminuido la frecuencia de los trastornos digestivos. No obstante, según los valores dados a las cuatro casillas de la tabla, puede aparecer asociación directa, inversa o independencia.

Otro hecho que complica esta tarea es que el número de ancianos en ambos grupos no es el mismo, esto es, que la distribución marginal de la variable (tomar o no tomar el medicamento) no tiene la misma frecuencia para sus diferentes valores. Otras posibles variables que influyen en la dificultad este problema son la intensidad de la asociación y la concordancia o no concordancia entre la asociación empírica en la tabla y las creencias previas del estudiante sobre la asociación que debe esperarse en el contexto dado.

En *la etapa IIIa,* Inhelder y Piaget encuentran que los sujetos analizan solamente la relación entre los casos favorables positivos (casilla a en la tabla 1) en relación a los casos totales (valor n en la tabla 1). En nuestro ejemplo, estos sujetos deducirían incorrectamente la existencia de una asociación directa entre las variables ya que el número de ancianos con trastornos digestivos que toman el medicamento es superior a cualquiera de las otras tres categorías.

Los adolescentes de nivel IIIa sólo comparan las casillas dos a dos. En la tabla 1, una vez admitido que también los casos (d) (ausencia-ausencia) son favorables a la existencia de asociación, no calculan la relación entre los casos que confirman la asociación (a+d) y el resto de los casos (b+c), lo que se produce sólo a partir de los 15 años (etapa IIIb) según Piaget e Inhelder.

Estas mismas conclusiones son obtenidas en trabajos con estudiantes adultos. La mayor parte de los estudiantes adultos basan su juicio, bien en la casilla (a) o comparando (a) con (b), esto es, empleando sólo la distribución condicional de tener o no trastornos digestivos, en los que toman medicamento. Con los datos del ejemplo, esta estrategia llevaría a concluir incorrectamente la existencia de una relación directa entre las variables, puesto que si nos restringimos a las personas que toman el medicamento, hay más con trastornos que sin ellos. La dificultad de estos problemas se pone de manifiesto al tener en cuenta que incluso la estrategia de comparación de diagonales, considerada como correcta por Piaget e Inhelder para resolver estos problemas sólo es válida para tablas con iguales frecuencias marginales, como puede apreciarse con los datos de nuestro ejemplo. Una estrategia correcta sería examinar la diferencia entre las dos probabilidades condicionales de que ocurra A cuando B es cierta y de que ocurra A cuando B es falsa:

$$\delta = a/(a+b) - c/(c+d)$$

es decir, en nuestro caso, sería necesario comparar las razones 9/17 con 7/8 (frecuencias condicionales).

Como dificultad añadida al tema, Chapman y Chapman (1967) mostraron que hay expectativas y creencias sobre las relaciones entre variables que producen la impresión de contingencias empíricas. Este fenómeno ha sido llamado "correlación ilusoria", porque los sujetos mantienen sus creencias y sobreestiman la asociación cuando piensan que existe causación entre dos variables.

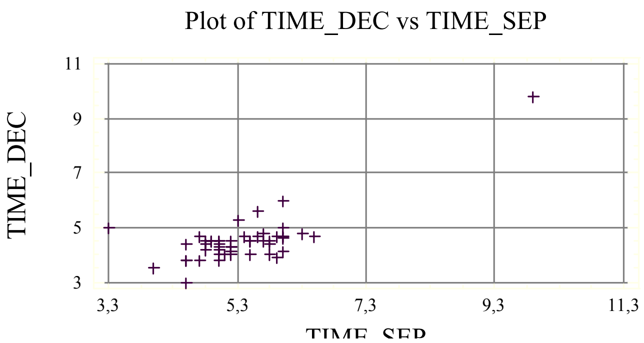
Actividad 3.30. A partir del ejemplo sobre trastornos digestivos en ancianos construye otros ítems tales que: a) los datos sean independientes; b) se presente una asociación directa.

- **Campos de problemas en el estudio de la asociación**

La tabla de contingencia es sólo un caso particular del campo de problemas del que emerge la asociación estadística. Para resolver este problema es preciso realizar operaciones y establecer relaciones con las diferentes frecuencias que aparecen o pueden calcularse en la tabla.

Actividad 3.31. Elabora una lista de las estrategias intuitivas que esperas sigan los alumnos para resolver el problema propuesto. Analiza las propiedades matemáticas utilizadas en cada una de estas estrategias.

Actividad 3.32. Un segundo tipo de problema de asociación es el estudio de la relación entre dos variables numéricas a partir de datos representados en un diagrama de dispersión. Describe las variables que crees que pueden afectar la dificultad de estas tareas y las posibles estrategias intuitivas de los alumnos para resolver estos problemas. ¿En qué propiedades matemáticas se diferencian estas estrategias de las utilizadas en las tablas de contingencia?



Actividad 3.33. Al medir la presión sanguínea antes y después de haber efectuado un cierto tratamiento médico a un grupo de 10 mujeres, se obtuvieron los valores siguientes:

	Presión sanguínea en cada mujer									
Mujer	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Antes del tratamiento	*115*	*112*	*107*	*119*	*115*	*138*	*126*	*105*	*104*	*115*
Después del tratamiento	*128*	*115*	*106*	*128*	*122*	*145*	*132*	*109*	*102*	*117*

¿Cuáles son las estrategias que esperarías usar un alumno sin instrucción en el tema? ¿Y si las muestras fuesen independientes?

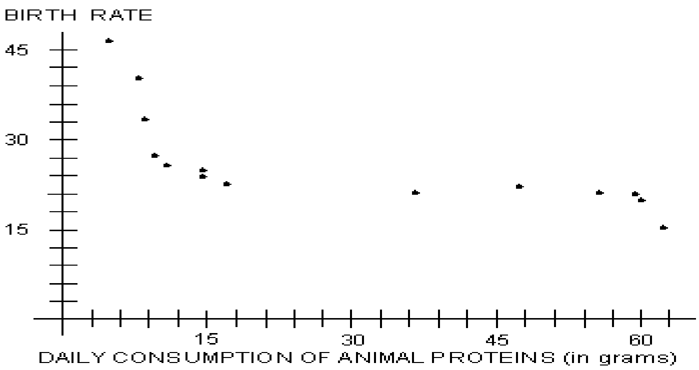
Otro problema diferente es valorar la correlación existente entre dos variables cuantitativas (Actividad 3.32), para lo cual podemos calcular la covarianza, el coeficiente de correlación o analizar la bondad del ajuste de una línea de regresión al diagrama de dispersión. Un tercer tipo de problema consiste en tratar de averiguar si una variable numérica tiene la misma distribución en dos muestras diferentes (actividad 3.33), que se puede resolver comparando la diferencia entre las medias o medianas, o bien, las representaciones gráfica o tabular de ambas distribuciones. Estos problemas y actividades son imprescindibles para construir progresivamente el concepto de asociación estadística y forman parte del significado matemático institucional del concepto dentro de un curso universitario introductorio al análisis de datos.

- **Algunas dificultades en el estudio de la asociación**

En Granada hemos llevado a cabo un estudio sobre el apprendizaje de la asoociación, con los tres tipos de problemas anteriores. Nuestro proyecto de investigación comenzó con el estudio del significado personal que los estudiantes daban al concepto de asociación antes de haber estudiado el tema. Después de algunas revisiones con muestras piloto, pusimos a punto un cuestionario con ítems semejantes a los presentados anteriormente, que propusimos a 213 alumnos de COU (17-18 años), antes de recibir enseñanza sobre la asociación. Se controlaron las siguientes variables de tarea: Signo e intensidad de la asociación, relación entre las creencias previas de los alumnos sobre el contexto del problema y el tipo de asociación presentada.

En cada ítem analizamos el tipo de asociación percibido por los estudiantes (asociación directa, inversa o independencia). Además clasificamos las estrategias de resolución utilizadas por los alumnos desde un punto de vista matemático. Esto nos permitió identificar estrategias intuitivas correctas que sugieren concepciones correctas o parcialmente correctas sobre la asociación estadística (Estepa et al., 1994; Batanero et al., 1996; Estepa, & Batanero, 1996; Estepa, & Sánchez Cobo, 1996). Algunos ejemplos se exponen a continuación:

(1) Utilizar la tendencia constante, creciente o decreciente de los puntos en los diagramas de dispersión para justificar el tipo de asociación (nula, positiva o negativa): "*Porque cuando el consumo diario de proteínas aumenta la tasa de natalidad disminuye*" (Ver figura). Como en caso de independencia no habría variación conjunta, estos estudiantes muestran una concepción correcta de la asociación.



(2) Utilizar las medias o los totales para comparar la distribución de una variable en dos muestras diferentes *"Porque la suma de todos los valores de la presión de la sangre antes del tratamiento es menor que la suma de valores de la presión de la sangre presiona después del tratamiento"* (actividad 3.33). Aquí los estudiantes usan implícitamente la idea correcta de que una diferencia en los totales implica asociación entre las variables.

(3) Comparar las frecuencias de casos a favor y en contra de la asociación en cada valor de la variable independiente o la razón de estas frecuencias en tablas de contingencia 2xr: "*No depende, porque 3/2 de los fumadores tienen trastornos bronquiales, y hay la misma proporción en los no fumadores*" (en la tabla que reproducimos a coontinuación). Esto indica una concepción correcta, ya que la razón de posibilidades se puede utilizar para evaluar la asociación en una tabla de contingencia 2xr.

	Fuma	No fuma	Total
Padece trastornos bronquiales	90	60	150
No padece trastornos bronquiales	60	40	100
Total	150	100	250

Otras estrategias de los estudiantes para resolver estos problemas eran inadecuadas, y les proporcionaban juicios incorrectos de asociación. A partir de ellas hemos descrito las siguientes concepciones erróneas sobre la asociación estadística:

(1) Concepción determinista de la asociación: Algunos estudiantes no admiten más de un valor de la variable independiente para cada valor de la variable dependiente. Cuando esto no ocurre, consideran que no hay dependencia entre las variables. En otras palabras, la relación entre las variables debe ser una función desde el punto de vista matemático. Por ejemplo*: "El tratamiento no tiene mucha influencia, ya que a algunas mujeres les aumenta la presión sanguínea, mientras que a otras les disminuye*" (actividad 3.33).

(2) Concepción unidireccional de la asociación. Si se percibe la dependencia solamente cuando es positiva (asociación directa), considerando la asociación inversa como independencia. El siguiente ejemplo ilustra un caso de asociación inversa interpretada como independencia y dificultades en el razonamiento proporcional: *"Personalmente creo que no hay dependencia, porque si tú miras la tabla hay mayor proporción de personas con trastornos bronquiales entre los fumadores*".

(3) Concepción local de la asociación. Utilizar solamente parte de los datos proporcionados por el problema para emitir el juicio de asociación. Si la parte de datos utilizados confirma un tipo de asociación, adoptan ese tipo de asociación en sus respuestas: *"Existe dependencia entre fumar y padecer trastornos bronquiales porque si observamos la tabla hay más fumadores con trastornos bronquiales que no fumadores 90>60".*

(4) Concepción causal de la asociación: Algunos estudiantes solamente consideran la existencia de asociación entre variables si se puede atribuir una relación causal entre ellas. Este tipo de concepción particularmente se encontró en un problema en el que se pedía que dos jueces puntuaran a un conjunto de individuos: *"Porque un juez no puede influir en el otro. Cada uno tiene sus preferencias; no puede haber mucha relación entre las puntuaciones otorgadas por cada uno".*

- **Influencia de los entornos informáticos en el aprendizaje de la asociación**

Después de llevar a cabo el estudio sobre concepciones iniciales de los estudiantes sobre la asociación, nuestra investigación se orientó a valorar el impacto que una experiencia de aprendizaje usando ordenadores tiene en dichas concepciones iniciales.

El contenido de los dos cursos incluyó los conceptos básicos sobre poblaciones y muestras, organización de datos, tipos de variables estadísticas y distribuciones de frecuencias, gráficos, parámetros y estadísticos (posición central, dispersión, estadísticos de orden, asimetría, curtosis), variables estadísticas bidimensionales: tablas de contingencia, covarianza, correlación y regresión lineal. En el segundo experimento se introdujeron también los conceptos de muestreo, distribución en el muestreo, intervalos de confianza, test de hipótesis de las medias para una y dos muestras y test Chi-cuadrado. Se identificaron las siguientes dificultades en el aprendizaje de los alumnos:

1. *La comparación de dos o más muestras, con objeto de estudiar la posible relación entre dos variables debe efectuarse en términos de frecuencias relativas*. Algunos alumnos comparan frecuencias absolutas de la distribución de una variable en dos muestras.
2. *La posible existencia o no de diferencias en la distribución de una variable entre dos o más muestras se deduce a partir de la comparación de toda la distribución de*

la variable en cada una de las muestras y no de una parte de la misma. Los estudiantes, sin embargo, comienzan con la comparación de valores aislados, al estudiar las dos muestras. Por ejemplo, en la primera sesión, los estudiantes solamente comparan los valores de máxima y mínima frecuencia en ambas muestras; aunque estas diferencias apuntan a la existencia de posible asociación, este modo de proceder es insuficiente para cuantificar la intensidad de la misma.

3. *A partir de una misma frecuencia absoluta pueden deducirse dos frecuencias relativas condicionales diferentes, según la variable que se emplee como condición. El papel de condición y condicionado en la frecuencia relativa condicional no es intercambiable*. Numerosos autores como Falk (1986) señalan la dificultad de interpretación de una probabilidad condicional porque los alumnos no diferencian a veces el papel jugado por la condición y el condicionado, con lo que pueden confundir P(A|B) con P(B|A), o no llegar a discriminarlas. Muchos alumnos mostraron confusiones similares a ésta en el estudio de las concepciones previas, que siguieron manifestándose durante el proceso de instrucción y que ha seguido manifestándose al finalizar la instrucción.
4. *Dos variables son independientes si la distribución condicional de una de ellas no cambia cuando se varían los valores de la otra variable.* Hasta llegar a la sesión 5, los estudiantes no descubren que una condición para la independencia es la invarianza de las distribuciones relativas condicionales, cuando varía el valor de la variable condicionante.
5. *En la determinación de la asociación entre dos variables, éstas juegan un papel simétrico. Por el contrario, en el estudio de la regresión, las variables desempeñan un papel asimétrico. Hay dos rectas de regresión diferentes, según cual de las dos variables actúe como variable independiente*. El hecho de que en la correlación no se distinga entre la variable explicativa y la variable explicada, mientras que en la regresión esta diferencia sea esencial (Moore, 1995) provocó gran confusión entre los estudiantes. Cuando necesitaron seleccionar la variable explicativa para calcular la línea de regresión, no supieron qué variable elegir. Al final del periodo de enseñanza estos estudiantes aun no habían descubierto que se pueden calcular dos líneas de regresión diferentes.
6. *Una correlación negativa indica dependencia inversa entre las variables*. Los alumnos se sorprenden al encontrar por primera vez un coeficiente de correlación negativo, hasta el punto de preguntar al profesor si ello es posible. Asimismo, aparece la duda en la comparación de dos coeficientes de correlación negativos, ya que, en este caso, un número menor corresponde a mayor intensidad en la asociación. Así, el conocimiento adquirido sobre el orden de los números negativos dificulta ahora la comprensión del signo negativo del coeficiente de correlación; se convierte en obstáculo epistemológico para dicha comprensión.

- **Diseño experimental**

Otro tema relacionado con la idea de asociación es el diseño de experimentos, que estudia los criterios estadísticos de planificación de los mismos que permitan alcanzar conclusiones acerca de un problema en el que un cierto número de variables pueden influir sobre otra.

Rubin y Rosebery (1990) planificaron y observaron un experimento de enseñanza dirigido a estudiar las dificultades de los profesores con las ideas estocásticas. Informaron que tanto los alumnos como su profesor interpretaron incorrectamente algunas de las ideas básicas del diseño experimental.

Una de las lecciones del mencionado experimento usó una actividad de lanzamiento a una canasta de baloncesto, en la que se varió la distancia de lanzamiento (de 1 a nueve metros) y el ángulo posicional del lanzador (para ángulos de 0, 45 y 90 grados). El objetivo de la lección era explorar los efectos separados de la distancia y el ángulo y la interacción entre las variables.

La observación de la discusión entre el profesor y los alumnos sobre la idea de variables independientes, dependientes y extrañas en el experimento de lanzamiento mostró la confusión entre estos conceptos. Algunos estudiantes sugirieron como posibles variables independientes características individuales del lanzador, como su altura o su habilidad para encestar. Incluso la altura de la canasta, que se conservó inalterable durante el experimento fue considerada como variable independiente por algunos estudiantes.

Otros estudiantes sugirieron que la iluminación del gimnasio podría ser diferente para las distintas combinaciones de ángulo y distancia, de modo que tanto el profesor como los alumnos quedaron con la creencia de que la presencia de tales influencias podría hacer imposible la obtención de conclusiones sobre el efecto de las variables ángulo y distancia. Finalmente, Rubin y Rosebery resaltaron la dificultad en distinguir entre las características de los sujetos que no tenían influencia sobre el resultado del experimento de otras variables que si podrían tenerla. El papel de la asignación aleatoria como medio de compensar estas diferencias individuales tampoco fue comprendido.

Actividad 3.33. Analiza llos distintos tipos de variables en el diseño experimental: explicativas, respuesta, variables concomitantes. Estudiar los posibles tipos de control para las variables concomitantes y las implicaciones de este control sobre los resultados del estudio estadístico.

Actividad 3.34. Analiza un ejemplo de investigación en educación estadística desde el punto de vista del diseño experimental empleado, tipos de variables y control efectuado. ¿Cuáles son las posibilidades de generalización de los resultados?

3.11. Introducción a la inferencia

- **Distribución normal y teorema central del límite**

El teorema central del límite es uno de los teoremas estadísticos más importantes que explica el uso generalizado de la distribución normal. Méndez (1991) estudia la coomprensión de este teorema en alumnos universitarios, indicandoo cuatro propiedades básicas en la comprensión del teorema:

(a) La media de la distribución muestral es igual a la media de la población, e igual a la media de una muestra cuando el tamaño de la muestra tiende al infinito.
(b) La varianza de la distribución muestral es menor que la de la población (cuando N>1).
(c) La forma de la distribución muestral tiende a ser acampanada a medida que se incrementa el tamaño muestral, y aproximadamente normal, independiente de la forma de la distribución en la población.

(d) La forma de la distribución muestral crece en altura y decrece en dispersión a medida que el tamaño muestral crece.

Define dos niveles de comprensión: El primer nivel esta formado por las habilidades y conocimientos que permiten resolver los ejercicios tipo presentados en los libros de texto. El segundo nivel representa la capacidad de aplicación a problemas reales. Observa que la comprensión del teorema se limita al nivel 1 y es excesivamente formal. Los estudiantes no son capaces de dar una explicación intuitiva de su significado o aplicarlo a casos reales.

Otras investigaciones estudian los factores que determinan el aprendizaje. Wilensky (1993, 1995, 1997), por ejemplo, muestra el influjo de los sentimientos y actitudes de los individuos frente a conceptos que conocen teóricamente y que manejan formalmente, pero cuyo significado no comprenden. Define la *ansiedad epistemológica*, como el sentimiento de descontento, confusión e indecisión que sienten la mayoría de los estudiantes frente a las distintas posibilidades o vías de resolución de un problema al no conocer los propósitos, orígenes o legitimidad de los objetos matemáticos que manipula y utiliza.

Esta *ansiedad* está reforzada por las prácticas de enseñanza empleadas en la clase de matemáticas y por la cultura matemática "protectora" que no promueve el diálogo entre docente y alumno y entre alumnos entre sí para ir construyendo el concepto.

Pensamos que hay muchas razones para que la *ansiedad epistemológica* sea particularmente pronunciada en el campo de la Estocástica. Todavía en la actualidad las nociones básicas de la teoría de probabilidad, como por ejemplo: aleatoriedad, distribución y probabilidad son bastante controvertidas. No existe un acuerdo generalizado sobre la interpretación de los conceptos de probabilidad o su aplicabilidad a casos particulares. Todo esto lleva a una confusión acerca de la aplicabilidad de la teoría de probabilidad a situaciones prácticas en nuestras vidas, lo que suscita dudas sobre su utilidad.

Wilensky concluye que la posibilidad de utilizar un medio informático que permita construir modelos de fenómenos probabilísticos que el alumno pueda revisar progresivamente ayuda a tener más seguridad y a reducir la ansiedad epistemológica. El autor enfatiza la importancia de la resolución de problemas, y principalmente sobre la proposición de problemas (o problem posing), ya que esta última permite proponer un problema de interés del propio alumno

- **Muestreo**

La idea central de la inferencia es que una muestra proporciona "alguna" información sobre la población y de este modo aumenta nuestro conocimiento sobre la misma. Se puede pensar en la inferencia estadística como una colección de métodos para aprender de la experiencia y, en la práctica, esto implica la posibilidad de acotar los valores de los parámetros de interés en las poblaciones, esto es, la obtención de intervalos de confianza para estos parámetros.

La comprensión de esta idea básica implica el equilibrio adecuado entre dos ideas aparentemente antagónicas: la representatividad muestral y la variabilidad muestral. La primera de estas ideas nos sugiere que la muestra tendrá a menudo características similares a las de la población, si ha sido elegida con las precauciones adecuadas. La segunda, el hecho de que no todas las muestras son iguales entre sí. El punto adecuado de equilibrio entre los extremos de información total e información nula respecto a la población es complejo, puesto que depende de tres factores: variabilidad de la población, tamaño de la muestra y coeficiente de confianza.

Los estudios sobre errores referidos al muestreo han tomado una gran importancia en el campo de la psicología, en el contexto de toma de decisiones. Un resumen de estos trabajos se presenta en Kahneman, Slovic, y Tversky (1982), quienes atribuyen estos errores al empleo de heurísticas en la resolución de problemas de decisión. En el libro citado, Kahneman y cols describen tres heurísticas fundamentales en los juicios probabilísticos: representatividad, disponibilidad y "ajuste y anclaje". También se estudian los sesgos asociados y sus implicaciones teóricas y prácticas.

Actividad 3.35. Estudiar en un libro de metodología de investigación los conceptos de validez y fiabilidad. ¿A qué conceptos de muestreo serían equivalentes? ¿Cuáles son las poblaciones y muestras empleadas?

- **La heurística de la representatividad**

La heurística de representatividad, descrita por Kahneman *et al.* (1982) consiste en calcular la probabilidad de un suceso sobre la base de la representatividad del mismo respecto a la población de la que proviene. En esta heurística, se prescinde del tamaño de la muestra, y de la variabilidad del muestreo, produciéndose una confianza indebida en las pequeñas muestras. Se supone que cada repetición del experimento, aunque sea limitada, ha de reproducir todas las características de la población. Por ejemplo, se espera que la frecuencia de una característica en la muestra coincida con la proporción de la misma en la población; por ello, tras una racha larga de aparición de un suceso se espera intuitivamente la aparición del suceso contrario, olvidando la independencia de los ensayos repetidos.

Mediante esta heurística, se resuelven juicios probabilísticos reduciéndolos a situaciones en las que se han dado ciertas características análogas a la presente y estableciendo comparaciones con ellas.

Actividad 3.36. La probabilidad de que nazca un varón es 1/2. A lo largo de un año completo, habrá más días en los cuales al menos el 60% de los nacimientos corresponden a varones:

- *en un hospital grande (100 nacimientos al día)*
- *en un hospital pequeño (10 nacimientos al día)*
- *no hay ninguna diferencia*

Razona la respuesta y analiza el razonamiento que seguiría un alumno según la opción elegida.

Actividad 3.37. Un taxi se ve implicado en un accidente de tráfico que ocurre de noche, en una ciudad con dos tipos de taxis de color verde y azul. Sabes los datos siguientes: a) 85% de los taxis son verdes y el resto azules; b) Un testigo que vio el accidente dice que el taxi era azul; c) El 80 % de los testigos hacen una identificación correcta de los datos del accidente. ¿Cuál es la probabilidad de que el taxi del accidente fuese azul?

Actividad 3.38. Dos recipientes etiquetados como A y B se llenan con canicas rojas y azules en las *siguientes proporciones:*

Recipiente	*Rojas*	*Azules*
A	*6*	*4*
B	*60*	*40*

¿Cuál recipiente da más probabilidad de obtener una bola azul? Analiza el razonamiento que seguiría un alumno que eligiese el recipiente B. ¿Por qué la solución de este problema es diferente al del ítem de la actividad 1?

Hay cuatro casos diferentes en los que se puede emplear la heurística de la representatividad, partiendo de un modelo M y un suceso X asociado a él:

(a) En el caso de que M sea una clase y X un valor de una variable aleatoria definida sobre esta clase (el sueldo de un profesor o la edad de un matrimonio). Aquí los valores más representativos son los que se acercan a las medidas de tendencia central (mediana, moda, media,...). La relación de representatividad depende del conocimiento que se tenga de la distribución de probabilidad que la variable X tenga en la clase M.

(b) En el caso en el que el valor que se estudia X sea un ejemplo de la clase M, como en el problema clásico en que se pregunta por la probabilidad de que una persona con unas ciertas características tenga una profesión dada. Se reconoce mejor a ejemplos más representativos que a otros menos representativos pero más frecuentes en una misma clase. Luego esta representatividad puede sesgar el reconocimiento de determinadas características.

(c) Cuando el valor a estudiar X es un subconjunto de M. En este caso el subconjunto representará no sólo los valores típicos, sino la variabilidad de la población o clase. Según esto, se puede decir que es más representativo que en cincuenta lanzamientos de monedas haya, por ejemplo 28 caras y 22 cruces a que haya 25 caras y 25 cruces.

(d) En el caso en que M sea un sistema causal y X sea una posible consecuencia del mismo. Por ejemplo, M puede ser la economía y X la tasa de inflación. La posible consecuencia será representativa del sistema causal M, debido a que, normalmente, esté asociada a ella o a que la gente crea correcta o incorrectamente que M causa X..

En estas cuatro clasificaciones, el grado de representatividad de M respecto a X es diferente. En el caso a) la representatividad está determinada por la percepción de la frecuencia relativa (probabilidad), en b) y c) por la similaridad entre la clase y el ejemplo o subconjunto, y por fin en la d) por la creencia que se tenga.

La evaluación de la probabilidad de un suceso incierto es un proceso complejo que comprende la interpretación del problema, la búsqueda de la información relevante y la elección de una respuesta adecuada. La representatividad se suele usar para predecir sucesos, ya que, normalmente, los acontecimientos más probables son más representativos que los menos probables, o bien se sobreestima la correlación entre una causa y su efecto (Kahneman *et al.* 1982). Pero su uso inapropiado da lugar a diferentes sesgos en los juicios probabilísticos. Estos sesgos no son debidos a la no comprensión de las normas probabilistas o estadísticas, sino a la facilidad que tiene el uso de la representatividad por su bajo coste de razonamiento frente al uso de cálculos normativos. Estos errores no son sólo típicos en estudiantes, sino que expertos en el tema llegan a cometerlos. Los sesgos más comunes que surgen de la utilización de esta heurística son:

(a) Insensibilidad al tamaño de la muestra.
(b) Insensibilidad a las probabilidades a priori
(c) Concepción errónea de las secuencias aleatorias. Estas se han descrito anteriormente
(d) Intuiciones erróneas sobre las probabilidades de experimentos compuestos.

- **Insensibilidad al tamaño de la muestra**

El valor esperado del estadístico de la muestra es el del parámetro en la población y no depende del tamaño de la muestra, aunque la varianza de la muestra cambia proporcionalmente a su tamaño, cambiando entonces las probabilidades de los sucesos.

Según indican Tversky y Kahneman (1971) se hace una extensión indebida de la ley de los grandes números, creyendo en la existencia de una "Ley de los pequeños números", por la que pequeñas muestras serían representativas en todas sus características estadísticas de las poblaciones de donde proceden. Este error se ha encontrado en individuos de niveles formativos diferentes; y puede tener importantes consecuencias de cara a la investigación experimental, ya que los científicos que creen en la "ley de los pequeños números" sobreestiman la potencia de sus métodos estadísticos, estiman a la baja la amplitud de sus intervalos de confianza y tienen unas expectativas injustificadas en la replicabilidad de experimentos realizados con pequeñas muestras.

- **Intuiciones erróneas sobre las probabilidades de experimentos compuestos.**

Otro error típico de esta heurística se suele dar en el cálculo de la probabilidad de un suceso en un espacio muestral producto. Se puede llegar a creer que es más probable la intersección de dos sucesos que su unión. Así, suele pensarse que es más probable encontrar un loco asesino a encontrar un hombre que solamente sea loco o que solamente sea asesino.

Actividad 3.39. Si observamos los siguientes 10 nacimientos, ¿Qué te parece más probable?

a) La fracción de varones será mayor o igual a 7/10.
b) La fracción de varones será menor o igual a 3/10.
c) La fracción de varones estará comprendida entre 4/10 y 6/10.
d) Las tres opciones anteriores a), b), c) son igual de probables.

Indica por qué das esta respuesta.

Actividad 3.40. Cinco caras de un dado equilibrado se pintan de negro y una se pinta de blanco. Se lanza el dado seis veces, ¿Cuál de los siguientes resultados es más probable?

a) Cara negra en cinco lanzamientos y blanca en el otro
b) Cara negra en los seis lanzamientos

Actividad 3.41. Diseñar una experimento que permita apreciar el efecto del tamaño de la muestra sobre la variabilidad de la media muestral: a) Usando un dispositivo físico; b) Usando una tabla de números aleatorio, la calculador o el ordendador.

- **La heurística de la disponibilidad**

Otra de las heurísticas usadas en el razonamiento probabilístico es la disponibilidad, por la que se juzgan más probables los sucesos más fáciles de recordar. Igual que con la representatividad, la presencia de errores derivados del uso de esta heurística se evalúa después de su utilización y no hay intento de predecir a priori bajo qué condiciones van a aparecer.

Actividad 3.43. Una persona debe seleccionar comités a partir de un grupo de 10 personas (Cada persona puede formar parte de más de un comité).

a) Hay más comités distintos formados por 8 personas
b) Hay más comités distintos formados por 2 personas
c) Hay el mismo número de comités de 8 que de 2

Razona la respuesta y analiza el razonamiento que seguiría un alumno según la opción elegida.

Actividad 3.44. Buscar ejemplos de situaciones en la vida diaria donde se ponga de manifiesto la heurística de la disponibilidad.

Esta heurística se apoya en el concepto de muestra, que es también uno de los conceptos claves en estadística. Al evaluar las probabilidades de un suceso, generalmente se usan muestras de datos para su estudio. Con ello corremos el riesgo de que la muestra no sea representativa de la población a estudiar o esté sesgada, cometiendo los consiguientes errores sistemáticos. Esta heurística también se debe a falta de razonamiento combinatorio.

Tversky y Kahneman asocian a esta heurística dos fenómenos característicos como la "correlación ilusoria" y la construcción de escenarios. El primero de ellos es el fenómeno que se atribuye a aquellos estudios en los que se pretende establecer una correlación entre variables que parecen estar relacionadas en sentido semántico, sin estarlo en sentido estadístico, juzgando la correlación sólo a partir de los casos en que los dos caracteres están presentes y no en aquellos en que los dos están ausentes.

Respecto al segundo, parte de la facilidad con que se pueden construir ejemplos o "escenarios" en la valoración de probabilidades y asignaciones de causas, frente a la heurística de disponibilidad que se basa en la facilidad para encontrar ejemplos.

Al comparar los errores cometidos en los juicios sobre el muestreo con las concepciones sobre el mismo que tienen los estadísticos expertos, Pollasek y cols. (1991) observan que éstos emplean "la extracción de bolas de una urna" para modelizar el proceso de muestreo. En este modelo, el muestreo aleatorio se ve como isomorfo al proceso de extracción con reemplazamiento de bolas de una. Los sujetos inexpertos podrían no tener un modelo para este proceso de muestreo o podrían tener un modelo inadecuado, lo que provocaría el empleo de la heurística de la representatividad incluso con muestras muy pequeñas

Otro problema relacionado con el muestreo son los diferentes niveles de concreción de un mismo concepto en estadística descriptiva e inferencia. En la estadística descriptiva la unidad de análisis es una observación (una persona, un objeto) y calculamos la media x de una muestra de tales objetos. En inferencia, estamos interesados por obtener información de la media teórica o esperanza matemática $E(\xi)$ de la población de la que ha sido tomada la muestra dada. Consideramos tal muestra como una observación de otra población diferente, la población de todas las posibles muestras de tamaño similar al dado, que podrían extraerse de la población de referencia. Hemos cambiado, en consecuencia, la unidad de análisis, que es ahora la muestra, y hablamos de que la media de la muestra es una variable aleatoria. Estudiamos la distribución de la media X en el muestreo y la media E(X) de esta variable aleatoria. Es preciso distinguir, por tanto, entre la media teórica en la población (que es una constante desconocida), la media particular obtenida en muestra; los posibles valores de las diferentes medias que se obtendrían en las diferentes muestras aleatorias de tamaño n (que es una variable aleatoria) y la media teórica de esta variable aleatoria, que coincide con la media de la población en el muestreo aleatorio. Esto supone una gran dificultad conceptual.

3.12. Contraste de hipótesis

En algunos países, uno de los temas introducidos en los últimos años de la enseñanza secundaria es el contraste de hipótesis. El campo de aplicación del contraste de hipótesis es muy amplio, pero esta parte de la inferencia es probablemente la peor comprendida, más confundida y de la que más se ha abusado en toda la estadística.

El término "contraste de hipótesis" abarca un gran número de procedimientos estadísticos: contrastes de diferencias de medias, análisis de la varianza, pruebas no paramédicas, contrastes multivariantes. Todos estos procedimientos tienen un núcleo común constituido por una serie de conceptos básicos (hipótesis nula y alternativa, estadístico de contraste, nivel de significación, etc.) y unos esquemas-procedimientos generales que se aplican a los casos particulares. La aplicación correcta de estos procedimientos precisa muchos tipos de elecciones, incluyendo: el tamaño de la muestra, el nivel de significación α y el estadístico apropiado.

La investigación sobre la comprensión de los métodos de inferencia muestra la existencia de concepciones erróneas ampliamente extendidas, tanto entre los estudiantes universitarios, como entre los científicos que usan la inferencia estadística en su trabajo diario. En particular los estudiantes tienen dificultades en los aspectos siguientes:

(a) La determinación de la hipótesis nula H_0 y la hipótesis alternativa H_1 ;
(b) La distinción entre los errores Tipo I y Tipo II;
(c) La comprensión del propósito y uso de las curvas características operativas o curvas de potencia; y
(d) La comprensión de la terminología empleada al establecer la decisión.

- **Errores comunes en la interpretación del nivel de significación y el p-valor**

Estas concepciones erróneas se refieren principalmente al nivel de significación α, que se define como la probabilidad de rechazar la hipótesis nula en caso de que sea cierta. La interpretación errónea más extendida de este concepto consiste en intercambiar los dos términos de la probabilidad condicional, es decir, en interpretar el nivel de significación como la probabilidad de que la hipótesis nula sea cierta si hemos tomado la decisión de rechazarla. Birnbaum (1982), por ejemplo, informó que sus estudiantes encontraban razonable la siguiente definición: "*Un nivel de significación del 5% indica que, en promedio, 5 de cada 100 veces que rechacemos la hipótesis nula estaremos equivocados*". Falk (1986) comprobó que la mayoría de sus estudiantes creían que α era la probabilidad de equivocarse al rechazar la hipótesis nula. Resultados semejantes se describen en el estudio de Pollard y Richardson (1987) realizado con investigadores.

Vallecillos (1994) planteó los siguientes ítems a una muestra de 436 estudiantes universitarios de diferentes especialidades (estadística, medicina, psicología, ingeniería y empresariales) que habían estudiado el tema:

> *Item 1: Un nivel de significación del 5% significa que, en promedio 5 de cada 100 veces que rechacemos la hipótesis nula estaremos equivocados (verdadero /falso). Justifica tu respuesta.*
> *Item 2: Un nivel de significación del 5% significa que, en promedio, 5 de cada 100 veces que la hipótesis nula es cierta la rechazaremos (verdadero / falso). Justifica tu respuesta.*

En el ítem 2 se presenta una interpretación frecuencial del nivel de significación (y es correcto), mientras que en el ítem 1 se han intercambiado los dos sucesos que definen la probabilidad condicional (y es incorrecto). Sin embargo, sólo el 32% de los estudiantes de la investigación de Vallecillos (1994) dio una respuesta correcta al ítem 1 y el 54% dio una respuesta correcta al ítem 2. De 135 estudiantes que justificaron su respuesta, el 41% dio un argumento correcto en los dos ítems. Un error prevalente en todos los grupos de estudiantes fue el intercambio de los términos de la probabilidad condicional, juzgando por tanto correcto el ítem 1 y falso el ítem 2. Entrevistas a un grupo reducido de estudiantes mostró que esta creencia aparecía en algunos estudiantes que eran capaces de discriminar entre una probabilidad condicional y su inversa (Vallecillos y Batanero, 1996). Otros estudiantes no distinguían las dos probabilidades condicionales, es decir, consideraban que ambos ítems eran correctos.

Que las probabilidades condicionales con términos intercambiados no coinciden, en general, se ilustra en la Tabla 1 que se refiere a la elección de estadística como tema optativo en una escuela. La probabilidad de que una chica tomada al azar estudie estadística y la probabilidad de que un estudiante de estadística sea una chica son diferentes:

P (estudie estadística / chica) = 3/4; P (chica / estudia estadística) =3/8

Tabla 1. Número de Chicos y Chicas en un Curso de Estadística

	Chicas	Chicos	Total
Estadística	300	500	800
No Estadistica	100	100	200
Total	400	600	1000

Es importante resaltar que, incluso cuando fijemos el nivel de significación α, es decir, la probabilidad de rechazar la hipótesis (supuesto que es cierta) y podamos calcular la probabilidad de obtener un valor del estadístico de contraste menor que un valor particular (supuesta la hipótesis nula cierta), la probabilidad de que la hipótesis nula sea cierta una vez la hemos rechazado y la probabilidad de que la hipótesis nula sea cierta una vez que hemos obtenido el valor del estadístico de contraste no pueden conocerse.

La probabilidad a posteriori de la hipótesis nula dado un resultado significativo depende de la probabilidad a priori de la hipótesis nula, así como de las probabilidades de obtener un resultado significativo, dadas las hipótesis nula y alternativa. Desafortunadamente estas probabilidades no pueden determinarse. Más aún, una hipótesis es o cierta o falsa y, por tanto, no tiene mucho sentido calcular su probabilidad en un paradigma inferencial clásico (donde damos una interpretación frecuencial a las probabilidades objetivas). Sólo en la inferencia Bayesiana pueden calcularse las probabilidades a posteriori de la hipótesis, aunque son probabilidades subjetivas, Lo más que podemos hacer, y es usando inferencia Bayesiana, es revisar nuestro grado de creencia personal en la hipótesis en vista de los resultados.

Otras interpretaciones erróneas del nivel de significación y el valor *p* son:

a) Algunas personas piensan que el valor *p* es la probabilidad de que el resultado se deba al azar. Podemos ver claramente que esta concepción es errónea del hecho de que incluso si la hipótesis nula es cierta (e.g. si no hubiera diferencias de rendimiento en el ejemplo 1) un resultado significativo puede ser debido a otros factores, como, por ejemplo, que los estudiantes del grupo experimental trabajasen más que sus compañeros al prepararse para la evaluación. Podemos observar aquí la importancia del control experimental para intentar asegurar que todas las condiciones (excepto el tipo de enseñanza) se mantienen constantes en los dos grupos. El valor *p* es la probabilidad de obtener el resultado particular u otro más extremo cuando la hipótesis nula es cierta *y* no hay otros factores posibles que influencien el resultado. Lo que rechazamos en un contraste de hipótesis es la hipótesis nula y, por tanto, no podemos inferir la existencia de una causa particular en un experimento a partir de un resultado significativo.

b) Otro error común es la creencia en la conservación del valor del nivel de significación cuando se realizan contrastes consecutivos en el mismo conjunto de datos, lo que produce el problema de las comparaciones múltiples. A veces aplicamos un alto número de pruebas de significación a un mismo conjunto de datos. El significado del nivel de significación (ver el Item 2 anterior) es que, si llevamos a cabo 100 comparaciones sobre el mismo conjunto de datos y usamos en todos ellos el nivel de significación .05, habrá que esperar que 5 de las 100 pruebas sean significativas por puro azar, incluso cuando la hipótesis nula sea cierta. Esto dificulta la interpretación de los resultados (Moses, 1992).

El uso frecuente de niveles de significación .05 y .01 es cuestión de convenio y no se justifica por la teoría matemática. Si consideramos el contraste de hipótesis como proceso de decisión (la visión de Neyman y Pearson), debemos especificar el nivel de significación antes de llevar a cabo el experimento y esta elección determina el tamaño de las regiones críticas y de aceptación que llevan a la decisión de rechazar o no la hipótesis nula. Neyman y Pearson dieron una interpretación frecuencial a esta probabilidad: Si la hipótesis nula es cierta y repetimos el experimento muchas veces con probabilidad de Error Tipo I igual a .05 rechazaremos la hipótesis nula el 5% de las veces que sea cierta.

En su libro "Diseño de experimentos" Fisher(1935) sugirió seleccionar un nivel de significación del 5%, como convenio para reconocer los resultados significativos en los experimentos. En sus trabajos posteriores, sin embargo, Fisher consideró que cada investigador debe seleccionar el nivel de significación de acuerdo a las circunstancias, ya que:

> "*de hecho ningún investigador mantiene un nivel de significación fijo con el cual rechaza las hipótesis año tras año y en todas las circunstancias"* (Fisher, 1956, p. 42).

Por el contrario, Fisher sugirió publicar el valor *p* exacto obtenido en cada experimento particular, lo que, de hecho, implica establecer el nivel de significación después de llevar a cabo el experimento.

A pesar de estas recomendaciones, la literatura de investigación muestra que los niveles arbitrarios de .05, .01, .001 se usan casi en forma universal para todo tipo de problemas. Skipper, Guenter y Nass (1970) sugirieron que esto trae como consecuencia la diferenciación de los resultados de investigación que se publicarán o no y llama la atención sobre las posibles implicaciones sobre los problemas investigados. A veces, si la potencia del contraste es baja y el error Tipo II es importante, sería preferible una probabilidad mayor de Error Tipo I.

(c) La interpretación incorrecta del nivel de significación se une normalmente a una interpretación incorrecta de los resultados significativos, un punto donde hubo también desacuerdos entre Fisher y Neyman -Pearson. Un resultado significativo implica para Fisher que los datos proporcionan evidencia en contra de la hipótesis nula, mientras que para Neyman y Pearson solo establece la frecuencia relativa de veces que rechazaríamos la hipótesis nula cierta a la larga (Error Tipo I). Por otro lado, debemos diferenciar entre significación estadística y significación práctica. En el Ejemplo 1 obtuvimos una diferencia media en puntuaciones entre los dos grupos de 13.32, que fue significativa. Sin embargo, podríamos haber obtenido una significación estadística mayor con un efecto experimental menor y una muestra de tamaño mayor. La significación práctica implica significación estadística más un efecto experimental suficientemente elevado.

d) Por otro lado, hay dos concepciones diferentes sobre los contrastes estadísticos que a veces se mezclan o confunden, como hemos razonado en el capítulo 2. Fisher concibió las pruebas se significación para confrontar una hipótesis nula con las observaciones y para él un valor p indicaba la fuerza de la evidencia contra la hipótesis (Fisher, 1958). Sin embargo, Fisher no creyó que los contrastes estadísticos proporcionaran inferencias inductivas de las muestras a las poblaciones, sino más bien una inferencia deductiva de la población de todas las muestras posibles a la muestra particular obtenida en cada caso.

Para Neyman (1950), el problema de contraste de una hipótesis estadística se presenta cuando las circunstancias nos fuerzan a realizar una elección entre dos formas de actuar. Aceptar una hipótesis sólo significa decidir realizar una acción en lugar de otra y no implica que uno necesariamente crea que la hipótesis es cierta. Para Neyman y Pearson, un contraste estadístico es una regla de comportamiento inductivo; un criterio de toma de decisión que nos permite aceptar o rechazar una hipótesis al asumir ciertos riesgos.

Hoy muchos investigadores emplean los métodos, herramientas y conceptos de la teoría de Neyman-Pearson con un fin diferente, medir la evidencia a favor de una hipótesis dada (Royal, 1997). El razonamiento más interior de la Tabla 3 del Capíitulo 2 (Implicación 5, observación y conclusión 5) puede describir el razonamiento usual en los contrastes de hipótesis hoy día, que consiste en:

(a) Una decisión binaria - decidir si el resultado es o no significativo. Esta decisión se lleva a cabo comparando el valor p con el nivel de significación, que se establece antes de recoger los datos.
(b) Un procedimiento inferencial que engloba un silogismo condicional (implicación 5 en la Tabla 3): Si $\mu_e = \mu_c$, entonces un valor significativo $\bar{x}_e - \bar{x}_c$ es muy improbable; $\bar{x}_e - \bar{x}_c$ es significativo, por tanto rechazamos $H_0 \equiv \mu_e = \mu_c$.
(c) Otro procedimiento inferencial que incluye un silogismo disyuntivo. O bien $\mu_e > \mu_c$, o $\mu_e = \mu_c$; si rechazamos que $\mu_e = \mu_c$, entonces $\mu_e > \mu_c$

Como consecuencia, la práctica actual de los contrastes estadístico tiene elementos de Neyman-Pearson (ya que es un procedimiento de decisión) y de Fisher (es un procedimiento inferencial, en el que usamos los datos para proporcionar evidencia a favor de la hipótesis), que se aplican en diferentes fases del proceso.

Otras características tomadas de Neyman-Pearson son que H_0 es la hipótesis de no diferencia, que el nivel de significación α debe escogerse antes de analizar los datos y debe mantenerse constante, así como los dos tipos de error. De Fisher conservamos la sugerencia de que la inferencia se basa en una probabilidad condicional; la probabilidad de obtener los datos supuesta cierta H_0, y que H_0 y H_1 son mutuamente exclusivas y complementarias. Deberíamos añadir que algunos investigadores suelen dar una interpretación Bayesiana a los resultados de los contrastes de hipótesis (clásicos), a pesar de que el enfoque de la estadística Bayesiana es muy diferente de las teorías tanto de Fisher como de Neyman y Pearson.

Vallecillos (1994) analizó también el concepto de nivel de significación y su relación con el resto de conceptos que intervienen en un contraste de hipótesis. Distinguió cuatro aspectos diferenciados para la comprensión de este concepto e identifica errores relacionados con cada uno de estos aspectos:

- *El contraste de hipótesis como problema de decisión:* El contraste de hipótesis es un problema de decisión entre dos hipótesis complementarias y excluyentes, con la posible consecuencia de cometer dos tipos de error, incompatibles pero no complementarios. Respecto a este apartado, algunos alumnos interpretan los errores tipo I y II como sucesos complementarios, por lo que la probabilidad de cometer alguno de los errores sería 1.
- *Las probabilidades de error y relación entre las mismas:* Los dos tipos de error tienen probabilidades α (Tipo I) y ß (Tipo II). Es necesario la comprensión de las probabilidades condicionales que intervienen en la definición de α y ß, de la dependencia de ß como función del parámetro desconocido y de las relaciones entre α y ß. Aparte del error del cambio en los condicionales se han encontrado otras interpretaciones erróneas de la probabilidad condicional que define el nivel de significación: suprimir la condición en la probabilidad condicionada que se emplea

para definir α; interpretar α como probabilidad de error (tanto tipo I como tipo II) en la decisión tomada.

- *Nivel de significación como riesgo del decisor*: Los valores de α y ß determinan los riesgos que el decisor está dispuesto a asumir y servirán, junto con las hipótesis, para adoptar un criterio de decisión. Se han hallado alumnos que creen que el cambio del nivel de significación no afecta al riesgo de error en la decisión.
- *Interpretación de un resultado significativo*. La obtención de un resultado estadísticamente significativo lleva al rechazo de la hipótesis nula, aunque no implica necesariamente ninguna relevancia desde el punto de vista práctico. Por ejemplo, una pequeña diferencia entre la media en dos poblaciones puede dar un resultado significativo si se toma una muestra de gran tamaño. Algunos estudiantes confunden la significación estadística y práctica o bien asocian un resultado significativo como uno que corrobora la hipótesis nula.

Actividad 3.45. Un contraste estadístico de hipótesis correctamente realizado establece la verdad de una de las dos hipótesis nula o alternativa. Analiza este enunciado y razona si es verdadero o falso.

Actividad 3.46. El nivel de significación en un contraste de hipótesis no puede ser justificado dentro de la estadística y es una elección de los estadísticos. Analiza este enunciado y razona si es verdadero o falso.

Actividad 3.47. ¿Qué ocurre con la potencia del contraste cuando pasamos de un nivel de significación de 0.01 a otro de 0.05?

Actividad 3.48. ¿Cuál de los tipos de error en un contraste está sujeto a control por parte del investigador?

Actividad 3.49. ¿Cuál es la diferencia entre significación estadística y significación práctica?

Actividad 3. 50. ¿Qué tipos de errores piensas están asociados al establecer las hipótesis en un contraste?

- **Factores Psicológicos que Contribuyen a la Prevalencia de Errores Comunes**

La anterior práctica de los contrastes estadísticos ha sido llamada marco ortodoxo de Neyman-Pearson (Oakes, 1986) o lógica híbrida de Neyman-Pearson (Gingerenzer, 1993). Este último autor piensa que dicha lógica puede explicar la creencia en que la inferencia estadística proporciona una solución algorítmica al problema de la inferencia inductiva, y el consecuente comportamiento mecánico que con frecuencia se presenta en relación con los contrastes estadísticos.

Como hemos descrito, Fisher y Neyman/ Pearson tuvieron diferentes interpretaciones de los contrastes estadísticos, incluyendo la forma es que se deben determinar los niveles de significación y la interpretación de un resultado significativo. Según Gingerenzer et al. (1989), la disputa entre estos autores se ha ocultado en las aplicaciones de la inferencia estadística en psicología y otras ciencias experimentales, donde se ha asumido una única solución para la inferencia. Libros de texto como el de

Guilford (1942) contribuyeron a difundir una mezcla de las lógicas de los contrastes de significación de Fisher con algunos componentes de Neyman-Pearson, dando una interpretación Bayesiana al nivel de significación y otros conceptos relacionados.

Con una analogía esclarecedora, Gingerenzer et al. (1989) comparan las características de Neyman-Pearson en la práctica actual de los contrastes estadísticos con el superego del razonamiento estadístico, porque prescriben los que debería hacerse y no da libertad a los investigadores. Requieren la especificación de hipótesis precisas, niveles de significación y potencia antes de recoger los datos y la probabilidad de error debe interpretarse en el contexto de muestreo repetido. Los componentes de Fisher se comparan al ego del razonamiento estadístico. Es conveniente para los investigadores, que quieren llevar a cabo su investigación y publicar sus trabajos, incluso a costa de determinar el nivel de significación después del experimento, establecer una hipótesis alternativa difusa o no establecerla antes de coger los datos, e interpretar la probabilidad de error como la probabilidad de error en su propio experimento.

El tercer componente en el comportamiento del investigador descrito por Gingerenzer et al. (1989) es el deseo Bayesiano de asignar probabilidades a las hipótesis en base a los datos de investigación (el id de la lógica híbrida). Cuando encontramos un resultado significativo nos preguntamos si este resultado puede ser debido al azar o si por el contrario es consecuencia de nuestra manipulación experimental. Falk (1986) encuentra natural la interpretación del nivel de significación como la probabilidad a posteriori de error, una vez que hemos rechazado la hipótesis en la que el investigador está interesado. Gingerenzer et al. (1989) sugieren que es el conflicto entre estos tres componentes psicológicos lo que explica nuestros usos incorrectos de la inferencia estadística y la institucionalización del nivel de significación como medida de la calidad de la investigación en revistas científicas y manuales de estadística.

Por otro lado, los sesgos en el razonamiento inferencial son sólo un ejemplo del pobre razonamiento de los adultos en los problemas probabilísticos, que ha sido extensamente estudiado por los psicólogos en relación con otros conceptos, como la aleatoriedad, probabilidad y correlación (Kahneman, Slovic, y Tversky, 1982; Nisbett y Ross, 1980). En el caso específico de interpretación incorrecta de los resultados de la inferencia estadística, Falk y Greenbaum (1995) sugieren la existencia de mecanismos psicológicos profundos que llevan a las personas a creer que eliminan el azar y minimizan su incertidumbre cuando obtienen un resultado significativo. Describen *la ilusión de la prueba probabilística por contradicción* o *ilusión de alcanzar la improbabilidad,* que consiste en la creencia errónea de que la hipótesis nula se vuelve improbable cuando se obtiene un resultados significativo, basada en una generalización abusiva del razonamiento lógico a la inferencia estadística (Birnbaum, 1982; Lindley, 1993).

Mientras una contradicción definitivamente prueba la falsedad de la premisa de partida, la creencia que al obtener datos cuya probabilidad condicional bajo una hipótesis dada es baja implica que la hipótesis condicionante es improbable es una falacia. La ilusión de la prueba probabilística por contradicción es, sin embargo, aparentemente difícil de erradicar, a pesar de la clarificación que se hace en muchos libros de estadística. En otros casos, esta concepción errónea está implícita en los libros de texto, como muestran Falk y Greenbaum (1995).

Según Falk (1986), las concepciones erróneas respecto al nivel de significación se relacionan también con las dificultades en discriminar las dos direcciones de las probabilidades condicionales, lo que se conoce como *la falacia de la condicional transpuesta* (Diaconis y Friedman, 1981), que desde hace tiempo se reconoce como extendida entre los estudiantes e incluso los profesionales. Además, Falk (1986) sugiere que la ambigüedad verbal al presentar α como *P(Error Tipo I)* puede provocar confusión entre las dos direcciones opuestas de la probabilidad condicional entre los estudiantes, que pueden pensar estar tratando con la probabilidad de un suceso simple. Falk sugiere que "Error Tipo I" es una expresión desafortunada que no debería usarse en si misma. Un "suceso condicional" no es un concepto legítimo y sólo las probabilidades condicionales están inequívocamente definidas, aunque esta confusión también aparece en algunos libros de texto.

Aunque α es una probabilidad condicional bien definida, la expresión "Error Tipo I" no está redactada como una condicional, ni indica cual de las dos combinaciones posibles de los sucesos que intervienen se refiere. En consecuencia, cuando rechazamos H_0 y nos preguntamos por el tipo de error que podríamos cometer, el concepto "Error Tipo I" nos viene a la mente en forma inmediata ya que la distinción crucial entre las dos direcciones opuestas de la probabilidad condicional ha sido difuminada. Esto nos lleva a interpretar el nivel de significación como la probabilidad de la conjunción de dos sucesos "la hipótesis nula es cierta" y "la hipótesis nula es rechazada" (Menon, 1993).

Durante muchos años se han lanzado críticas en contra del contraste estadístico y ha habido muchas sugerencias de eliminar este procedimiento de la investigación académica. Sin embargo, los resultados significativos siguen siendo publicados en las revistas de investigación, y los errores referentes a los contrastes estadísticos siguen llenando los cursos y libros de estadística, así como los informes publicados de investigación. Falk (1986) sugiere que los investigadores experimentan una confidencia ilusoria en los contrastes estadísticos debido a la sofisticación de los términos y fórmulas matemáticas, que contribuyen a nuestra sensación de que la significación estadística garantiza la objetividad. Un problema adicional es que otros procedimientos estadísticos sugeridos para reemplazar o complementar los contrastes estadísticos (como los intervalos de confianza, la estimación de la magnitud de los efectos experimentales, el análisis de la potencia y la inferencia Bayesiama) no resuelven los problemas filosóficos y psicológicos que hemos descrito.

- **La enseñanza y Aprendizaje de Conceptos de Inferencia**

En el capítulo 2 hemos descrito la lógica de los contrastes de significación, su papel en la investigación experimental, y en las secciones anteriores las dificultades psicológicas y filosóficas relacionadas con ellos. También en el capitulo 2 hemos revisado algunas críticas frecuentes contra las pruebas de significación estadística. Estas críticas no pueden aplicarse a los procedimientos matemáticos en el contraste de hipótesis, donde no hay contradicciones. Por el contrario se relacionan con los usos incorrectos del contraste de significación y son consecuencias de errores conceptuales y de problemas filosóficos y psicológicos que hemos descrito a lo largo de este trabajo.

Como describe Ito (1999), hay tres niveles diferentes en la controversia respecto a los contrastes estadísticos:

- La disputa dentro de la misma estadística, donde diferentes métodos e interpretaciones varias de los mismos métodos fueron recomendadas por los enfoques de Fisher, Neyman-Pearson y en el enfoque Bayesiano.

- La controversia en la aplicación de la estadística, donde, en la práctica el contraste de significación es una mezcla informal de los contrastes de significación originales de Fisher, la teoría de Neyman-Pearson y conceptos e interpretaciones que no son partes de esta última. Mas aún, los editores de revistas y las sociedades profesionales están sugiriendo cambios en las políticas de publicación científica, respecto a los métodos estadísticos (Lecoutre, 1999).
- La controversia en la enseñanza acerca de cuándo, como y con qué profundidad deberíamos enseñar la inferencia estadística.

Estos tres niveles diferentes están de hecho interrelacionados, porque nuestras concepciones acerca de la teoría estadística, también afectan a nuestra aplicación y enseñanza de la estadística. Esto es importante, ya que, con la creciente investigación sobre la enseñanza y aprendizaje de la estadística, el análisis de datos se está introduciendo cada vez más en el nivel escolar (Shaughnessy, Garfield, y Greer, 1997) incluyendo también en muchos países los rudimentos de inferencia (Dahl, 1999). Nuestra opinión es que el contraste de hipótesis no debiera abandonarse en las ciencias sociales y educación, sino que debería cambiarse su enseñanza y su práctica hasta llevar a un "proceso significativo" (Levin, 1998b), que incluya replicaciones independientes de los estudios, elección de tamaños óptimos de muestras, la combinación del contraste de hipótesis con intervalos de confianza y/o estimación del tamaño de los efectos y la especificación de criterios de "éxito" antes de realizar el experimento.

Nuestro análisis muestra con claridad la complejidad conceptual de los contrastes estadísticos, y la atención particular que debe darse a la enseñanza de la inferencia si queremos prevenir en nuestros estudiantes futuras faltas de comprensión como las descritas por Vallecillos (1999). Se ha sugerido la revisión de la metodología de enseñanza en los cursos introductorios de estadística (Moore, 1997 y la discusión relacionada) para pasar a un modelo constructivista de aprendizaje, en el que el profesor guíe a sus estudiantes hacia unas competencias y conocimientos estadísticos más específicos. Nuevos libros de texto que cambian el papel del estudiante de simple oyente a una participación más activa en actividades estructuradas (e.g., Rossman, 1996) facilitan este enfoque.

Puesto que los ordenadores hacen posible una variedad de cálculos y representaciones gráficas, Moore (1997) recomienda dar a los estudiantes la oportunidad de experimentar con datos y problemas reales. Específicamente, la simulación con ordenador puede contribuir a mejorar la comprensión del estudiante de las ideas de variabilidad muestral, estadístico y su distribución, respecto a las cuales hay muchas concepciones erróneas (Rubin, Bruce, y Tenney, 1991; Well, Pollatsek, y Boyce, 1990) y que son esenciales para comprender la lógica del contraste de significación. Por ejemplo, delMas, Garfield y Chance (1999) describen el software Sampling Distribution y actividades educativas diseñadas para guiar a sus estudiantes en la exploración de las distribuciones muestrales. En su experimento, los estudiantes podían cambiar la forma de la distribución teórica de la población (normal, sesgada, bimodal, uniforme, en forma de U) y simular la distribución muestral del diferentes estadísticos para varios tamaños de muestra. Las actividades estuvieron orientadas a enfocar la atención de los estudiantes hacia el Teorema Central del Límite.

Sin embargo, e incluso cuando los resultados demostraron un cambio significativo positivo en los estudiantes como consecuencia de la instrucción, delMas, Garfield y Chance (1999) avisan que el uso de la tecnología y actividades basadas en resultados de la investigación no siempre produce una comprensión efectiva de las distribuciones muestrales. Los autores sugieren que las nuevas actividades y el aprendizaje del software puede ser demasiado exigente para algunos, y la nueva información sobre el software puede interferir con el aprendizaje de los estudiantes sobre las distribuciones muestrales, cuya comprensión requiere la integración de las ideas de distribución, promedio, dispersión, muestra y aleatoriedad. Es claro que necesitamos más investigación para comprender como podemos usar la tecnología para ayudar a los estudiantes en su proceso de aprendizaje (ver el trabajo de Ben-Zvi, 2000). En particular necesitamos encontrar buenas situaciones didácticas en las que los estudiantes sean confrontados con sus concepciones erróneas, como la confusión entre una probabilidad condicional y su inversa o su creencia en la posibilidad de calcular la probabilidad de una hipótesis (dentro de la concepción objetiva de la probabilidad).

Por otro lado, el contraste estadístico sólo es una parte de proceso más general de inferencia científica, como se indica en las tablas 2 y 3. Sin embargo, frecuentemente encontramos que la estadística se enseña aisladamente, sin conectarla con un marco más general de metodología de investigación y diseño experimental. Desde nuestro punto de vista, es necesario discutir el papel de la estadística en la investigación experimental con los estudiantes y hacerlos conscientes de las posibilidades y limitaciones de la estadística en el trabajo experimental. Aún más, coincidimos con la sugerencia de Wood (1998) de enfocar el curso introductorio de estadística alrededor del razonamiento estadísticos, es decir el ciclo de aprendizaje Planificación-Conjetura-Comprobación-Acción. El análisis estadístico de datos no es un proceso mecánico y, por tanto, no debería ser enseñado o aplicado de esta forma. Puesto que la estadística no es una forma de hacer sino una forma de pensar que nos puede ayudar a resolver problemas en las ciencias y la vida cotidiana, la enseñanza de la estadística debería empezar con problemas reales mediante los cuales los estudiantes puedan desarrollar sus ideas, trabajando las diferentes etapas en la resolución de un problema real (planificar la solución, recoger y analizar los datos, comprobar las hipótesis iniciales y tomar una decisión en consecuencia).

Finalmente, recomendamos a los investigadores que reconozcan la complejidad de la aplicación de la inferencia estadística para resolver problemas reales y que necesitan la colaboración de estadísticos profesionales, además del uso de su conocimiento profesional para juzgar hasta qué punto sus cuestiones de investigación pueden contestarse por medio del análisis estadístico.

4.

EL CURRICULO DE ESTADÍSTICA

4.1. Introducción

Como se indicó en la introducción, la enseñanza de la Estadística ha cobrado gran desarrollo en los últimos años y algunos países han dedicado grandes esfuerzos a diseñar el currículo y los materiales de enseñanza. El mayor peso que se da a la Estadística en los diferentes niveles educativos, tanto en España como en otros países, requiere una intensa preparación de los profesores, para permitirles abordar con éxito los objetivos educativos correspondientes. Muchos profesores precisan incrementar su conocimiento, no sólo sobre la materia, sino también sobre los aspectos didácticos del tema. En este capítulo analizaremos los contenidos de estadística y probabilidad en los currículos de primaria y secundaria, haciendo primero una introducción a las bases en que se apoya este diseño curricular. Tomaremos el término "currículo" en la acepción de Stenhouse (1984) como actividad de planificar una formación.

Rico y Sierra (1997) sugieren que al reflexionar sobre el currículo se debe tener en cuenta el colectivo de personas a formar, el tipo de formación que se quiere proporcionar, la institución donde se lleva a cabo la formación, las finalidades a alcanzar y los mecanismos de valoración. En nuestro caso, fijados la institución de enseñanza (enseñanza secundaria en España) y los alumnos de estos niveles, analizaremos los fines de la enseñanza de la estadística, el tipo de conocimiento que se espera adquieran los alumnos, así como el proceso de evaluación.

- **Razones y fines de la educación estadística**

Las razones para el interés hacia la enseñanza de la estadística han sido repetidamente señaladas por diversos autores, desde comienzos de la década de los ochenta. Por ejemplo en Holmes (1980) encontramos las siguientes:

- La estadística es una parte de la educación general deseable para los futuros ciudadanos adultos, quienes precisan adquirir la capacidad de lectura e interpretación de tablas y gráficos estadísticos que con frecuencia aparecen en los medios informativos.
- Es útil para la vida posterior, ya que en muchas profesiones se precisan unos conocimientos básicos del tema.
- Su estudio ayuda al desarrollo personal, fomentando un razonamiento crítico, basado en la valoración de la evidencia objetiva.
- Ayuda a comprender los restantes temas del curriculum, tanto de la educación obligatoria como posterior, donde con frecuencia aparecen gráficos, resúmenes o conceptos estadísticos.

Otro aspecto que fue señalado por Fischbein (1975) es el carácter exclusivamente determinista que el currículo de matemáticas tuvo hasta 1975, y la necesidad de mostrar al alumno una imagen más equilibrada de la realidad, en la que hay una fuerte presencia de fenómenos aleatorios.

Más recientemente, Begg (1997) señala que la estadística es un buen vehículo para alcanzar las capacidades de comunicación, tratamiento de la información, resolución de problemas, uso de ordenadores, trabajo cooperativo y en grupo, a las que se da gran importancia en los nuevos currículos.

Además, la probabilidad y la estadística pueden ser aplicadas a la realidad tan directamente como la aritmética elemental puesto que no requieren técnicas matemáticas complicadas. Por sus muchas aplicaciones, proporcionan una buena oportunidad para mostrar a los estudiantes las aplicaciones de la matemática para resolver problemas reales, siempre que su enseñanza se lleve a cabo mediante una metodología heurística y activa, enfatizando la experimentación y la resolución de problemas.

Ejemplo 4.1. Castro, Rico y Castro (1987) describen las siguientes destrezas relacionadas con el pensamiento numérico que pueden ejercitarse en la clase de estadística: anotar números, hacer el recuento por clases, ordenarlas, comparar las frecuencias, obtener las sumas acumuladas y el total, representar gráficamente, calcular porcentajes, proporciones, medias y desviaciones típica, usar tablas numéricas de doble entradas, substituir números en fórmulas, aproximar y redondear resultados e interpretar datos numéricos.

Actividad 4.1. Analizar la presencia de la estadística en la formación profesional y universitaria en diversas especialidades. ¿Cómo es la preparación matemática de los alumnos que acceden a esta formación? ¿Qué problemas didácticos se plantean?

Actividad 4.2. Estudiar las sesiones que componen la revista Teaching Statistics. ¿Qué podemos concluir sobre los conocimientos de un profesor para enseñar estadística a los alumnos de 9 a 19 años?

Cuando tenemos en cuenta el tipo de estadística que queremos enseñar y la forma de llevar a cabo esta enseñanza debemos reflexionar sobre los fines principales de esta enseñanza que son:

- Que los alumnos lleguen a comprender y a apreciar el papel de la estadística en la sociedad, incluyendo sus diferentes campos de aplicación y el modo en que la estadística ha contribuido a su desarrollo.
- Que los alumnos lleguen a comprender y a valorar el método estadístico, esto es, la clase de preguntas que un uso inteligente de la estadística puede responder, las formas básicas de razonamiento estadístico, su potencia y limitaciones.

4.2. Fenomenología estocástica

Las situaciones de tipo aleatorio tienen una fuerte presencia en nuestro entorno. Si queremos que el alumno valore el papel de la probabilidad y estadística, es importante que los ejemplos y aplicaciones que mostramos en la clase hagan ver de la forma más amplia posible esta fenomenología que analizamos a continuación. Al tratar de clasificar la fenomenología del azar vamos a utilizar cuatro grandes grupos de fenómenos que rodean al hombre: su mundo biológico, físico, social y político.

- **Nuestro mundo biológico**

Dentro del campo biológico, puede hacerse notar al alumno que muchas de las características heredadas en el nacimiento no se pueden prever de antemano: el sexo, color de pelo, peso al nacer, etc. Algunos rasgos como la estatura, número de pulsaciones por minuto, recuento de hematíes, etc., dependen incluso del momento en que son medidas.

Otras aplicaciones se refieren al campo de la medicina. La posibilidad de contagio o no en una epidemia, la edad en que se sufre una enfermedad infantil, la duración de un cierto síntoma, o la posibilidad de un diagnóstico correcto cuando hay varias posibles enfermedades que presentan síntomas parecidos varían de uno a otro chico. El efecto posible de una vacuna, el riesgo de reacción a la misma, la posibilidad de heredar una cierta enfermedad o defecto, o el modo en que se determina el recuento de glóbulos rojos a partir de una muestra de sangre son ejemplos de situaciones aleatorias.

Cuando se hacen predicciones sobre la población mundial o en una región dada para el año 2050, por ejemplo, o sobre la posibilidad de extinción de las ballenas, se están usado estudios probabilísticos de modelos de crecimiento de poblaciones, de igual forma que cuando se hacen estimaciones de la extensión de una cierta enfermedad o de la esperanza de vida de un individuo.

En agricultura y zootecnia se utilizan estos modelos para prever el efecto del uso de fertilizantes o pesticidas, evaluar el rendimiento de una cosecha o las consecuencias de la extensión de una epidemia, nube tóxica, etc. Por último, y en el ámbito de la psicofisiologia, observamos el efecto del azar sobre el cociente intelectual o en la intensidad de respuesta a un estímulo, así como en los tipos diferentes de caracteres o capacidades de los individuos.

- **El mundo físico**

Además del contexto biológico del propio individuo, nos hallamos inmersos en un medio físico variable. ¿Qué mejor fuente de ejemplos sobre fenómenos aleatorios que los meteorológicos?. La duración, intensidad, extensión de las lluvias, tormentas o granizos; las temperaturas máximas y mínimas, la intensidad y dirección del viento son variables aleatorias. También lo son las posibles consecuencias de estos fenómenos: el volumen de agua en un pantano, la magnitud de daños de una riada o granizo son ejemplos en los que se presenta la ocasión del estudio de la estadística y probabilidad.

También en nuestro mundo físico dependemos de ciertas materias primas como el petróleo, carbón y otros minerales; la estimación de estas necesidades, localización de fuentes de energía, el precio, etc., están sujetos a variaciones de un claro carácter aleatorio.

Otra fuente de variabilidad aleatoria es la medida de magnitudes. Cuando pesamos, medimos tiempo, longitudes, etc., cometemos errores aleatorios. Uno de los problemas que se puede plantear es la estimación del error del instrumento y asignar una estimación lo más precisa posible de la medida. Por último, citamos los problemas de fiabilidad y control de la calidad de los aparatos y dispositivos que usamos: coche, televisor, etc.

- **El mundo social**

El hombre no vive aislado: vivimos en sociedad; la familia, la escuela, el trabajo, el ocio están llenos de situaciones en las que predomina la incertidumbre. El número de hijos de la familia, la edad de los padres al contraer matrimonio, el tipo de trabajo, las creencias o aficiones de los miembros varían de una familia a otra.

En la escuela, ¿podemos prever las preguntas del próximo examen?; ¿quién ganará el próximo partido? Para desplazarnos de casa a la escuela, o para ir de vacaciones, dependemos del transporte público que puede sufrir retrasos. ¿Cuantos viajeros usarán el autobús? ¿Cuantos clientes habrá en la caja del supermercado el viernes a las 7 de la tarde?

En nuestros ratos de ocio practicamos juegos de azar tales como quinielas o loterías. Acudimos a encuentros deportivos cuyos resultados son inciertos y en los que tendremos que hacer cola para conseguir las entradas. Cuando hacemos una póliza de seguros no sabemos si la cobraremos o por el contrario perderemos el dinero pagado; cuando compramos acciones en bolsa estamos expuestos a la variación en las cotizaciones,...

- **El mundo político**

El Gobierno, a cualquier nivel, local, nacional o de organismos internacionales, necesita tomar múltiples decisiones que dependen de fenómenos inciertos y sobre los cuales necesita información. Por este motivo la administración precisa de la elaboración de censos y encuestas diversas. Desde los resultados electorales hasta los censos de población hay muchas estadísticas cuyos resultados afectan las decisiones de gobierno y todas estas estadísticas se refieren a distintas variables aleatorias relativas a un cierto colectivo. Entre las más importantes citaremos: el índice de precios al consumo, las tasas de población activa, emigración - inmigración, estadísticas demográficas, producción de los distintos bienes, comercio, etc., de las que diariamente escuchamos sus valores en las noticias.

4.3. La naturaleza de las matemáticas

- **Matemáticas: ¿construcción o descubrimiento?**

Actualmente se reconoce la importancia que tiene la visión de las matemáticas sobre la práctica docente. Podemos describir dos concepciones extremas sobre la matemática y sus aplicaciones a problemas de la vida real.

La primera consiste en pensar que los objetos matemáticos (conceptos, teoremas, propiedades, algoritmos) tienen una existencia objetiva, ideal, independiente del sujeto y de la realidad a la que se aplican y de la cultura. Por ejemplo, se pensaría que las propiedades de los conjuntos numéricos tienen una *"existencia real"* y no pueden ser cambiadas por otras diferentes. Si se tiene esta creencia, quedaría justificada una enseñanza en la que se de un peso fuerte a los conceptos y técnicas. Las aplicaciones, problemas no serían tan importantes, ya que se supone que el alumno sería capaz de aplicar correctamente los conceptos una vez aprendidos.

Una segunda concepción considera que las matemáticas son *una construcción humana* que surge como consecuencia de la necesidad de resolver problemas; los objetos matemáticos serían consecuencia de un proceso de negociación social y están a evolución. Por ejemplo, se pensaría que podríamos haber construido los conjuntos numéricos con otros convenios y otras propiedades y que podemos definir nuevas propiedades sobre los conjuntos numéricos, que sean relevantes para algún problema nuevo.

Si se tiene esta postura, estaremos convencidos que las matemáticas y sus aplicaciones deben mantenerse en íntima relación a lo largo del currículo. Los estudiantes deberían ver la necesidad de cada concepto antes de que le sea presentado (o incluso mejor, antes de que los estudiantes lo creen por sí mismos).

Las matemáticas deben aparecer como una respuesta natural al entorno físico, biológico y social en que el hombre vive. Los estudiantes deben ver, por sí mismos, que las matemáticas son necesarias con el fin de comprender los problemas de la Naturaleza y la Sociedad.

- **Tres características de la matemática**

Por otro lado, autores que se preocupan por la filosofía de las matemáticas, como Wittgestein y Lakatos, Davis y Hersch (1982) y Ernest (1991) distinguen en las Matemáticas al menos tres aspectos esenciales que deben ser tenidos en cuenta en la enseñanza:

- Las matemáticas constituyen una actividad de resolución de problemas, socialmente compartida. Los problemas pueden tener relación con el mundo natural y social o bien ser internos a la propia matemática. Como respuesta o solución a estos problemas externos surgen y evolucionan progresivamente los objetos matemáticos (conceptos, procedimientos, teorías, ...).
- Las matemáticas son un lenguaje simbólico en el que se expresan los problemas y las soluciones encontradas. Como todo lenguaje, implica unas reglas de uso que hay que conocer y su aprendizaje ocasiona dificultades similares al aprendizaje de otro lenguaje no materno.
- Las matemáticas constituyen un sistema conceptual, lógicamente organizado y socialmente compartido. La organización lógica de los conceptos, teoremas y propiedades explica también gran número de las dificultades en el aprendizaje, ya que es difícil trocear y secuenciar convenientemente los contenidos matemáticos para ser enseñados.

En nuestra concepción de las matemáticas, que se describió en el Capítulo 3 se asumen estos tres principios, así como el carácter constructivo del conocimiento matemático, pero se enriquece con la clasificación de los diferentes elementos del significado de los objetos matemáticos y con la distinción entre las perspectivas institucional y personal.

4.4. Algunas teorías educativas

Entre las diversas teorías educativas que han influenciado la educación matemática, describiremos brevemente algunas que pueden ser potencialmente útiles en la organización de la enseñanza de la estadística y que pensamos son compatibles con una teoría de la enseñanza derivada de nuestra concepción sobre el significado de los objetos matemáticos.

- **Constructivismo**

Como indicamos en el Capítulo 3, Piaget postula que la experiencia, la actividad y el conocimiento previo son los que determinan el aprendizaje. El conocimiento es construido activamente por el sujeto y no recibido pasivamente del entorno. El niño trata de adaptarse al mundo que le rodea y cuando una idea nueva se presenta sobre otras ya existentes se crea un *"conflicto cognitivo"* o "*desequilibrio*" en su estado mental, que se resuelve mediante un proceso de *"equilibración"* (*asimilación* y *acomodación).*

La posibilidad de aprender depende del conocimiento previamente adquirido y del desarrollo intelectual del alumno, que ha sido objeto de otras teorías sobre *niveles o etapas* de desarrollo, especialmente las de Piaget y Van Hiele. Las etapas son fases, de modo que los sujetos que están en una misma fase tienen un modo de razonamiento similar y la progresión de una etapa a otra siempre sigue un cierto patrón. En la teoría de Piaget se contemplan principalmente tres etapas: preoperatoria, operaciones concretas y operaciones abstractas, que son precedidas por un periodo sensorio-motor.

La idea de que el aprendizaje es un proceso gradual de asimilación y acomodación de concreta en el *currículo en espiral* propugnado por Bruner. Para este autor, un mismo tema puede ser enseñado con diversos niveles de formalización a un mismo alumno según su nivel de desarrollo. Por ello, se habla de la espiral curricular por la cual se van tratando sucesivamente los mismos temas a lo largo del currículo.

- **Resolución de problemas**

Una forma de llevar a cabo el aprendizaje por descubrimiento es mediante la *resolución de problemas* por parte del alumno. Se trata no de problemas rutinarios sino que requieren una verdadera actividad de resolución por parte del alumno quien, al resolverlo, ha aprendido algo nuevo, aunque no está claro que haya transferencia de este aprendizaje a otros problemas diferentes. Polya ha sido uno de los impulsores de la resolución de problemas, recomendando el uso de estrategias generales y su método en cuatro fases:

- Comprensión del problema;
- Concepción de un plan de resolución;
- Ejecución del plan;
- Examen retrospectivo de la solución hallada.

- **El papel de las herramientas semióticas**

También Vigostky basa su teoría del aprendizaje en la *actividad*, considerando que el sujeto no sólo responde a los estímulos sino que actúa sobre ellos y los transforma, usando instrumentos mediadores de dos tipos: herramientas y símbolos (o signos). La herramienta actúa sobre el estímulo y lo modifica. El signo actúa como mediador entre la persona y el entorno y no modifica el estímulo, sino a la persona que lo usa como mediador. El aprendizaje consiste en la interiorización progresiva de instrumentos mediadores. Diferencia dos niveles de desarrollo de la persona:

1. El desarrollo efectivo que es lo que el individuo es capaz de hacer por sí mismo, sin ayuda de mediadores externos o personas, es decir con los mediadores ya interiorizados;
2. El desarrollo potencial que es loa que se es capaz de hacer con la ayuda de otras personas o mediadores externos. La diferencia entre estos dos niveles de desarrollo es la *zona de desarrollo potencial*.

Actividad 4.3. Van Hiele describe los siguientes niveles de aprendizaje, refiriéndose particularmente al caso de la geometría:

- *Visual: se distinguen las figuras por su forma. No se ven componentes ni relaciones entre las figuras y sus atributos.*
- *Descriptiva: se diferencian los componentes de las figuras y se establecen algunas de sus propiedades en forma experimental.*
- *Abstracto: se establecen propiedades mediante razonamiento informal, se establecen definiciones abstractas y se pueden distinguir condiciones necesarias y suficientes para determinar un concepto.*
- *Deductivo lógico-formal: comienza el razonamiento matemático y s opera con leyes lógicas, axiomas y teoremas.*

¿Piensas que podríamos hallar estos niveles en el aprendizaje de algún concepto estadístico?

- **La Teoría de Situaciones Didácticas**

Esta teoría toma elementos de las anteriores, aunque se apoya en el carácter específico del conocimiento matemático y en la importancia particular de las situaciones de enseñanza y la gestión de las mismas por parte del profesor.

Para Brousseau (1986) una situación didáctica se establece entre un grupo de alumnos y un profesor que usa un medio didáctico -incluyendo los problemas, materiales e instrumentos, con el fin específico de ayudar a sus alumnos a reconstruir un cierto conocimiento. Para lograr el aprendizaje el alumno debe interesarse personalmente por la resolución del problema planteado en la situación didáctica. Se diferencian cuatro tipo de situaciones didácticas:

- Situación de acción. donde se resuelve el problema planteado;
- Situaciones de formulación/comunicación: en las que el alumno debe poner por escrito para otra persona la solución hallada, lo que le hace usar el lenguaje matemático;
- Situaciones de validación: donde se pide a los alumnos las pruebas de que su solución es la correcta. En caso de que no sea así, el debate con los compañeros les permite descubrir los puntos erróneos;
- Situaciones de institucionalización: Tienen como fin dar un estatuto "oficial" al

nuevo conocimiento aparecido, ponerse de acuerdo en la nomenclatura, formulación, propiedades, para que pueda ser usado en el trabajo posterior.

De acuerdo con Brousseau, el trabajo intelectual del alumno debe ser en ciertos momentos comparable al de los propios matemáticos: El alumno debería tener oportunidad de investigar sobre problemas a su alcance, formular, probar, construir modelos, lenguajes, conceptos, teorías, intercambiar sus ideas con otros, reconocer las que son conformes con la cultura matemática, adoptar las ideas que le sean útiles.

Por el contrario, el trabajo del profesor es en cierta medida inverso del trabajo del matemático profesional: En lugar de "inventar" métodos matemáticos adecuados para resolver problemas, debe "inventar" problemas interesantes que conduzcan a un cierto conocimiento matemático. Esta formulación del aprendizaje matemático se corresponde con las teorías constructivistas, ampliamente asumidas. Por ejemplo, en el National Council of Teachers of Mathematics (U.S.A.) se indica:

> "*El aprendizaje debe venir guiado por la búsqueda de respuestas a problemas -primero a un nivel intuitivo y empírico; más tarde generalizando; y finalmente justificando (demostrando)"* (NCTM, 1989).

Actividad 4.4. Investigar en un libro de historia de la Matemática cuáles fueron los problemas que dieron origen al concepto de media aritmética. Busque en un libro de estadística matemática a nivel universitario un ejemplo de presentación decontextualizada de este concepto. Analizar en un libro de texto de bachillerato los contextos empleados por el autor para "contextualizar" la enseñanza de la media aritmética.

Puesto que para nosotros el significado de un objeto matemático emerge de la actividad de resolución de problemas, mediada por herramientas semióticas, es claro que el aprendizaje tiene un carácter constructivo y también asumimos la teorías de resolución de problemas y de Vigotsky. La teoría de enseñanza que mejor se acomoda a nuestra visión de las matemáticas es la de Brousseau, ya que tiene en cuenta los cinco tipos de elementos que consideramos en el significado de un objeto matemático.

Todavía añadiríamos más y es que las situaciones didácticas deben ser diseñadas a partir del análisis previo del significado institucional que se desea transmitir a los alumnos, en sus diversos tipos de elementos. Trataremos de ejemplificar estos principios de enseñanza en la construcción de los Proyectos presentados en el Capítulo V.

4.5. Factores que condicionan el currículo

Cuando queremos comprender y analizar el currículo de estadística en la enseñanza primaria y secundaria, es preciso tener en cuenta una serie de factores, lo que explican la dificultad e importancia que tiene el diseño curricular.

- **Diferentes sentidos del termino curriculum**

En primer lugar, la palabra currículo puede usarse son diferentes significados. Podemos definir el currículo como *"plan operativo que detalla qué matemáticas* *necesitan conocer los alumnos, qué deben hacer los profesores para conseguir que sus alumnos desarrollen sus conocimientos matemáticos y cuál debe ser el contexto en el que tenga lugar el proceso de enseñanza-aprendizaje*" (N.C.T.M., 1989). Sin embargo, podemos hacer referencia al término con diferentes matices:

- *Currículo pretendido:* Plan escrito para el sistema escolar, que incluye especificaciones de lo que se tiene que enseñar, en que secuencia y a qué edades. Puede también contener sugerencias de métodos de enseñanza y de modos de evaluación del aprendizaje. El curriculum pretendido se conoce a través de la lectura de los documentos oficiales.
- *Currículo enseñado*. Lo que se enseña realmente en las aulas, Puede ser evaluado mediante la observación de las clases o entrevistas o informes de los profesores.
- *Currículo aprendido*. Ideas y habilidades que los estudiantes realmente aprenden en clase y en el trabajo personal. Puede ser evaluado mediante pruebas o exámenes a los estudiantes.
- *Currículo retenido*. Las ideas y habilidades que los estudiantes recuerdan algún tiempo después de la instrucción.
- *Currículo ejercitado*. Lo que los estudiantes pueden luego aplicar en sus estudios posteriores o en su vida profesional.

Dimensiones del currículo

Rico (1997) considera cuatro dimensiones principales en torno a las cuales se organiza la reflexión curricular, y que han sido analizadas a lo largo de este capítulo:

- *La dimensión cultural/ conceptual:* La propia disciplina académica es una de las fuentes, ya que el conocimiento se organiza con frecuencia en las propias disciplinas. Para el caso de la estocástica, el fundamento del currículum estaría en la propia ciencia estadística. El conocimiento a aprender, las capacidades para practicar y la lógica de la instrucción tendrían en cuenta los principios de esta disciplina.
- *La dimensión cognitiva*: la reflexión sobre el aprendizaje y la comprensión y cómo se producen ha sido analizado en la sección sobre teorías de enseñanza y aprendizaje. El diseño del currículo se basa en el estudio de la forma en que aprenden los sujetos, requiere un conocimiento de como los estudiantes piensan, como cambian sus concepciones y como podemos dirigir este cambio. El curriculum se diseña para mover al estudiante de su posición actual, anticipar su aprendizaje natural y dirigirlo hacia la meta deseada. La base del curriculum son las concepciones ingenuas de los estudiantes sobre la incertidumbre. Los temas se distribuirían a lo largo del tiempo, en función del desarrollo de la capacidad de cálculo de los alumnos.
- *La dimensión ética/formativa*: el análisis de la educación, sus agentes y funciones, planificación, control y optimización. Un punto importante es el conocimiento de los profesores: Se ha mostrado que los profesores son los que finalmente determinan el currículo enseñado y que, a veces, presentan también dificultades en el razonamiento estocástico e inseguridad por esta materia. Por tanto, la

implantación real pasa por la motivación y formación de los profesores que deben llevar a cabo la enseñanza.

- La dimensión social: La utilidad social del conocimiento y su valoración social, que ha sido analizada indirectamente en las secciones sobre fines de la educación estadística y sobre fenomenología de la estocástica.La naturaleza del mundo, la presencia de una cierta clase de fenómenos (los estocásticos) justifica la necesidad de su estudio a través de la disciplina en cuestión. Las necesidades sociales fundamentan la necesidad de educar ciudadanos que comprendan los informes de las encuestas y estudios y que hagan decisiones inteligentes en situaciones de incertidumbre.

4.6. Estadística en los currículos oficiales de educación secundaria

Recogemos, a continuación las disposiciones oficiales sobre esta material

- **Real Decreto 1345/1991 de 6 de Septiembre, que establece el currículo de la Educación Secundaria Obligatoria.**

En el Bloque 4, *Interpretación, representación y tratamiento de la* información encontramos las siguiente directrices

Contenidos:

1. *Información sobre fenómenos causales:* Dependencia funcional, Características de las gráficas; Funciones elementales
2. *Información sobre fenómenos aleatorios:* Obtención de información sobre fenómenos aleatorios; Las muestras y su representatividad; Frecuencias absolutas, relativas y porcentajes, Gráficas estadísticas usuales
3. *Parámetros estadísticos:* Los parámetros centrales y de dispersión como resumen de un conjunto de datos estadísticos. Algoritmos para calcular parámetros centrales y de dispersión sencillos.
4. *Dependencia aleatoria entre dos variables*

Procedimientos

1. Utilización de distintos lenguajes:

- Utilizar e interpretar el lenguaje gráfico, teniendo en cuenta la situación que se representa y utilizando el vocabulario y los símbolos adecuados.
- Interpretar y elaborar tablas numéricas a partir de conjuntos de datos, gráficas o expresiones funcionales, teniendo en cuenta el fenómeno al que se refieren.
- Utilizar los parámetros de una distribución y analizar su representatividad, en relación con el fenómeno a que se refieren.

2. Algoritmos y destrezas

- Utilizar distintas fuentes documentales (anuarios, revistas, bancos de datos, etc.) para obtener información de tipo estadístico.
- Analizar la representatividad de las muestras estadísticas.
- Elección de los parámetros más adecuados para describir una distribución en función del contexto y de la naturaleza de los datos y obtención de los

mismos utilizando los algoritmos tradicionales o la calculadora.
- Detección de falacias en la formulación de proposiciones que utilizan el lenguaje estadístico.
- Construcción de gráficas a partir de tablas estadísticas o funcionales, de fórmulas y de descripciones verbales de un problema, eligiendo en cada caso el tipo de gráfica y medio de representación más adecuado.
- Detección de errores en las gráficas que pueden afectar su interpretación.

3. Estrategias generales
- Planificación y realización individual y colectiva de tomas de datos, utilizando técnicas de encuesta, muestreo, recuento y construcción de tablas estadísticas.
- Formulación de conjeturas sobre el comportamiento de una población, de acuerdo con los resultados relativos a una muestra de la misma.

Actitudes
1. Referentes a la apreciación de la matemática:
- Reconocer y valorar la utilidad del lenguaje gráfico y estadístico para resolver problemas científicos y de la vida cotidiana.
- Valoración de la incidencia de los nuevos medios tecnológicos en el tratamiento y representación gráfica de informaciones de índole estadística
- Reconocimiento y valoración de las relaciones entre el lenguaje gráfico y otros conceptos y lenguajes matemáticos.
- Curiosidad por investigar relaciones entre magnitudes o fenómenos.
- Sensibilidad, interés y valoración crítica del uso de los lenguajes gráfico y estadístico en informaciones y argumentaciones sociales, políticas y económicas.

2. *Referentes a la organización y hábitos de trabajo:*
- Reconocimiento y valoración del trabajo en equipo como la manera más eficaz para realizar determinadas actividades (planificar y llevar a cabo experiencias, tomas de datos, etc.).
- Sensibilidad y gusto por la precisión, el orden y la claridad en el tratamiento y presentación de datos y resultados relativos a observaciones, experiencias y encuestas.

- **Junta de Andalucía: Orden de 28 de Octubre de 1993 por la que se establecen criterios y orientaciones para la elaboración de Proyectos Curriculares de Centro, secuenciación de contenidos, así como distribución horaria y de materias optativas en la Educación Secundaria Obligatoria.**

En el Bloque sobre *Tratamiento de la Información Estadística y del Azar* se indica que el manejo de datos ya organizados y bien presentados, su representación e interpretación, constituyen actividades de gran importancia en nuestra época tan marcada por la información y la tecnología. Los alumnos deben llegar a conseguir, a lo largo de la etapa, una cierta destreza en el manejo de datos .Esta adquisición se basa en conceptos y procedimientos seleccionados entre aquellos que, apareciendo como incorporables a su desarrollo cognitivo, resultan de mayor eficacia y potencia para:

- Interpretar informaciones que aparecen en los anuarios, periódicos y otros medios de comunicación.
- Enunciar y poner en práctica estrategias en sencillas situaciones de juego.
- Expresar predicciones razonables ante fenómenos aleatorios y tomar decisiones coherentes con esas predicciones.
- Comenzar a reconocer la dificultad de una toma de datos para propósitos estadísticos.

La mencionada selección de conceptos y procedimientos necesita el equilibrio entre las destrezas cuya adquisición se va a proponer a los alumnos y las actitudes de éstos, por ejemplo:

- El dominio de técnicas de cálculo irá acompañado de la expresión clara de los que se está haciendo.
- Se propondrán análisis de datos basados en las diferentes medidas de centralización: media, mediana y moda.
- El trabajo con probabilidades facilitará la adquisición de convicciones acerca de los axiomas de probabilidad.
- La adquisición y uso, por los alumnos, de una terminología precisa en estadística y probabilidad constituye un proceso acumulativo cuatrienal.

La especial motivación que presentan los alumnos de esa etapa en temas relacionados con el entorno, deportes, modas o juegos, favorece la realización de investigaciones y estudios de carácter estadístico. Puede ser útil el que los alumnos participen en el proceso completo:

- Formulación y refinamiento de las preguntas.
- Planificación y recogida de los datos.
- Organización y representación de los datos mediante tablas y gráficos.
- Análisis y resumen de la información.
- Elaboración de conjeturas y, en su caso, toma de decisiones.
- Comunicación de la información y crítica de las conclusiones.

Se tratará siempre de un proceso a largo plazo en el que las tres primeras fases necesitarán de sucesivos refinamientos, con el consiguiente efecto sobre las restantes fases: por ello, en este supuesto parece imperativo sugerir que el tema de investigación sea cercano al alumno y permita tales refinamientos. El contenido se divide en dos ciclos:

Primer Ciclo: Estadística descriptiva:

Los datos reales que son relevantes par los alumnos o, alternativamente, los datos recogidos por ellos mismo, proporcionan un primer acercamiento práctico a las técnicas estadísticas: elección de muestras, redacción de preguntas, agrupamientos de los datos, elaboración de tablas de frecuencias, redacción de informes provisionales.

Algunas representaciones gráficas utilizadas en estadística. Aunque es la primera ocasión en que se estudia, los alumnos han conocido ejemplos de gráficas a través de los medios de comunicación, etapa anterior o en este ciclo; en ese mismo área se están iniciando a la representación gráfica de fenómenos, de manera fundamentalmente cualitativa.

Reconocer, interpretar y elaborar distintas representaciones: Diagramas de barras, diagramas de sectores, histogramas, pictogramas, troncos y hojas, gráficos de las cajas. Los alumnos deberían llegar a enunciar criterios para seleccionar aquellas que consideren más pertinentes para dar una buena imagen de los datos que están manejando y conseguir identificar algunas incorrecciones que suelen cometerse en la elaboración de las gráficas.

Las medidas de centralización: moda, mediana. Se orientarán las intervenciones más que a realizar cálculos reiterados a reconocer la necesidad de este tipo de medidas y a indicar en qué sentido hay dependencia entre una descripción de los datos y el uso de una u otra medida.

Al comparar tablas de datos con ayuda de las medidas de centralización se propicia la elaboración de conjeturas y la necesidad de establecer otros tipos de medidas que describan más adecuadamente la colección de informaciones disponibles. La posibilidad o no de calcular las diferentes medidas de centralización a partir de tablas de datos, permitirá establecer distinciones entre los datos (cuantitativos y cualitativos, agrupados o no, extremos definidos y un acercamiento intuitivo a la noción de dispersión).

4.7. La evaluación del aprendizaje

- **Significados del término**

La palabra evaluación se usa con diferentes significados: Como examen o como sistema de calificación; como sistema de seleccionar los alumnos para el acceso a un centro o para ser promocionados de un curso a otro o simplemente, como método de informar a los alumnos sobre su progreso.

Desde nuestra concepción de las matemáticas la evaluación es el estudio de la correspondencia entre el significado institucional de los conceptos que se trata de enseñar en una cierta institución escolar y el significado personal construido por los estudiantes y tiene, por tanto, un carácter social.

La evaluación se considera hoy día una parte importante del proceso de instrucción. En NCTM (1991) se concibe la evaluación como un proceso dinámico y continuo de producción de información sobre el progreso de los alumnos hacia los objetivos de aprendizaje. El principal propósito es mejorar el aprendizaje de los alumnos. Otros fines secundarios de la evaluación son:

- Proporcionar a los alumnos información individual sobre qué han aprendido y en qué puntos tienen dificultades.
- Proporcionar información al profesor, a los padres y al centro escolar sobre el progreso y la comprensión de sus alumnos, en general y sobre las dificultades de estudiantes particulares
- Proporcionar a las autoridades educativas o a cualquier agente educativo un indicador global del éxito conseguido en los objetivos educativos.

- **Componentes de la evaluación**

Garfield (1995) señala que un marco para la evaluación debe tener en cuenta los siguientes componentes:

- Qué se evalúa: los conceptos, habilidades, aplicaciones actitudes y creencias concretas sobre los que versa la evaluación;
- Fin de la evaluación: qué se pretende; por qué se recoge la información (por ejemplo, para informar a los alumnos);
- Quien evalúa: el profesor habitual u otro; los mismos alumnos; un alumno a otro;
- El método e instrumento usados, de los descritos en el apartado anterior.

La importancia de la evaluación dentro del funcionamiento del sistema de enseñanza en su conjunto, y de los propios sistemas didácticos es ampliamente reconocida (Giménez, 1997).. Si las situaciones y formas de evaluación son deficientes, se corre el riesgo de que esta evaluación ejerza una influencia negativa ocasionando serias distorsiones entre los objetivos educativos y lo que finalmente aprenden por los estudiantes. Por tanto, un componente imprescindible de cualquier currículo es la propuesta de criterios sobre qué evaluar, cómo (procedimientos y circunstancias) y cuándo.

Un punto importante es reconocer la complejidad de la función evaluadora, debido a que ésta debe atender a las múltiples facetas del conocimiento estadistico (comprensión conceptual y procedimental, actitudes). Es necesario usar todo un sistema para recoger datos sobre el trabajo y rendimiento del alumno y no es suficiente evaluarlo a partir de las respuestas breves dadas a preguntas rutinarias en una única evaluación (o examen). Por ejemplo, en los "Estándares" de evaluación del NCTM, que resumimos a continuación, reflejan bien los diversos aspectos del conocimiento matemático, que se deben tener en cuenta en la planificación de la instrucción y en su correspondiente evaluación.

El primer estándar se refiere a la *coherencia*: los métodos y tareas que se usen para evaluar el aprendizaje de los alumnos deben ser coherentes con el curriculo, esto es, con el plan global de formación matemática diseñado. Deben corresponderse con las situaciones didácticas propuestas, con los objetivos, contenidos y métodos de enseñanza y aprendizaje. Asimismo, la valoración del aprendizaje de los alumnos debe basarse en información obtenida a partir de diversas fuentes, métodos y formas de evaluación. La evaluación de los conocimientos matemáticos de los estudiantes, de acuerdo con los "Estándares" citados, debe tener en cuenta las siguientes dimensiones:

Comprensión conceptual:

- Dar nombre, verbalizar y definir conceptos;
- Identificar y generar ejemplos válidos y no válidos;
- Utilizar modelos, diagramas y símbolos para representar conceptos;
- Pasar de un modo de representación a otro;
- Reconocer los diversos significados e Interpretaciones de los conceptos;
- Identificar propiedades de un concepto determinado y reconocer las condiciones que determinan un concepto en particular;
- Comparar y contrastar conceptos.

C*onocimiento procedimental*:

- Reconocer cuándo es adecuado un procedimiento;
- Explicar las razones para los distintos pasos de un procedimiento;
- Llevar a cabo un procedimiento de forma fiable y eficaz;
- Verificar el resultado de un procedimiento empíricamente o analíticamente;
- Reconocer procedimientos correctos e incorrectos;
- Generar procedimientos nuevos y ampliar o modificar los ya conocidos;
- Reconocer la naturaleza y el papel que cumplen los procedimientos dentro de las matemáticas.

Resolución de problemas:

- Formular y resolver problemas;
- Aplicar diversas estrategias para resolver problemas;
- Comprobar e interpretar resultados;
- Generalizar soluciones.

Formulación y comunicación matemática.:

- Expresar ideas matemáticas en forma hablada, escrita o mediante representaciones visuales;
- Interpretar y juzgar ideas matemáticas, presentadas de forma escrita, oral o visual;
- Utilizar el vocabulario matemático, notaciones y estructuras para representar ideas, describir relaciones.

Razonamiento matemático.

- Utilizar el razonamiento inductivo para reconocer patrones y formular conjeturas;
- Utilizar el razonamiento deductivo para verificar una conclusión, juzgar la validez de un argumento y construir argumentos válidos;
- Analizar situaciones para hallar propiedades y estructuras comunes;
- Reconocer la naturaleza axiomática de las matemáticas.

Actitud o disposición hacia las matemáticas:

- Confianza en el uso de las matemáticas para resolver problemas, comunicar ideas y razonar;

- Flexibilidad al explorar ideas matemáticas y probar métodos alternativos para la resolución de problemas;
- Deseo de continuar hasta el final con una tarea matemática;
- Interés, curiosidad e inventiva al hacer matemáticas;
- Inclinación a revisar y reflexionar sobre su propio pensamiento y su actuación;
- Valorar la aplicación de las matemáticas a situaciones que surjan de otras materias y de la experiencia diaria;
- Reconocer el papel que cumplen las matemáticas en nuestra;
- cultura, y el valor que tienen como herramienta y como lenguaje.

- **Instrumentos de evaluación:**

Los instrumentos y situaciones de evaluación deben estar en consonancia con la enseñanza. Por tanto, será necesario que las situaciones de evaluación tengan en cuenta las actividades realizadas en clase y el conocimiento matemático que se derivan de ellas. Garfield (1995) indica las siguientes posibilidades de instrumentos de evaluación para el caso de la estadística:

- Observación sistemática de las intervenciones de los alumnos en clase a lo largo del curso;
- Revisión periódica de los cuadernos y apuntes de los alumnos;
- Pruebas específicas escritas tipo examen;
- Preguntas realizadas en clase a alumnos particulares o a toda la clase;
- Encuestas breves sobre lo que han aprendido o lo que han encontrado confuso en una clase particular, sobre la actitud de los estudiantes, el contenido del curso o su visión de la estadística;
- Trabajos de síntesis sobre un tema o una colección de lecturas. que muestren la comprensión y capacidad de síntesis;
- Proyectos de análisis de datos individuales o colectivos;
- Test de opciones múltiples;
- Problemas para realizar en la clase o como trabajo de casa;
- "Dossier" donde el profesor va recogiendo información diversa acerca del alumno;
- "Diario" elaborado por los alumnos con resúmenes de lo aprendido en clase.

Actividad 4.6. Se desear determinar la capacidad de los alumnos de 1º Curso de ESO para calcular promedios a partir de un conjunto de datos sencillos, para decidir si es preciso reforzar el aprendizaje. Identifica los componentes del proceso de evaluación y prepara un posible instrumento de evaluación para esta situación.

Actividad 4.7. El ítem siguiente es un ítem de respuesta abierta. Analiza el conocimiento que se trata de evaluar. Da una lista de las posibles respuestas correctas e incorrectas que esperas de tus alumnos y transforma el enunciado en un ítem de opciones múltiples que tenga en cuenta estas respuestas previsibles.

Item. *Durante un mes 500 alumnos llevan a cabo un registro diario del número de horas que pasan viendo la televisión. En promedio el número de horas por semana que los estudiantes* *dedicaron a ver la televisión fue de 23. Pidieron al profesor las calificaciones de cada uno de los estudiantes y descubrieron que los estudiantes que dedicaban más horas a ver la televisión tenían peores calificaciones. Indica algunas razones por las que este experimento no demuestra que ver la televisión es la causa de las bajas calificaciones de los alumnos.*

Actividad 4.8. Imagina que vas a evaluar a los alumnos en un curso de estadística a partir de un Proyecto de análisis de datos. Cada alumno elige y elabora personalmente su Proyecto. Elabora una pauta de calificación de los Proyectos que te permita discriminar los conocimientos y actitudes del alumno hacia la estadística.

Actividad 4.9. A continuación incluimos una escala de actitudes hacia la estadística (algunas preguntas puntúan en sentido inverso).Identifica las dimensiones en las actitudes que se trata de evaluar y las preguntas que puntúan en sentido inverso.

Para cada una de las siguiente preguntas indica en la escala 1 a 5 tu grado de acuerdo, según el siguiente convenio

1	2	3	4	5
Fuertemente en desacuerdo	No estoy de acuerdo	Indiferente	De acuerdo	Fuertemente de acuerdo

1. Uso a menudo la información estadística para formar mis opiniones o tomar decisiones
1 2 3 4 5
2. Es necesario conocer algio de estadistica para ser un consumidor inteligente.
1 2 3 4 5
3. Ya que es fácil menir con la estadísticao, no me fío de ella en absoluto.
1 2 3 4 5
4. La estadística es cada vez más importante en nuestra sociedad y saber estadística será tan necesario como saber leer y escribir.
1 2 3 4 5
5. Me gustaría aprendear más estadística si tuviese oportunidad.
1 2 3 4 5
6. Debes ser bueno en matemáticas para comprender los conceptos estadísticos básicos.
1 2 3 4 5
7. Cuando te compras un coche nuevo es preferible preguntar a los amigos que consultar una encuesta sobre la satisfacción de usuarios de distintas marcas, en una revista de información al consumidor.
1 2 3 4 5
8. Me parecen muy claras las frases que se refieren a la probabilidad, como, por ejemplo, las probabilidades de ganar una lotería.
1 2 3 4 5
9. Entiendo casi todos los términos estadísticos que encuentro en los periódicos o noticias.
1 2 3 4 5
10. Podría explicar a otra persona como funciona una encuesta de opinión.
1 2 3 4 5

4.8. Materiales y recursos didácticos

Para Alsina y cols. (1988), el término "material" agrupa todos aquellos objetos, aparatos o medios de comunicación que pueden ayudar a descubrir, entender o consolidar conceptos fundamentales en la diversas fases del aprendizaje, es decir, el material manipulativo, software didáctico y no didáctico, libros, problemas, juegos, y, en general, todos los instrumentos que facilitan el trabajo en la clase de estadística.

- **Material manipulativo**

El material manipulativo desempeña un papel básico en los primeros niveles de enseñanza por la necesidad que tienen los niños de contar con referentes concretos de los conceptos abstractos que tratamos de enseñarles. Por ejemplo, en los Estándares Curriculares y de Evaluación para la Educación Matemática del NCTM (1991) se recomienda que los estudiantes experimenten y simulen modelos de probabilidad. Glayman y Varga (1975) sugieren que para iniciar al niño en la probabilidad en ideas específicas es preciso echar mano de medios específicos. No basta contar con que en la vida cotidiana el niño encuentre sucesos aleatorios que le permitan descubrir las leyes de las probabilidades; es necesario introducir estrategias, apelar a la actividad para suscitar la curiosidad natural del niño, conducirle a que se enfrente con la realidad y luchar contra las ideas falsas que pueda tener.

Una característica particular de los experimentos aleatorios es su carácter irreversible, destacado por Piaget e Inhelder (1951). Una vez producido un resultado aleatorio, no es posible volver al estado inicial con seguridad. Por ejemplo, supongamos que una ruleta numerada con los dígitos 1 a 10 tiene su aguja parada en el número 1 y después de impulsar la aguja ésta se para sobre el número 5. ¿Quiere decir esto que al impulsar de nuevo la aguja en el sentido contrario volveremos al número 1?

Esta falta de reversibilidad impide un apoyo directo del material concreto en el estudio de los fenómenos aleatorios. Mientras que en el aprendizaje de las operaciones básicas en aritmética o de los conceptos geométricos, el niño puede explorar, mediante composición y descomposición una operación o propiedad dada (por ejemplo, al estudiar la suma de dos números o la descomposición de un polígono en triángulos) con ayuda de material concreto (mediante las regletas o mediante modelos de polígonos), ello no es posible en el caso de los fenómenos aleatorios. El análisis de un experimento aleatorio va más allá del resultado inmediato (suceso producido) y requiere la consideración de todos los sucesos que podrían haber ocurrido, es decir de la consideración del espacio muestral del experimento.

Por otro lado, en aritmética o geometría el niño puede repetir las veces que quiera el trabajo con el modelo, obteniendo siempre el mismo resultado. Una repetición de la experiencia aleatoria, debido a su mismo carácter, no puede servir para comprobar un resultado. Es necesario establecer un sistema de registro que permita reflexionar sobre las experiencias.

Es decir, el estudio de los experimentos aleatorios implica la organización de una situación de recogida y análisis de datos estadísticos. Es únicamente cuando se recogen los resultados de una serie larga de tales experimentos, cuando se producen los fenómenos de convergencia y se aprecian regularidades en el comportamiento de los fenómenos aleatorios. De la misma forma, si lo que partimos es de una clase de estadística (situación de análisis de datos) será difícil olvidar completamente los problemas de variabilidad, muestra, población, generalizabilidad de las conclusiones, posibilidad de predicción, aleatoriedad. De este modo estadística y probabilidad son dos caras de una misma moneda, de ahí que prefiramos hablar de "estocástica".

Para el estudio de las probabilidad y estadística los profesores y alumnos pueden preparar con gran facilidad materiales similares a los comercializados. Los generadores aleatorios de tipo físico pueden reducirse a cuatro tipos principales: "dados", "bolas" , "ruletas" y "barajas de cartas".

- *Dados*: Cualquier objeto que presente un número finito de posiciones distintas, como "trompos", monedas, fichas bicolores, etc. Concretizan el experimento aleatorio más simple posible: Cuando el espacio muestral es finito y los resultados son equiprobables. Una generalización sería que el dispositivo fuese sesgado, por lo cual los dos resultados no fuesen equiprobables (tales como en el caso de la chincheta, un dado sesgado, etc.).
- *Bolas en urnas*: cualquier colección de objetos (fichas, cartas, regletas, bloques, etc.) que se puedan mezclar antes de extraer de una urna, caja, etc., de modo que todas tengan la misma posibilidad de salir. La diferencia con el caso anterior es que permite, por un lado introducir el número de elementos diferentes que se desee, en lugar de estar restringido a un número dado de elementos. Por otro lado, la mayor o menor proporción de elementos de cada tipo en la urna permite cambiar a voluntad las probabilidades de los distintos sucesos elementales. La extracción de bolas en urnas un número dado de veces puede dar origen a cuatro tipo de experimentos diferentes: extracciones con reemplazamiento y sin reemplazamiento, ordenadas y no ordenadas (que se corresponden con las cuatro operaciones combinatorias básicas). Finalmente, la extracción de bolas en urnas puede dar lugar al estudio de problemas de inferencia si se usa una urna opaca y el alumno desconoce la composición de la urna
- *Ruletas,* u otro dispositivo que permita plantear problemas de probabilidades geométricas: Pueden servir las ruletas construidas con cartulina por los propios alumnos, con áreas rayadas de formas diversas. Como eje de giro de estas ruletas puede utilizarse un lápiz y como aguja un clip sujetapapeles desplegado por uno de sus laterales. Este tipo de material permite utilizar sectores con áreas iguales o desiguales, y sucesos con probabilidades iguales o diferentes. Permite también plantear problemas de probabilidades compuestas mediante un juego de coronas y sectores circulares. Podemos generar variables continuas, tales como la distancia al radio más cercano.
- *Barajas de cartas,* o colecciones de tarjetas con datos referentes a más de un atributo. Estos generadores contienen datos por lo menos de dos variables aleatorias diferentes (palo de la baraja y número/figura). Por lo tanto, pueden dar origen al estudio de la asociación/ independencia de variables, además del estudio de experimentos compuestos.

Otros materiales utilizables:

- Chinchetas(o cualquier otro objeto irregular).
- Tablas de números aleatorios, o generadores de números al azar con ordenador, tablas de las distribuciones de probabilidades básicas.

- Diagramas de barras, cartesianos, en árbol, de Venn, y cualquier tipo de gráfico estadístico.
- Botellas de muestreo.
- Canales, por los que se dejan caer bolas (máquina de Galton).
- Dianas y dardos para tirar al blanco.
- Anuarios estadísticos, tablas y colecciones de datos estadísticos tomados de la prensa o recogidos por los propios alumnos, junto con proyectos asociados a los mismos.
- Plantillas para el dibujo de funciones de densidad o modelos planos o espaciales de dichas distribuciones (por ejemplo, representaciones gráficas con ordenador de dichas distribuciones).
- Materiales curriculares específicos; en particular los desarrollados por proyectos centrados en áreas de interés para el alumno (por ejemplo, materiales del School Council Projects, School Mathematics Projects, Quantitative Literacy Project, etc.).
- Otros materiales, como la regla para tomar tiempos de reacción, instrumentos de medida para comparar medidas y estimaciones de las mismas, etc.

- ***Simulación***

Un uso característico del material en estocástica es la simulación, es decir, sustituir un experimento aleatorio difícil de observar en la realidad, por otro equivalente. La simulación permite condensar el experimento en el tiempo y en el espacio y operar con el experimento simulado para obtener conclusiones válidas para el experimento original.

Además, la simulación proporciona un método "universal" para obtener una estimación de la solución de los problemas probabilísticos, que no tiene paralelo en otras ramas de la matemática. En la enseñanza de la estocástica en secundaria la simulación cobra papel importante, ya que ayuda al alumno a conocer las diferencias entre la probabilidad experimental y la teórica.

Cualquier problema probabilístico implica una serie de experimentos aleatorios compuestos de una determinada manera. Cada uno de estos experimentos puede ser "simulado" con un modelo de urnas convenientemente escogido (incluso los modelos continuos de probabilidad). El experimento compuesto se obtiene componiendo las "urnas" correspondientes a los experimentos simples (hiperurna) y la repetición del experimento global, junto con el análisis de los datos producidos permite una solución aproximada del problema. En este sentido la urna con bolas de colores (fichas, tarjetas) es un "material universal", válido para estudiar cualquier problema o concepto probabilístico. Muchos problemas complejos se resuelven hoy día mediante simulación y el mostrar a los alumnos ejemplos sencillos de esta técnica puede servir para mostrar su aplicabilidad a campos y problemas reales.

- **Calculadoras gráficas**

Las calculadoras gráficas se consideran a veces como la tecnología del futuro, debido a que su conste cada vez menor hace creíble que, en el futuro, cada estudiante pueda disponer de su propia calculadora. Entre las nuevas posibilidades que ofrecen a la enseñanza de la estadística, citamos:

- Transmisión de datos (entre calculadoras o calculadora y ordenador). Es posible, por ejemplo, tomar datos de internet, sobre un tema de interés y transmitirlo a la calculadora, sin necesidad de tener que grabarlos a mano;
- Opciones de manejo de listas;
- Posibilidad de transformación de los datos;
- Cálculos estadísticos básicos para una y varias variables;
- Gráficos estadísticos usuales;
- Posibilidad de ser programadas;
- Generador de números aleatorios y tablas estadísticas básicas.

- **Juegos**

Como indica Corbalán (1994), el esfuerzo por hacer atractiva la clase de matemática tiene que ser mayor precisamente en los temas más abstractos, entre los que cita la asignación de probabilidades. Justamente en este campo no hay que olvidar que el desarrollo matemático comenzó precisamente a partir de los juegos de azar, que además constituyen en la actualidad una actividad socialmente muy extendida, incluso en forma institucional. Otro objetivo importante puede ser el hacer a los alumnos conscientes del problema social de la ludopatía.

4.9. Ordenadores y enseñanza de la estadística

Es un hecho evidente la influencia que los ordenadores han tenido en el desarrollo de la Estadística como disciplina y en facilitar el acceso a la estadística a un número y diversidad cada vez mayor de usuarios, aumentando las demandas de formación básica en estadística. Los ordenadores han aumentado, por un lado, el número de contenidos estadísticos a enseñar, incluyendo el uso adecuado del software, sin el cual es hoy día impensable la realización del análisis de datos en cualquier campo de aplicación. También ha habido un cambio en los contenidos, prestándose mayor importancia a los aspectos interpretativos y conceptuales y menor a los procedimentales y algoritmos de cálculo.

El uso de ordenadores en la enseñanza de la Estadística está recibiendo atención creciente, tanto por parte de los profesores como de los investigadores, como se puede constatar en Shaughnessy et al., (1996) y Shaughnessy, Garfield y Greer (1997). Por ejemplo, el Instituto Internacional de Estadística organizó una *Round Table Conference* sobre el ordenador en la enseñanza de la estadística en Austria en 1970 y otra en Camberra en 1984. La *Round Table Conference*, organizada por IASE en Granada en 1996 se centró en el rol de la tecnología en la enseñanza y aprendizaje de estadística (Garfield y Burrill, 1997).

Los ordenadores han reducido muchas horas dedicadas al cálculo, permitiendo el estudio de mayores conjuntos de datos. Esto ha permitido agregar nuevos tópicos a la enseñanza de la estadística. La gran ventaja de los ordenadores es su naturaleza dinámica, su velocidad, y el creciente rango de software que permite a los estudiantes experimentar y explorar todos los aspectos de los procesos estadísticos, desde la planificación de la muestra o del diseño experimental hasta la recolección y el manejo de datos, la simulación y el análisis, para interpretar y comunicar los resultados.

También en los congresos ICOTS I, II, III, IV y V se ha dado gran importancia a los ordenadores y a la investigación basada en los ordenadores. En el último ICOTS V el uso de ordenadores en la enseñanza de la estadística fue uno de los núcleos temáticos del congreso y en muchos de los trabajos presentados, no sólo dentro de este núcleo sino en los restantes se analizaron desarrollos de software, materiales para la enseñanza, instrumentos de evaluación basados en el uso de ordenadores, experiencias didácticas e investigaciones sobre el efecto del uso del ordenador en el aprendizaje y las actitudes de los alumnos (Pereira-Mendoza y cols., 1998).

Hay varias formas de uso del ordenador en la enseñanza de la estadística, todas ellas de gran importancia y que suponen una revolución sobre la forma en que se debe enseñar y se debe aprender estadística. Biehler (1997) clasifica el software estadístico según sus funciones educativas. Los *micromundos* posibilitan la realización de experimentos interactivos, mediante simulaciones y visualizaciones exploratorias, que ayudan a los estudiantes a conceptualizar la estadística. Las *herramientas* permiten a los estudiantes practicar la estadística del mismo modo que lo hacen los estadísticos profesionales. En el caso particular del análisis exploratorio de datos, estas herramientas deberían facultar a los estudiantes para hacer un trabajo interactivo, exploratorio y abierto, utilizando software flexible, fácil de usar y aprender.

Además, las clases de estadística proporcionan actividades interesantes para introducir al alumno en el uso de recursos informáticos habituales, como procesadores de texto y hoja de cálculo, así como para el aprendizaje del manejo de la calculadora científica y gráfica. Es interesante animar a los chicos a escribir un informe sobre su análisis, ya que la habilidad para producir informes comprensivos y estructurados donde la información estadística se incorpore y presente adecuadamente para apoyar la argumentación será sin duda útil en su futura vida profesional, sea cual fuere y es un medio también para el aprendizaje de los procesadores de texto.

Podemos diferenciar los siguientes tipos de software para la enseñanza:

- *Paquetes estadísticos profesionales*. como por ejemplo: *SPSS, STATGRAPHICS*, etc. Cada uno de estos paquetes tiene una amplia disponibilidad de presentación gráfica y numérica.
- *Software didáctico,*como *Fathom*, (Ben-Zvi, 2000), que es un medio de aprendizaje para análisis exploratorio de datos y álgebra, y se utiliza en secundaria y en cursos introductorios de estadística a nivel de bachillerato y *Sampling Distributions* (DelMas, Garfield y Chance, 1998; Chance, Garfield y DelMas, 1999).
- *Software de uso general*, como las hojas de cálculo, como por ejemplo, *EXCEL*
- *Tutoriales,* que son programas desarrollados para enseñar a los estudiantes sobre habilidades estadísticas específicas o evaluar su conocimiento. Ejemplos de tutoriales utilizados en la enseñanza de la estadística son: *ActivStats* y *ConStats* (Cohen y Chechile, 1997),

- *Software en Internet* material que puede obtenerse "on-line"

- **Cálculo y representación gráfica**

En primer lugar, el ordenador puede y debe usarse en la enseñanza como *instrumento de cálculo y representación gráfica,* para analizar datos recogidos por el alumno o proporcionados por el profesor.

Un problema tradicional en la enseñanza de la Estadística ha sido la existencia de un desfase entre la comprensión de los conceptos y los medios técnicos de cálculo para poder aplicarlos. La solución de los problemas dependía en gran medida de la habilidad de cálculo de los usuarios, que con frecuencia no tenían una formación específica en matemáticas. Hoy día la existencia de programas facilmente manejables permite salvar este desfase y realizar cálculos complejos en pocos segundos sin posibilidad de error. No tiene pues, sentido, hacer perder el tiempo a los alumnos ocupándoles en repetir una y otra vez cálculos tediosos para intentar aumentar su destreza de cálculo, sino que es preferible dedicar ese tiempo a actividades interpretativas y a la resolución de problemas.

La capacidad de graficación de los ordenadores permite también incorporar la filosofía del análisis exploratorio de datos, en que los gráficos y el cambio de uno a otro sistema de representación se usa como herramienta de descubrimiento y análisis. El manejo de diversas formas de representación dinámica e interactiva enriquece el significado de los conceptos mostrados a los estudiantes.

Sin embargo, esta mayor facilidad actual de empleo de procedimientos estadísticos, implica, sin embargo, el peligro del uso no adecuado de la estadística. Acostumbremos, pues, a los alumnos a planificar el análisis que quieren realizar incluso antes de finalizar la construcción de su sistema de datos. Si, por ejemplo, quieren hacer un estudio en su escuela para comparar la intención de voto de chicos y chicas en las próximas elecciones al consejo escolar, deben recoger una muestra lo suficientemente representativa de chicos y chicas en los diferentes cursos escolares y deben recoger datos sobre las principales variables que influyan en esta intención de voto. De otro modo, sus conclusiones pudieran estar sesgadas o ser poco explicativas.

Los paquetes estadísticos tienen, además, la posibilidad de seleccionar datos y variables, y se pueden tener diversas representaciones gráficas y tabulares simultáneamente en la pantalla, y manipular ciertas características de los gráficos como anchura, formato, escala, etc.

Por otro lado, el hecho mismo de realizar los cálculos con ayuda del ordenador, introduce cambios en los elementos intensionales del significado que al alumno debe usar en la resolución de los problemas.

Actividad 4.10. Analiza el razonamiento subyacente en la respuesta de Luisa al resolver el problema con ayuda del ordenador y compara con el razonamiento seguido al calcular a mano los percentiles.

Problema: En el fichero MMF20 (referido a datos sobre capacidad lectora de niños de 6 años), ¿Cuál es el número máximo de errores en comprensión para el 90 por ciento de los alumnos?

Respuesta escrita de Luisa: "*He usado la opción STATS del menú principal del programa y luego he elegido DESCRIPTIVE METHODS y PERCENTILES. Puse el nombre de la variable en la ventana de entrada de datos, escribiendo: MFF20.errorpalab. Después puse 90 en el cuadro de porcentaje y obtuve 5. El número máximo de errores para el 90 por ciento de los niños es 5*".

- **Trabajo con datos reales**

Nos enfrentamos a diario a la necesidad de recoger, organizar e interpretar sistemas complejos de datos y esta necesidad aumentará en el futuro, debido al desarrollo de los sistemas de comunicación y las bases de datos. El punto de comienzo de la estadística debería ser el encuentro de los alumnos con sistemas de datos reales: resultados deportivos de sus equipos favoritos, medios de transporte usados para ir a la escuela, temperatura máxima y mínima a lo largo de un mes; color o tipo de vehículo que pasa por delante de la ventana, etc.

Uno de los objetivos que debiera incluirse en un curso de estadística es capacitar al alumno para recoger, organizar, depurar, almacenar, representar y analizar sistemas de datos sencillos. Este objetivo comienza por la comprensión de las ideas básicas sobre organización *de datos: codificación grabación y depuración*.

De este modo podrán ver que construir un sistema de datos propio y analizarlo no es lo mismo que resolver un problema de cálculo rutinario tomado de un libro de texto. Si quieren que el sistema de datos sea real, tendrán que buscar información cuando les falte, comprobar y depurar los errores que cometen al recoger los datos, añadir nueva información a la base de datos cuando se tenga disponible, Aprenderán a comprender y apreciar más el trabajo de los que realizan las estadísticas para el gobierno y los medios de comunicación. Si comprenden la importancia de la información fiable, se mostrarán más dispuestos a colaborar cuando se les solicite colaboración en encuestas y censos.

En la mayor parte de los conjuntos de datos hay al menos tres componentes: la descripción de las variables, los valores de las variable (campos), que es el cuerpo principal de los datos, y los resúmenes estadísticos de cada variable. Los campos pueden ser de longitud fija o variable, y puede haber campos vacíos. Asimismo, clasificamos las variables según diversas tipologías: cualitativas o cuantitativas; discretas, continuas; nominales, ordinales, datos de intervalo, de razón.

Sobre cada una de estas componentes pueden realizarse operaciones o transformaciones internas (clasificación, recodificación, agrupamiento) y externas (insertar, borrar, seleccionar...). Podemos clasificar variables, clasificar los casos dentro de una variable o clasificar los resúmenes estadísticos, por ejemplo, por su magnitud. Podemos seleccionar casos por los valores de una variable, o seleccionar variables porque sus valores coinciden en una serie de casos. También es posible determinar relaciones entre estos componentes, por ejemplo, de dependencia, implicación, similaridad (dependencia entre variables; similaridad de sujetos; similaridad de variables). Estos tipos de operaciones deben ser presentadas para casos sencillos a los estudiantes, de modo que sean comprendidas.

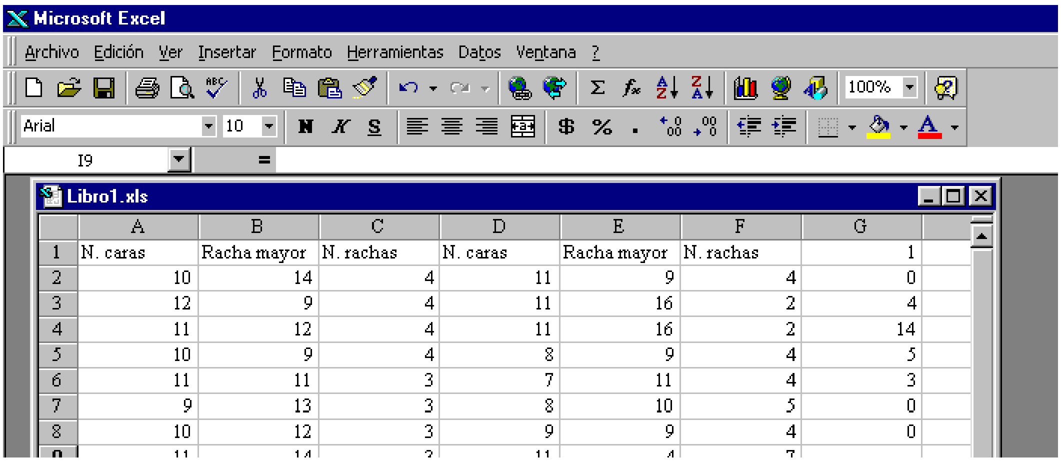
Estos sistemas de datos pueden ser la base de trabajos interdisciplinares en geografía, ciencias sociales, historia, deportes, etc. En el caso de que los datos se tomen de los resultados de experimentos aleatorios realizados en la clase, estaremos integrando el estudio de la estadística y probabilidad.

- **Ficheros de datos y proyectos**

El profesor también puede proporcionar ficheros de datos a los alumnos, para introducir algún tema particular o porque sea difícil de recoger por los propios alumnos. En la actualidad existen colecciones de ficheros de datos disponibles para los profesores, en revistas como Teaching Statistics o en la internet. En los proyectos que se sugieren como ejemplos en el Capítulo 5 incluimos algunos datos tomado del servidor de la revista Journal of Statistical Education.

Los ficheros de datos son fácilmente analizables desde una hoja electrónica, que podrían ser utilizadas por los chicos de 14 o 15 años. En la Figura 3.1. mostramos una hoja electrónica, donde hemos introducido los datos del proyecto 1 (Capítulo V).

Figura 3.1. Pantalla de una hoja electrónica con datos del Proyecto 1

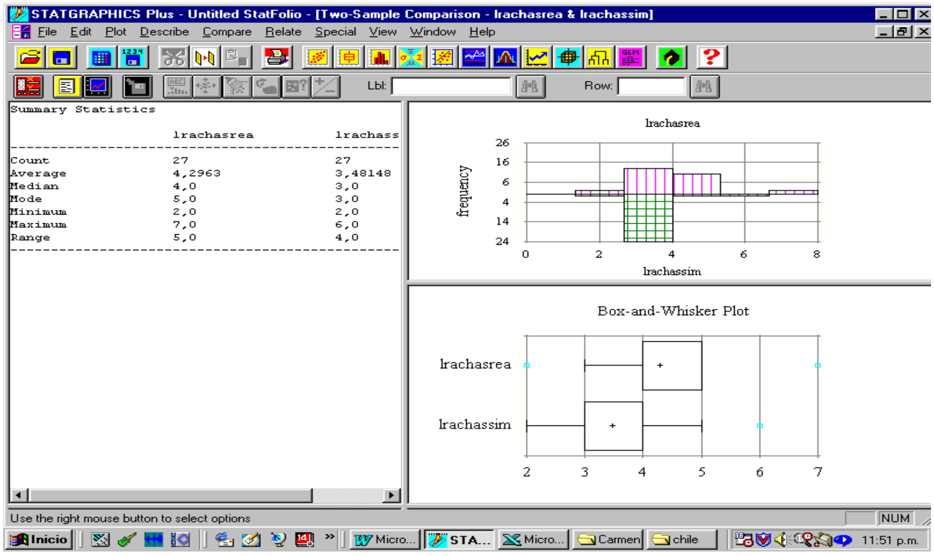


	A	B	C	D	E	F	G
1	N. caras	Racha mayor	N. rachas	N. caras	Racha mayor	N. rachas	1
2	10	14	4	11	9	4	0
3	12	9	4	11	16	2	4
4	11	12	4	11	16	2	14
5	10	9	4	8	9	4	5
6	11	11	3	7	11	4	3
7	9	13	3	8	10	5	0
8	10	12	3	9	9	4	0

Este tipo de recurso proporciona una variedad de gráficos y funciones estadísticas. Por ejemplo, podemos usar la función CONTAR.SI, para calcular las frecuencias absolutas con que aparecen los diferentes valores en la tabla de datos. En la columna G de la hoja electrónica mostrada en la Figura 3.1 hemos calculado las frecuencias de la variable presentada en la columna A (número de caras en la secuencia simulada). La primera celda de esta columna presenta la frecuencia de apariciones del valor 7, que se calcularía mediante la fórmula: =CONTAR.SI(A2:A28; 7).

En el caso de disponer de un paquete estadístico sencillo de manejar, y también con los chicos mayores la capacidad de representación gráfica y análisis se enriquece notablemente. Como vemos en la Figura 3.9 para Statgraphics este tipo de software permite tener al mismo tiempo en la pantalla salidas numéricas y gráficas, cuyas opciones pueden ser modificadas fácil y rápidamente y aumenta por tanto las posibilidades exploratorias.

Figura 3.9. Comparación de la longitud de la racha más larga en secuencias reales y simuladas con Statgraphics



- **Simulación**

Aunque es importante que los alumnos realicen algunas actividades de simulación con apoyo de material manipulativo, como monedas, dados o ruletas y con tablas de números aleatorios, es realmente el ordenador el que proporciona una mayor potencia de simulación.

Las hojas electrónicas proporcionan generadores de números aleatorios, así como de valores de diferentes distribuciones de probabilidad, que pueden, una vez generados, ser analizados con ayuda de los recursos de cálculo y representación gráfica de la hoja.

Las posibilidades son todavía mayores con los programa estadísticos que suelen incluir módulos de estudio de las diferentes distribuciones de probabilidad con representación gráfica y cálculo de valores críticos y áreas bajo la función de densidad. Unido esto a la posibilidad de extracción de muestras de valores de estas distribuciones de tamaño dado, almacenamiento de las mismas en nuevos ficheros de datos, que pueden ser analizados, proporciona una herramienta muy interesante para la introducción de ideas de inferencia. En el capítulo V hemos introducido el uso de estas herramientas en uno de los Proyectos presentados como ejemplos.

Finalmente existen programas didácticos específicos para explorar conceptos estocásticos, desde los más elementales a los más avanzados, como, por ejemplo los procesos estocásticos.

4.10. Recursos en internet

Una nueva dimensión en la enseñanza y la práctica estadística está siendo marcada por Internet. En esta sección realizamos un resumen de los trabajos de Snell (1996), Batanero (1998) y Galmacci (2001) que presentan algunos de los recursos disponibles en la red.

- **Cursos y materiales didácticos**

El prototipo de los cambios previsibles con las nuevas tecnologías es el curso

Chance, desarrollado en cooperación por varias universidades americanas. Este curso presenta el uso de los conceptos básicos de estadística en la prensa. Un boletín electrónico proporciona trimestralmente resúmenes de artículos de prensa que usan conceptos de estadística. Adicionalmente una base de datos en WWW contiene planificación de cursos que han utilizado este material y una guía para el profesor.

Las clases de un curso de este tipo se organizan del modo siguiente: se elige un articulo reciente y se preparan algunas preguntas relacionadas. Los estudiantes, en grupos, leen el artículo e intentan contestar las preguntas formuladas u otras relacionadas que surjan durante la discusión. Todo ello se utiliza como base para introducir un tema de estadística relacionado con el contenido del artículo. Otros recursos similares se listan a continuación:

Baker R. (University of Saskatchewan, CA), *Basic principles of statistical analysis*, http://duke.usask.ca/~rbaker/stats.html.
de Leeuw J. (University of California, Los Angeles), *Statistics UCLA*, http://www.stat.ucla.edu/textbook/.
Department of Statistics (University of South Carolina), *Interactive Statistics and the GASP Initiative - Globally Accessible Statistical Procedures,* http://www.stat.sc.edu/rsrch/gasp/.
Hopkins W.G. (Sportscience), *A New View of Statistics*, http://www.sportsci.org/resource/stats/.
Lane D.M. (Rice University), *HyperStat*, http://davidmlane.com/hyperstat/index.html .
Martínez de Lejarza J. & Martínez de Lejarz I. (Universidad de Valencia) http://www.uv.es/~lejarza/estadistic.htm.
NWP Associates, Inc. (The Pennsylvania State University), *Investigating Statistics*, http://espse.ed.psu.edu/statistics/investigating.htm.
Snell J.L., Peter Doyle, Joan Garfield, Tom Moore, B. Peterson, N. Shah (Dartmouth College), *CHANCE*, http://www.dartmouth.edu/~chance/.
Stirling W. D. (Massey University, NZ), *CAST - Computer Assisted Statistics Teaching*, http://cast.massey.ac.nz/.
Stockburger D. W. (Southwest Missouri State University), *Introductory Statistics*, http://www.psychstat.smsu.edu/sbk00.htm.
StatSoft, Inc. (Tulsa, OK), *Electronic Statistics Textbook*, http://www.statsoft.com/textbook/stathome.html.

- **Revistas electrónicas**

*The Journal of Statistics Education e*s una revista publicada desde 1993, electrónicamente, cuyo tema es la enseñanza de la estadística a nivel universitario. Se recibe libre de costo por internet y también puede conseguirse los números atrasados o un artículo suelto directamente a partir de la dirección donde están archivados. La universidad de North Carolina mantiene una base de datos relacionada con esta revista donde se contiene otra serie de recursos para la enseñanza de la estadística. Una diferencia con una revista convencional es que es posible a los lectores mandar comentarios a un artículo o hacer búsquedas automatizadas de artículos sobre un cierto tema. Muchos de estos comentarios serán seleccionados para pasar a ser parte del archivo y, por tanto, del propio artículo. Incluye "teaching bits’ que proporciona resúmenes de artículos de interés para los profesores de estadística.

- **Conjuntos de datos**

Un recurso interesante es el llamado *"data sets and stories*", donde se acumulan conjuntos de datos, junto con su descripción y algunas indicaciones de sus posibles usos en la enseñanza. Los datos se pueden recuperar en formato útil para la mayor parte de paquetes estadísticos, hojas de cálculo y calculadoras gráficas. Otros servidores que presentan colecciones de datos para la enseñanza se listan a continuación:

American Statistical Association, *JSE Data Archive*, http://amstat.org/publications/jse/archive.htm.
Behrens J. (Arizona State University), *Dr. B's Data Gallery*, http://seamonkey.ed.asu.edu/~behrens/classes/data_gallery/.
Cornell University, DASL - The Data and Story Library, http://lib.stat.cmu.edu/DASL/.
de Leeuw J. (University of California, Los Angeles), *Case Studies*, http://www.stat.ucla.edu/cases/.

- **Grupos de discusión o trabajo**

Se puede ser miembro activo de un grupo de discusión sobre la enseñanza de la estadística por correo electrónico. Los temas pueden ir desde pedir la solución a un problema o sugerencias sobre como resolver la dificultad que tiene un estudiante, anuncios de nueva bibliografía, intercambio de material didáctico, etc. Estas discusiones también se archivan para posibles búsquedas sobre un tema. Dos grupos relevantes son el grupo de discusión *Edstat-L* mantenido en el servicio JSE y el *IASE Statistics Education Research Group* (Stated-List) mantenido actualmente en la universidad de Granada. Este último publica también un boletín distribuído por correo electrónico con resúmenes de trabajos de investigación de sus miembros y noticias sobre reuniones, proyectos y recursos de interés para la investigación en el área.

- **Centros de recursos**

Algunas páginas web preparan listas de recursos para la enseñanza y aprendizaje de nociones estocásticas, como las siguiente:

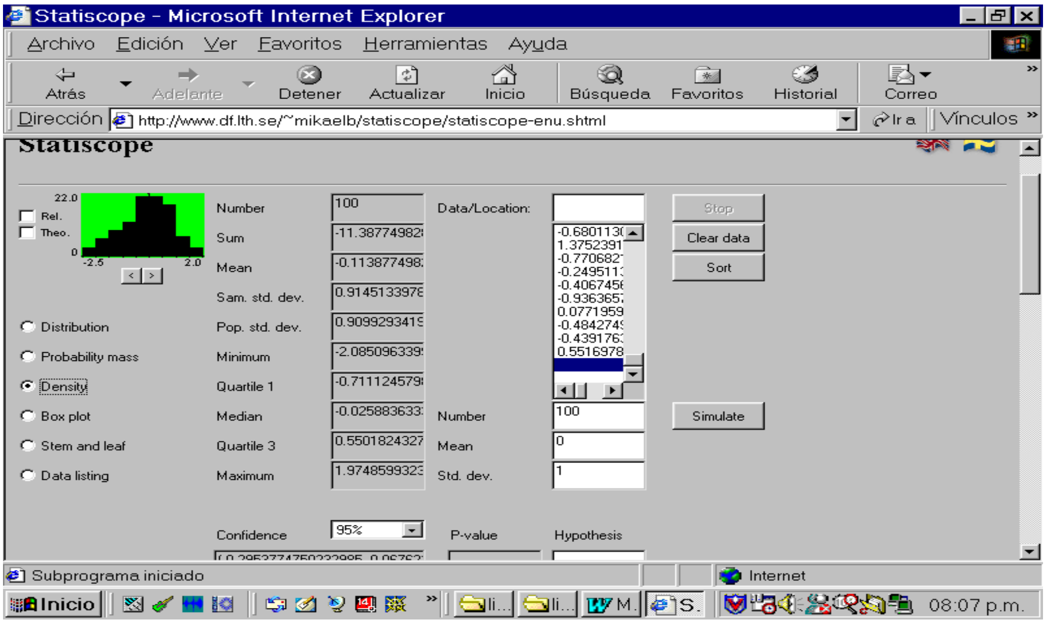
ASA, *Center for Statistics Education*, http://amstat.org/education/index.html.
CIRDIS, *Centro Interuniversitario di Ricerca per la Didattica delle Discipline Statistiche*, http://www.stat.unipg.it/CIRDIS/.
CTI Statistics (University of Glasgow), http://www.stats.gla.ac.uk/cti/.
IASE, *International Association for Statistical Education*, http://www.cbs.nl/isi/iase.htm.
NCTM, *National Council of Teachers of Mathematics*, http://www.nctm.org/.
Grupo de Educación Estadística (Universidad de Granada), http://www.ugr.es/~batanero/.
Royal Statistical Society Centre for Statistical Education, http://science.ntu.ac.uk/rsscse/.
SIIP, *The Statistical Instruction Internet Palette*, http://research.ed.asu.edu/siip/.

- **Software**

Existe una gran cantidad de software disponible en Internet, especialmente para la exploración y simulación. Por ejemplo, en la figura 1 mostramos el Statiscope que permite simular datos, introducir datos manualmente o a partir de internet y analizarlos, mostrando la generación de muestras y distribuciones muestrales. Los resultados se puede representar de forma diversa y calcula también intervalos de confianza. Una de las salidas de), *Virtual Laboratories in Probability and Statistics* see muestra en la figura 2. Otros recursos similares se listan a continuación:

Bonnier M. (Värpinge, Lund, Scania, Sweden), *Statiscope*, http://www.df.lth.se/~mikaelb/statiscope/statiscope-enu.shtml.
Claremont Colleges' "Web Interface for Statistics Education", *Statistics Applet made for the WISE Project*, http://www.grad.cgs.edu/wise/appletsf.shtml.
de Leeuw J. (University of California, Los Angeles), *Statistics UCLA*, http://www.stat.ucla.edu/textbook/.
Lewis B. (Kent State University), *Elementary Statistical Java Applets and Tools*, Marden J. (University of Illinois at Urbana-Champaign), The CUWU Statistics Program, http://www.stat.uiuc.edu/~stat100/cuwu/.
Siegrist K. (University of Alabama in Huntsville), *Virtual Laboratories in Probability and Statistics*, http://www.math.uah.edu/stat/.

Figura 1. Statiscope

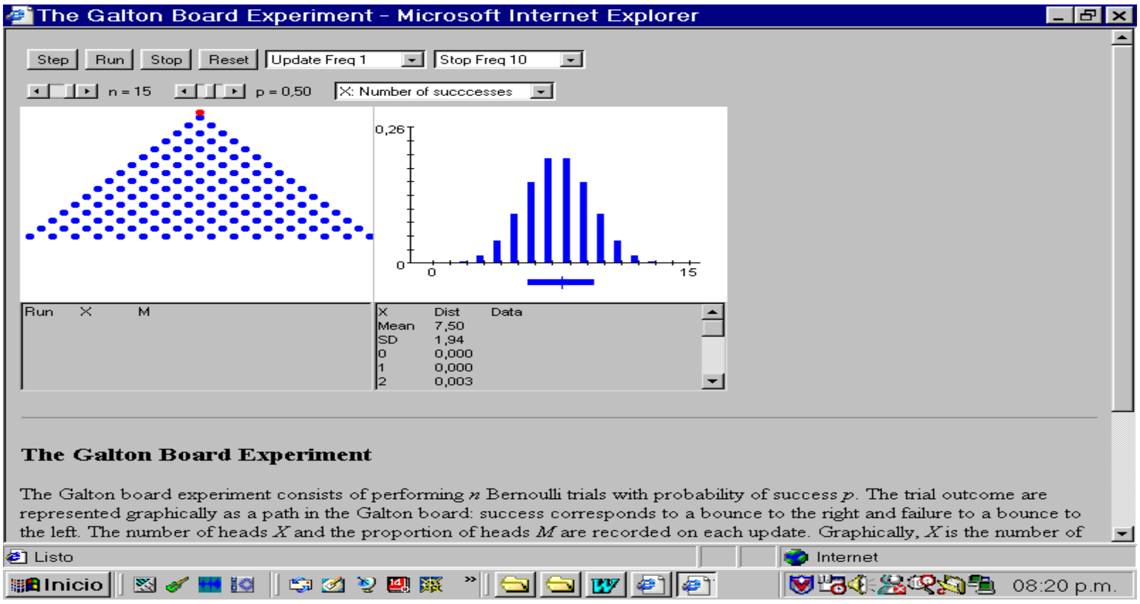


.

En particular, el siguiente software es especialmente adecuado para simulación:

Stark P.B. (University of California, Berkeley), *SticiGui: : Statistical Tools for Internet and Classroom Instruction with a Graphical User Interface*, http://www.stat.berkeley.edu/~stark/SticiGui/index.htm.
Statistics Department (University of Glasgow), *STEPS - STatistical Education through Problem Solving*, http://www.stats.gla.ac.uk/steps/home.html
West R. W. & Ogden R. T. (University of South Carolina), *WebStat*, http://www.stat.sc.edu/webstat/.
Young F. W. (University of North Carolina), *Vista - The Visual Statsistical System*, http://forrest.psych.unc.edu/research/index.html.
Zielman B. (University of), *Statistical Page*, http://huizen.dds.nl/~berrie/.

Figura 2. Simulación del aparato de Galton



5

Ejemplos de Proyectos para la Clase de Estadística

5.1. Introducción

En este último capítulo presentamos algunos ejemplos de proyectos que pueden ser desarrollados en la clase de estadística, describiendo los datos y la forma en que han sido recogido y sugiriendo algunas posibles actividades que propicien la reflexión sobre los conceptos estadísticos y permitan la ejercitación de las diversas representaciones, técnicas y tipos de argumentación. Dependiendo de la edad y conocimientos previos de sus alumnos, de sus intereses, tiempo disponible, el profesor puede suprimir o añadir otras actividades o proyectos.

Los proyectos están concebidos para introducir en la clase una filosofía exploratoria y participativa, en tendencias con las recomendaciones recientes sobre metodología de enseñanza de la estadística. Tienen una estructura común que analizamos en la sección 5.2 y pretenden, en su conjunto, dar una visión del contenido que podría abordarse en la enseñanza secundaria.

Para finalizar el capítulo introducimos una sección donde se sugieren temáticas para otros posibles proyectos. Lo deseable sería que los propios alumnos eligieran el tema en el que quieren trabajar y elaborasen sus propios proyectos en grupos de dos o tres alumnos. Para ser realistas, hemos de reconocer que son pocos los alumnos que se interesan por la estadística y que ésta es una materia aburrida para ellos. Por el contrario, los alumnos pueden interesarse en muchos temas diferentes y llegar a valorar la estadística como instrumento de investigación de los problemas que les gustaría resolver. En algunos países como Estados Unidos o Inglaterra es ya tradicional el celebrar en las escuelas competiciones de proyectos estadísticos y es posible que esta tendencia no tarde en llegar a nuestros centros.

5.2. Estructura de los proyectos y análisis de su contenido

Los proyectos que presentamos se conciben como verdaderas investigaciones asequibles al nivel del alumno, donde tratamos de integrar la estadística dentro del proceso más general de investigación. Es decir, planteamos unos objetivos y preguntas que el alumno debe tratar de contestar. Para ello el alumno necesita recoger datos, que, pueden provenir de diversas fuentes, ser obtenidos mediante diferentes técnicas, y corresponder a diversas escalas de medida y tipos de variables estadísticas (Tabla 5.2.1).

Consideramos importante que, a lo largo de la educación secundaria el alumno tenga oportunidad de apreciar esta diversidad de datos estadísticos. Algunas veces los datos se encuentran disponibles, pero hay que saber localizarlos de diferentes fuentes, como libros o anuarios estadísticos. La Internet proporciona en la actualidad datos para cualquier tema por el que los alumnos estén interesados, bien a partir de servidores estadísticos específicos como Data Sets and Stories library donde los profesores de estadística han puesto sus datos al servicio de la enseñanza, bien recurriendo a organismos oficiales como el INE, Eurostat, Unesco u otros.

Tabla 5.2.1. Tipos de datos en los Proyectos

	P1	P2	P3	P4	P5
Procedencia de los datos					
Anuarios estadísticos			x	x	
Encuestas		x			
Experimento realizado en la clase	x			x	
Internet			x		x
Prensa	x			x	
Simulación	x			x	
Técnica de recogida de datos					
Observación	x	x	x	x	
Encuesta		x		x	
Medida		x			x
Naturaleza de la escala de medida					
Nominal		x		x	
Ordinal		x			
Razón		x		x	x
Variables estadísticas incluídas					
Cualitativa		x	x	x	x
Cuantitativa discreta, pocos valores	x	x		x	
Cuantitativa discreta, necesidad de agrupar			x	x	
Continua		x	x		x

En otras ocasiones los datos son recogidos por los alumnos mediante la realización de una encuesta o a través de un experimento. La encuesta requirirá la elaboración de un cuestionario, fijando los objetivos del mismo, eligiendo las variables explicativas y redactando las preguntas que permitan obtener la información deseada de una forma clara y concisa. La selección de una muestra representativa plantea problemas de tipo teórico y práctico, relacionados con la población objetivo y alcanzada, el marco de muestro, los métodos de selección, la administración del cuestionario y los problemas de no respuesta.

La información que queremos recoger puede corresponder a diversos niveles que se corresponden con diferentes técnicas de obtención de datos: información consciente y conocida (encuesta), información desconocida, pero que puede deducirse de la observación e información no consciente ni observable (medida).

Finalmente es importante considerar la naturaleza de las escalas de medida y tipo de variable estadística, puesto que de ellas depende el método de análisis de datos que se puede aplicar. Puesto que el software estadístico requiere la codificación de los datos en forma numérica, con frecuencia los usuarios que no tienen unos conocimientos estadísticos sólidos aplican métodos de inferencia (por ejemplo un test T de student) a variables para la que este tipo de métodos no son aconsejados.

Los proyectos estadísticos permiten mostrar a los alumnos los campos de aplicacion de la estadística y su utilidad en muchas facetas de la actividad humana. No hay nada que haga más odiosa la estadística que la resolución de ejercicios descontextualizados, donde se pida al alumno calcular la media o ajustar una recta re regresión a un conjunto de números. No hay que olvidar que la estadística es la ciencia de los datos y los datos no son números, sino números en un contexto. En la tabla 5.2.2 mostramos los campos de aplicación que hemos elegido para estos proyectos, a los cuales hay que agregar todos los que se describieron en la sección sobre Fenomenología (4.2). En los proyectos complementarios descritos brevemente en la sección 5.8 presentamos otros proyectos en temas de psicología, relaciones laborales, producción, deportes, educación, salud, comunidad europea y elecciones.

Tabla 5.2.2. Campo de aplicación en los Proyectos

	P1	**P2**	**P3**	**P4**	**P5**
Aleatoriedad, Probabilidad	x			x	
Botánica					x
Consumo				x	
Demografía			x		
Diseño de experimentos; prueba de hipótesis				x	
Economía			x		
Elecciones, votaciones				x	
Estadísticas oficiales; agencias de estadística			x		
Experimentación	x			x	
Fisiología		x			
Pruebas de hipótesis				x	
Sociología		x	x		

El trabajo con proyectos en la clase de estadística motiva a los alumnos, pero plantea el problema de la gestión de la clase, de modo que se oriente a los alumnos hacia el aprendizaje de conceptos y representaciones, la ejercitación de las técnicas y la mejora en sus capacidades de argumentación, formulación de conjeturas y validación. Hemos tratado de presentar en estos proyectos una panorámica de estos diferentes componentes. En la tabla 5.2,3. presentamos el contenido conceptual que podría ser abordado en la enseñanza media.

Aunque la estadística se suele enseñar separada de la probabilidad, durante la enseñanza media, nosotros creemos que esta separación es artificial, puesto que, detrás de cualquier estudio estadístico hay una componente aleatoria. Sin tratar de analizar completamente el tema de probabilidad, para el cual hicimos una propuesta didáctica en nuestro libro anterior (Godino, Batanero y Cañizares, 1987), hemos tratado de relacionar estos dos campos cuando ha sido posible, y en particular, en los proyectos 1 y 4.

Tabla 5.2.3. Conceptos y propiedades estadísticas en los Proyectos

	P1	P2	P3	P4	P5
Aleatoriedad y probabilidad					
Experimento aleatorio; secuencia de resultados aleatorios	x			x	
Sucesos equiprobables,	x			x	
Independencia de ensayos, rachas	x			x	
Codificación	x	x	x	x	x
Variable estadística					
Variable discreta	x	x		x	
Frecuencia absoluta; tabla de frecuencias; distribución de frecuencias	x	x		x	
Frecuencias acumuladas		x			
Agrupación; intervalos, extremos y marcas de clase		x		x	
Posición central					
moda, media,	x	x	x	x	x
mediana	x	x	x		
medias ponderadas, percentiles, rangos de percentiles			x	x	
Valores atípicos y su efecto sobre los promedios			x		
Dispersión					
rango, máximo, mínimo	x	x		x	
Casos centrales, 50% de casos centrales;	x	x		x	
Cuartiles; recorrido intercuartílico		x			
Desviación típica, varianza					x
Asociación					
Tablas de contingencia; frecuencias dobles, marginales y condicionadas		x			
asociación en tablas de contingencia		x			
correlación directa/ inversa; lineal no lineal; correlación y causalidad			x		x
Modelización					
ajuste de modelos a datos bivariantes o multivariantes			x		x
uso de modelos en la predicción			x	x	x
Inferencia					
Población, censo, muestra				x	x
Representatividad y sesgo en el muestreo				x	x
Variabilidad en las muestras; efecto del tamaño sobre la variabilidad				x	x
Parámetro y estadístico				x	
Hipótesis; contraste de una hipótesis				x	
Riesgo en una decisión				x	
Estadística multivariante					
Clasificación					x
Coordenadas, centro de gravedad					x
Distancia, simetría de una distancia					x
Error en la clasificación					x
Función discriminante					x
Similaridad, disimilaridad					x

Introducimos también unas nociones básicas de inferencia, cuya inclusión se recomienda en el currículo de Bachillerato. Somos contrarios a dar una excesiva formalización al tema en este nivel de enseñanza, debido a la complejidad y cantidad de conceptos estadísticos que el alumno debiera conocer para poder introducir, por ejemplo, el test de hipótesis en la teoría de Neyman y Pearson y aplicarlo al caso de comparación de las medias de las muestras.

Nos limitamos a plantear situaciones de estimación sencillas y las ideas del test de signifiación debidas a Fisher, quien no llega a considerar la idea de hipótesis alternativa, ni estudia el error Tipo II. Trataremos sólo de introducir algunos conceptos básicos y principios lógicos de inferencia que serán útiles a los alumnos para un estudio posterior formalizado del tema. Las simulaciones, cálculos probabilísticos sencillos, tablas de números aleatorios y ordenadores serán herramientas suficientes para la resolución de las actividades de inferencia, que no precisarán del estudio previo de las distribuciones en el muestreo.

Hemos introducido, también de forma intuitiva, las técnicas multivariante de clasificación, donde se presenta una buena oportunidad de aplicar los conocimientos de los alumnos sobre algebra lineal y geometría y relacionar, por tanto, estas ramas de la matemática, con la estadística. Puesto que los ordenadores facilitan hoy día la aplicación de estas técnicas, no hay razón para que no puedan ser introducidas desde la educación secundaria.

Tabla 5.2.4. Notaciones y representaciones en los Proyectos

	P1	P2	P3	P4	P5
Términos estadísticos	x	x	x	x	x
Símbolos	x	x	x	x	x
Tablas					
Tablas de frecuencias, datos no agrupados	x				
Tablas de frecuencias, datos agrupados, efecto de la agrupación			x		
Hojas de recogida de datos	x				
Tablas de contingencia		x			
Listado del fichero de datos	x	x	x		x
Gráficos					
Gráficos de puntos	x			x	x
Gráficos de barras, simples adosados o apilados	x		x	x	
Gráfico de sectores	x			x	
Histogramas		x	x		
Polígonos de frecuencias					
Diagramas acumulativos		x			
Polígono de frecuencias acumuladas			x		
Gráficos de tallo y hojas		x	x		
Gráficos de cajas	x	x	x		x
Diagramas de dispersión bivariantes			x		x
Duagrmas de dispersión espaciales					x
Curva empírica de distribución		x			
Dendograma					x
Simulación					
Con material manipulativo	x			x	
Con tablas de números aleatorios o calculadora				x	
Con ordenador				x	

Una parte importante de la estadística es la reducción y presentación de los datos en una variedad de formatos, desde tablas y listados hasta gráficos de tipo diverso (Tabla 5.2.4). Los gráficos estadísticos son, en general sofisticados y el manejo de un gráfico no supone simplemente el cambio de un tipo de representación a otra de un concepto dado.

Por el contrario, en cada gráfico estadístico se representa, además de la distribución una serie de conceptos que varían de un gráfico a otro: frecuencias en el diagrama de barras, densidad de frecuencias y moda en el histograma, mediana y cuartiles, valores atípicos en el gráfico de la caja, etc.

Los gráficos estadísticos presentan convenios de construcción que el alumno debe reconocer y recordar. Es también relativamente fácil producir un gráfico inadecuado o interpretar incorrectamente un gráfico, con lo que se produce una distorsión de la información, sea intencionada o no.Un último tipo de representación es la simulación, en cuanto sirve para representar experimentos aleatorios difícilmente reproducibles, como hemos ya señalado anteriormente. En la tabla 5.2. 5. incluímos las técnicas que se ejercitan en los diferentes proyectos, que van más allá del cálculo y representación gráfica.

Finalmente, en la Tabla 5.6. presentamos las diferentes actitudes que se destacan en los proyectos. Puesto que la estadística es una ciencia cambiante a una gran velocidad, es difícil saber cuáles de los contenidos que hoy impartimos serán últiles a nuestros alumnos.

Tabla 5.2.5. Técnicas y procedimientos estadísticos en los Proyectos

	P1	P2	P3	P4	P5
Recogida y registro de datos experimentales	x			x	x
Búsqueda de datos a partir de anuarios estadísticos o de la Internet			x		
Recogida de datos de observación		x			
Elaboración de encuesta y recogida de datos		x			
Obtención de datos mediante medida		x			
Codificación de datos		x			
Elaboración de tablas de frecuencia; recuento y cálculo de frecuencia	x		x	x	
Elaboración de tablas de doble entrada y cálculo de frecuencias condicionadas y marginales		x			
Elaboración de gráficos	x		x	x	x
Interpretación de tablas y gráficos	x	x	x	x	x
Elaboración de argumentos y conclusiones a partir del análisis de datos	x	x	x	x	x
Estudio de asociación entre variables		x	x		x
Uso de calculadora gráfica, hojas de cálculo o software estadístico	x		x	x	x

Por ello, más importante que el aprendizaje de un concepto o una técnica es enseñar a los alumnos a valorar la estadística, el papel que tiene en el desarrollo científico y económico y la importancia de su colaboración para la obtención de datos estadísticos fiables.

Otras actitudes igualmente importante son poseer un espíritu crítico frente a la información estadística y concienciarse sobre sus propias intuiciones incorrectas.

Tabla 5.2.6. Actitudes que se destacan en los Proyectos

	P1	P2	P3	P4	P5
Reflexión sobre las propias intuiciones incorrectas en relación a los experimentos aleatorios	x			x	
Valoración de la utilidad de la estadística para analizar datos obtenidos mediante experimentación, observación, encuesta o medida	x	x	x	x	x
Valorar la utilidad y complejidad de la elaboración de las estadísticas oficiales y la importancia de colaborar en encuestas y censos para obtener datos fiables.			x	x	
Valoración de la estética y la claridad en la construcción de tablas y gráficos estadísticos	x	x	x	x	x
Concienciar al alumno sobre la posibilidad de que se transmita información sesgada en una gráfica mal construida			x	x	
Reflexión sobre la dificultad de codificación y cómo ésta introduce siempre una simplificación en la realidad;		x			
Valoración de la utilidad de la estadística para identificar relaciones de asociación entre variables;		x			
Reflexión sobre las tendencias y dispersiones en los datos; sobre el excesivo énfasis en los prototipos y el hecho de que éstos con frecuencia son modelos que no se dan en la realidad.		x			
Fomentar un espíritu crítico en el uso de paquetes estadísticos y sus opciones por defecto			x	x	

A continuación describimos detalladamente los proyectos, inluyendo objetivos, descripción de los datos, actividades para la clase y su gestión, actividades de ampliación, análisis del contenido estadístico del proyecto y descripción de posibles errores y dificultades de los alumnos.

5.3. Proyecto 1. Comprueba tus intuiciones sobre el azar

5.3.1. Objetivos

Se trata de realizar un experimento para comprobar si tenemos buenas intuiciones respecto a los experimentos aleatorios. En concreto tratamos de comprobar si somos capaces de simular una secuencia de resultados aleatorios. Utilizaremos el dispositivo aleatorio más sencillo posible: una moneda equilibrada, comparando los resultados obtenidos al lanzar realmente una moneda con los simulados.

La finalidad principal es hacer reflexionar al alumno sobre el hecho de que nuestras intuiciones sobre el azar nos engañan con frecuencia. También se les quiere mostrar la utilidad de la estadística en la prueba de nuestras hipótesis o teorías (en este caso la hipótesis de que nuestras intuiciones sobre los fenómenos estocásticos son correctas).

Alumnos: Puesto que las variables a tratar son discretas y las actividades no introducen conceptos estadísticos complejos, el proyecto podría ser adecuado para alumnos a partir de 13-14 años, es decir, desde el comienzo de la educación secundaria.

Se sugiere empezar con preguntas similares a las siguientes y realizar en clase una discusión colectiva.

1. ¿Cómo piensas que deberían ser los resultados de lanzar una moneda 20 veces seguidas? ¿Serías capaz de escribir 20 resultados de lanzar una moneda (sin lanzarla realmente, sino como tú pienses que debieran salir) de forma que otras personas piensen que has lanzado la moneda en realidad. O, ¿podría otra persona adivinar que estás haciendo trampa?

5.3.2. Los datos

Los datos son producidos como resultado del experimento que será realizado por cada uno de los alumnos de la clase. Se le proporciona a cada alumno una pauta cuadriculada, dándole la siguiente consigna:

2. Vamos a comprobar qué tal son tus intuiciones respecto a los resultados aleatorios. Abajo tienes dos cuadrículas. En la primera de ellas escribe 20 resultados sin realizar realmente el experimento. En la segunda mitad lanza la moneda 20 veces y escribe los resultados obtenidos. Pon C para cara y + para cruz.

Una vez que los alumnos han realizado el experimento tendrán diferentes resultados. Mis resultados han sido los siguientes:

C	C	+	C	+	+	+	C	C	+	C	+	C	+	+	C	C	C	+	+

+	C	+	C	+	+	C	+	C	C	C	C	+	+	+	+	+	C	+	+

5.3.3. Preguntas, actividades y gestión de la clase

Cuando todos los alumnos han finalizado la realización del experimento se puede plantear preguntas similares a las que reproducimos a continuación.

3. ¿Cómo podremos distinguir una secuencia realmente aleatoria de otra que hemos inventado?

Se dejará algún tiempo para pensar y a continuación se organiza una discusión colectiva. Seguramente algún alumno sugerirá contar el número de caras y cruces que debe ser aproximadamente igual, ya que hay las mismas posibilidades para la cara que para la cruz.

4. *Pero, ¿hemos de obtener exactamente 10 caras y 10 cruces? ¿Qué pasa si obtenemos 11 caras y 9 cruces ? ¿Y si obtenemos 18 y 2? ¿Que os parece si comparamos el número de caras en las secuencias real y simulada de todos los alumnos de la clase?*

Para realizar esta comparación se recogen los datos de todos los alumnos de la clase, tanto del número de caras en las secuencias simuladas como en las reales, para proceder, primeramente al análisis de cada una de estas dos variables y luego a la comparación de las principales diferencias en su distribución. Para ejemplificar la realización de la actividad utilizaremos los resultados obtenidos en una clase de 27 alumnos, quienes obtuvieron los siguientes números de caras en las secuencias simuladas.

10, 12, 11,10, 11, 9, 10, 11, 9, 10, 10, 10, 7, 10, 10, 10, 10, 12, 11, 10, 9, 10, 10, 9, 10, 12, 11

5. Hemos recogido el número de caras en las secuencias simuladas por cada alumno de la clase ¿Como podríamos organizar y resumir estos datos? ¿Cuáles son el valor mínimo y máximo obtenido? ¿Cómo representar los datos de modo que sepamos cuántas veces aparece cada valor? ¿Cuál es el valor más frecuente?

El profesor ayudaría a los chicos a identificar el valor máximo y mínimo y a organizar un recuento y tabla de frecuencias como la Tabla 5.3.1, haciéndoles ver su utilidad para resumir la información. Para chicos mayores se podría agregar las columnas de frecuencias relativas y frecuencias acumuladas.

Tabla 5.3.1. Número de caras en las secuencias simuladas

Número de caras	Recuento	Frecuencia
7	x	1
8		0
9	xxxx	4
10	xxxxxxxxxxxxxx	14
11	xxxxx	5
12	xxx	3
Total		27

Figura 5.3.1. Número de caras en secuencias simuladas. Gráfico de puntos

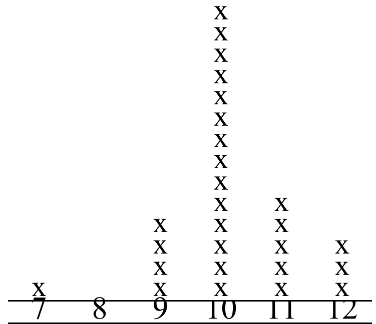


Figura 5.5.2. Gráfico de barras

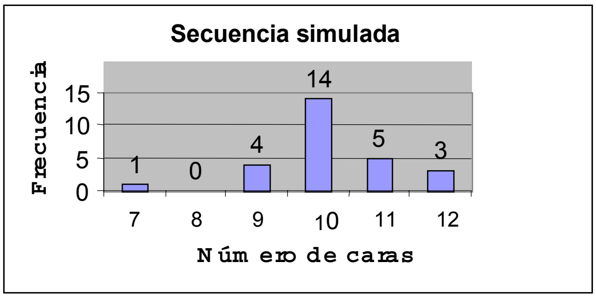
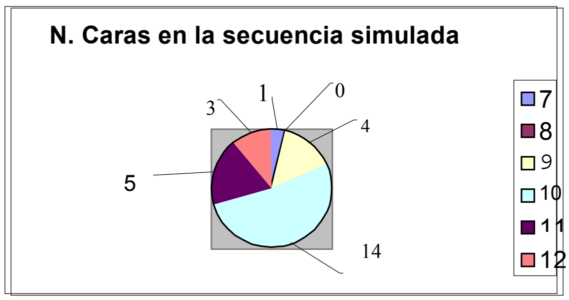


Figura 5.5.3. Gráfico de sectores



El gráfico de puntos (Figura 5.3.1) es muy sencillo de construir con ayuda de un papel cuadriculado y puede ser un paso previo a la introducción del gráfico de barras (Figura 5.3.2) y gráfico de sectores (Figura 5.3.3). Mientras en los dos primeros se visualiza mejor el carácter numérico de la variable, la moda, la dispersión y la forma de la distribución, en el gráfico de sectores se visualiza mejor la importancia relativa de cada valor respecto al conjunto de datos. Puede mostrarse también como una aplicación en el tema de las fracciones y servir para introducir o repasar los conceptos de sector circular y amplitud del mismo, así como de aplicación en el tema de la proporcionalidad.

De igual modo se realizaría el estudio del número de caras en las secuencias reales, para finalmente comparar las dos distribuciones y analizar si existen algunas diferencias importantes que indiquen que nuestra intuición respecto a la aleatoriedad nos engaña.

11, 11, 11, 8, 7, 8, 9, 11, 10, 9, 9, 9, 9, 14, 7, 10, 9, 10, 11, 13, 11, 8, 8, 11, 12, 9, 8

5. Compara ahora los gráficos del número de caras en las secuencias reales y simuladas. ¿En qué se parecen? ¿En que se diferencian? ¿Es el valor más frecuente el mismo? ¿Hay el mismo rango de variación de valores? ¿Cuál de las dos variables tiene mayor variabilidad? ¿Piensas que nuestras intuiciones sobre el número de caras que se obtienen al lanzar 20 veces una moneda equilibrada es totalmente correcta? ¿Podrías idear algún tipo de gráfico en que se viesen más claramente las diferencias?

Una característica del número de caras en una secuencia real es que, en general es más variable de lo que nuestra intuición nos sugiere, mientras que los valores medios coinciden, aproximadamente en ambas distribuciones, ya que, en general, somos muy exactos al reflejar la equiprobabilidad de resultados, incluso más exactos de lo debido, puesto que la secuencia simulada tiene menos dispersión que la real.

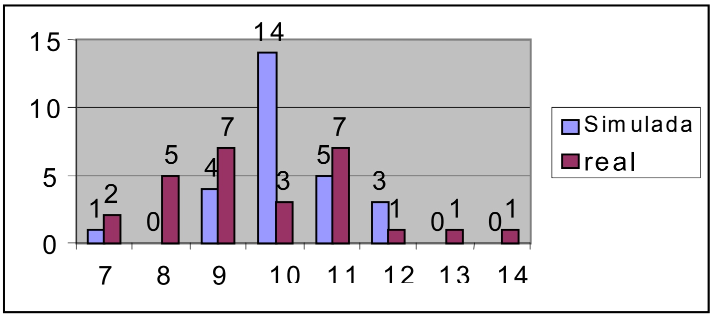
En el ejemplo de nuestra clase, hemos obtenido una distribución *bimodal* (Figura 4) lo cual sugiere la necesidad de introducir aquí las ideas de media ($\overline{x}$=10,14 para las secuencias simuladas $\overline{x}$=9,74 para las reales en nuestro ejemplo).

También surge en esta actividad la idea de dispersión de una forma sencilla. Bien a través del recorrido o del 50 % de casos centrales se observa mayor dispersión en la secuencia real, donde el 50% de los casos centrales se presentan en el intervalo (9-11) y el recorrido es 7, mientras que en la secuencia simulada el 50% de casos centrales se reduce al valor 10 y el recorrido es 5. Es conveniente llevar a los alumnos a realizar gráficos simultáneos para las dos distribuciones, como el presentado en la Figura 5.3.4 o el gráfico de barras adosado (Figura 5.3.5).

Figura 5.3.4. Comparación del número de caras en secuencias reales y simuladas

Secuencia simulada	*N. Caras*	***Secuencia real***
X	*7*	*XX*
	8	*XXXXX*
XXXX	*9*	*XXXXXXX*
XXXXXXXXXXXXXX	*10*	*XXX*
XXXXX	*11*	*XXXXXXX*
XXX	*12*	*X*
	13	*X*
	14	*X*

Figura 5.3.5. Gráficos de barras adosados del número de caras en secuencias reales y simuladas



El número de caras es sólo una de las variables que podemos analizar en una secuencia de resultados aleatorios, en la que aparecen otros muchos modelos probabilísticos. Uno de ellos es la longitud de las rachas que, intuitivamente esperamos que sean cortas. Es bien conocida la *falacia del jugador* por la que esperamos que, tras una corta racha de, por ejemplo caras, la probabilidad de que aparezca una cruz aumente. En otro plano, si un matrimonio tiene ya dos hijos varones, tendrá una gran seguridad en que el siguiente sea una niña, sin darse cuenta que no es demasiado raro (1 caso de cada ocho) los matrimonios con tres varones ni de que, de los matrimonios que ya tienen dos varones, aproximadamente la mitad de los que tengan un nuevo hijo, deben esperar que sea varón, exactamente lo mismo que cuando esperaban a su hijo mayor.

En este proyecto proponemos analizar dos nuevas variables en las secuencias producidas por los alumnos: el número de rachas y la longitud de la racha más larga. Para aclarar el lenguaje llamaremos racha a una secuencia de resultados iguales, de modo que, si después de una cara aparece una cruz (o viceversa) la racha tiene longitud 1.

Para clarificar volvemos al ejemplo inicial y coloreamos las rachas que aparecen. Vemos que en la secuencia simulada, la racha más larga es de longitud 3 (3 caras) y que el número de rachas es 12, mientras que en secuencia real hay una racha de 5 cruces y el número de rachas es 11.

C	C	+	C	+	+	+	C	C	+	C	+	C	+	+	C	C	C	+	+

+	C	+	C	+	+	C	C	+	C	C	C	+	+	+	+	+	C	+	+

Para motivar el estudio de estas variables, el profesor puede preguntar si el resultado obtenido arriba, donde aparecen 5 cruces seguidas, parece razonable. Probablemente algún alumno sugiera que la moneda utilizada no está bien construida y se plantea el estudio de las rachas en las secuencias.

El profesor explicará cómo identificar las rachas y sugerirá a los niños que busquen cuál es la racha más larga en cada una de sus dos secuencias, así como que cuenten el número de rachas, procediendo de nuevo al estudio y comparación de estas variables en las dos secuencias, tal y como se ha hecho con el número de caras y finalizando con una discusión sobre sus diferencias y si nuestras intuiciones sobre las rachas son o no correctas.

El profesor puede usar una hoja de registro como la que reproducimos a continuación (Figura 5.3.6) donde cada niño anota sus resultados. Luego la hoja se fotocopia y se reparte a los chicos. Si hay poco tiempo, la clase puede dividirse en grupos para que cada uno de ellos se encargue de analizar una de las variables y posteriormente, una vez disponibles los gráficos se realiza la discusión conjunta.

Figura 5.3.6. Hoja de recogida de datos de la clase con datos recogidos en una clase

Secuencia simulada			Secuencia real		
N. caras	N. rachas	Racha mayor	N. caras	N. rachas	Racha mayor
10	14	4	11	9	4
12	9	4	11	16	2
11	12	4	11	16	2
10	9	4	8	9	4
11	11	3	7	11	4
9	13	3	8	10	5
10	12	3	9	9	4
11	14	3	11	4	7
9	13	3	10	12	3
10	8	5	9	9	5
10	12	3	9	10	5
10	12	3	9	10	5
7	10	6	9	10	5
10	11	3	14	11	5
10	13	4	7	7	5
10	11	3	10	10	3
10	12	4	9	12	3
12	10	4	10	11	4
11	12	4	11	14	3
10	13	3	13	12	4
9	7	3	11	5	4
10	13	3	8	11	5
10	11	4	8	10	7
9	14	3	11	11	4
10	7	2	12	4	4
12	13	3	9	10	5
11	14	3	8	8	5

5.3.4. Actividades de ampliación

Se pueden plantear a los estudiantes otros problemas que les permitan mostrar sus intuiciones sobre la aleatoriedad, por ejemplo, el ítem siguiente:

6. La probabilidad de que un niño nazca varón es aproximadamente 1/2. ¿Cuál de las siguientes secuencias de sexos es más probable que ocurra en tres nacimientos?

a) MMM; b) VMM; c) las dos son igual de probables.

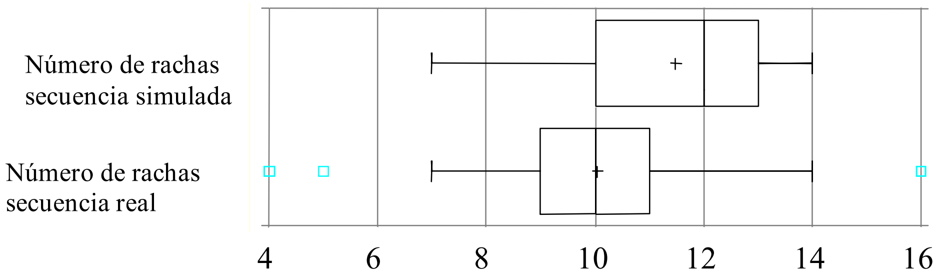
Si el alumno piensa que b) es más probable, pueden organizarse en clase experimentos de simulación con ayuda de tres monedas, donde la cara representa, por ejemplo el varón y la cruz la mujer. También pueden usarse diagramas en árbol para escribir todas las posibilidades en una familia de 3 hijos y enumerar el espacio muestral.

Al introducir la media, se puede hacer ver a los alumnos alguna de sus propiedades sencillas como:

1. La media es un valor comprendido entre los extremos de la distribución;
2. El valor medio es influenciado por los valores de cada uno de los datos;
3. La media no tiene por qué ser igual a uno de los valores de los datos;
4. El valor obtenido de la media de números enteros puede ser un decimal, como en este ejemplo que no tenga sentido en el contexto de los datos;
5. Hay que tener en cuenta los valores nulos en el cálculo de la media.

Con alumnos mayores, puede ampliarse el tipo de gráficos y resúmenes estadísticos utilizables para comparar las distribuciones. Podríamos, por ejemplo, calcular las medianas y cuartiles de las diferentes variables y construir gráficos de cajas paralelos para cada par de variables a comparar. En la Figura 5.5.7 mostramos los gráficos de caja de la longitud de la racha mayor en las dos secuencias, donde se observa como la media, mediana y cuartiles de la variable es menor en la secuencia real. Asimismo el 50 % central de valores está por debajo, lo que nos indica que esperamos demasiadas rachas en una secuencia aleatoria. La menor dispersión nos indica que en esto somos menos variables de lo que ocurre en la realidad.

Figura 5.3.7. Gráficos de cajas paralelos



5.3. 5. Algunas dificultades y errores previsibles

Los alumnos podrían tener dificultades en la realización de los gráficos, construyendo, por ejemplo, unas escalas no homogéneas u omitiendo las escalas o etiquetas que identifiquen claramente el propósito del gráfico. Es importante concienciar a los alumnos de que un gráfico mal construido proporciona una información engañosa. Una actividad complementaria podría ser buscar ejemplos en la prensa de tablas estadísticas o gráficos que presenten errores de construcción o que induzcan a obtener conclusiones equivocadas y posteriormente elaborar una lista de los principales tipos de errores detectados.

Al calcular la media a partir de la tabla de frecuencias, los alumnos podría omitir el ponderar los valores de la variables por las frecuencias, ya que los alumnos tienen con frecuencia dificultad en el cálculo de medias ponderadas. Pueden planteárseles problemas como el siguiente, para hacerles ver la necesidad de ponderación:

Hay 10 personas en un ascensor, 4 mujeres y 6 hombres. El peso medio de las mujeres es de 60 kilos y el de los hombres de 80. ¿Cuál es el peso medio de las 10 personas del ascensor?

Los alumnos tienen a veces dificultades en comprender la idea de mediana; sugerimos el cálculo de la mediana a partir del conjunto de datos ordenados y pasar a los algorítmos de cálculo sólo cuando el alumno ha comprendido bien el significado del concepto.

5.3.6. Análisis del contenido estadístico

En este proyecto podemos identificar, explícita o implícitamente los siguientes contenidos:

1. *Aplicaciones de la estadística:*
- Diseño de un experimento;
- Análisis de datos experimentales; comparación de datos experimentales con patrones teóricos.

2. C*onceptos y propiedades*:
- Aleatoriedad: experimento aleatorio; secuencia de resultados aleatorios, sucesos equiprobables, independencia de ensayos, rachas;
- Variable estadística discreta, frecuencia absoluta; tabla de frecuencias; distribución de frecuencias;
- Posición central, moda, media, mediana;
- Dispersión: rango, casos centrales, 50% de casos centrales.

3. Notaciones y representaciones:
- Palabras como frecuencia, media, mediana, moda, recorrido;
- Símbolos como $\overline{x}$, Me, Mo;
- Tablas de frecuencia; Gráficos de puntos, barras, barras adosados, sectores, cajas.

4.*Técnicas y procedimientos:*
- Recogida y registro de datos experimentales;
- Elaboración de tablas de frecuencia; recuento y cálculo de frecuencias;
- Elaboración de gráficos de puntos, diagramas de barras, diagramas de barras adosados y gráficos de sectores;
- Interpretación de tablas y gráficos; elaboración de conclusiones a partir del análisis de tablas y gráficos;
- Elaboración de argumentos y conclusiones a partir del análisis de datos obtenidos en un experimento;
- Uso de calculadora gráfica, hojas de cálculo o software estadístico.

5. A*ctitudes:*
- Reflexión sobre las propias intuiciones incorrectas en relación a los experimentos aleatorios;

- Valoración de la utilidad de la estadística para analizar datos obtenidos mediante experimentación;
- Valoración de la estética y la claridad en la construcción de tablas y gráficos estadísticos.

5.4. Proyecto 2. ¿Cómo son los alumnos de la clase?

5.4.1. Objetivos

Se trata de elaborar un perfil de los alumnos, identificando el alumno típico y analizando si hay diferencias entre el chico y la chica típicos, así como identificar relaciones entre las variables analizadas. Un objetivo importante es introducir al alumno en las diferentes técnicas de recogida de datos observación, encuesta y medición. Además queremos hacerle ver que al usar datos cualitativos, surge la necesidad de categorización, que siempre supone una simplificación de la realidad y que existen diversos modos de categorizar la misma realidad. Se trata de poner al alumno en la situación de realizar una encuesta, concienciándolo de la importancia de la fiabilidad de los datos, la necesidad y dificultad de la categorización, de la importancia de la claridad en las preguntas y de la serie de pasos que van desde la idea inicial de la investigación hasta la obtención de las conclusiones.

Alumnos: El proyecto podría ser adecuado para alumnos a partir de 14-15 años, ya que hacemos una primera introducción a la idea de asociación y estudio de las tablas de contingencia.

5.4.2. Los datos

Se preparará una lista de las características que queremos incluir en el estudio, analizando las diferentes formas en que podrían obtenerse los datos:

- Por simple observación: como el sexo, color de pelo y ojos, si usa o no gafas;
- Se requiere una medición: como el peso, talla o longitud de brazos extendidos;
- Habría que preguntar a los alumnos; es decir realizar una pequeña encuesta: número de hermanos, cómo viene al instituto; cuánto deporte practica, etc.

Los datos serán recogidos por los propios alumnos, mediante las diversas técnicas señaladas. Se requerirá un metro y una báscula, para tomar datos de los alumnos con un mismo instrumento.

5.4.3. Preguntas, actividades y gestión de la clase

Una vez planteado el proyecto, la actividad comienza con la recogida, codificación y registro de los datos. Algunas características a incluir, y las preguntas relacionadas con la obtención de los datos, se recogen a continuación.

1.Tomemos datos del sexo de cada alumno. ¿Qué tipo de variable es el sexo? ¿tendría sentido calcular la media de esta variable? ¿ y la moda? Para ello es interesante que nos pongamos *de acuerdo, sobre como vamos a codificar los chicos y chicas. De lo contrario, alguno de vosotros podría usar "chico/chica", otros "varón/ mujer" o "hombre/mujer", "V/M", etc. Un sistema posible de codificar los datos sería*

1= "chico"; 2= chica".

Es importante el hecho de que la codificación es un convenio, pero que debemos llegar a un acuerdo y describir el sistema empleado para que otros puedan comprender nuestros datos.

2. ¿Cómo se distribuye el sexo de los alumnos en esta clase? Prepara una tabla de frecuencias y un gráfico que describa la distribución. ¿Es el alumno típicoun chico o una chica?

Los alumnos prepararán una tabla de frecuencias similar a la 5.4.1 y elaborarán alguno de los gráficos que ya conocen, como el diagrama de barras o de sectores. El alumno típico de la clase es una chica, puesto que la moda es ser una chica (valor más frecuente).

Tabla 5.4.1. Distribución de frecuencia del sexo de los alumnos

Sexo	Frecuencia absoluta	Frecuencia relativa	Porcentaje
Chicos	23	0,3833	38.3
Chicas	37	0,6167	61.7
Total	60	1	100

3. Tratemos, ahora de recoger datos sobre la práctica de deporte. ¿Cómo podemos recoger estos datos? ¿Cómo los podríamos codificar?

Los alumnos se encuentran ahora con el problema de que la práctica de deporte no es una variable directamente observable, aunque cada uno de los alumnos conoce si practica o no deporte y la frecuencia con que lo practica. Surgirá la discusión de como codificar esta variable: una posibilidad sería preguntar por el número de días a la semana que se practica deporte (con lo cual tendríamos una variable cuantitativa discreta con valores 0 a 7). Pero es posible que no todos los alumnos sean sistemáticos en la práctica de deporte: unas semanas practiquen 3 días y otras ninguno, quizás dependiendo de si es época de exámenes o no.

Nosotros, al realizar este proyecto en clase, decidimos codificar simplemente con tres valores: 1 (poco, sólo de vez en cuando), 2 (con frecuencia, alguna vez cada semana), 3 (sistemáticamente, por ejemplo, 2 o más días en semana). Con este convenio (u otro como una escala 0-10) obtenemos una escala ordinal, porque lo que un alumno considera sistemático puede no coincidir con la opinión de otro y porque 2 no representa el doble que 1.

Una vez llegados a una decisión se recogerían los datos, elaboraría una tabla de frecuencias y se buscaría el valor típico. La tabla de frecuencias esta vez puede tener frecuencias acumuladas y el valor típico preferible sería la mediana, porque es más informativa que la moda y la media no sería muy precisa al ser la escala ordinal. La mediana en los alumnos de nuestra clase fue 2, por lo que el alumno típico es una chica que practica deporte con frecuencia.

4. ¿Cómo clasificamos a los alumnos según el color de pelo? ¿Y según el color de ojo?

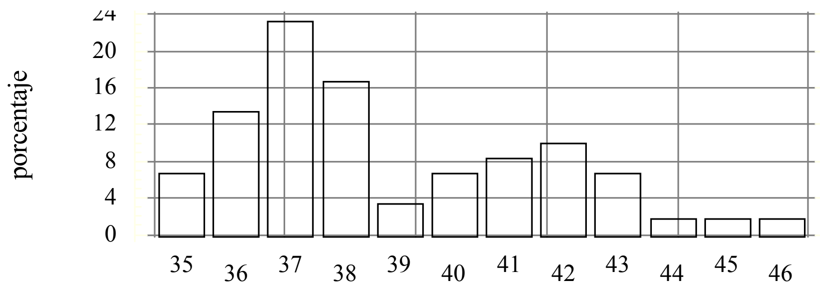
Esta vez se trata de variables cualitativas que pueden observarse directamente, sin necesidad de preguntar. Sin embargo, se plantea el problema de la clasificación. Para los ojos, por ejemplo, podríamos considerar ojos verdes, azules, grises, castaños y negros, e incluso diferenciar entre castaño y dorados. Incluso así, para algunos alumnos podría ser difícil decidir si sus ojos son azules o verdes y al final habría que tomar una decisión sobre como categorizar al alumno.

Estas variables dan lugar a reflexiones interesantes sobre que al categorizar siempre simplificamos la realidad y que los mismos datos podrían categorizarse en forma diferente. En estadística no hay una única solución a cada problema y tan importante o más que los cálculos son las decisiones que se toman sobre cómo recoger y categorizar los datos. En nuestro caso nos decidimos por clasificar simplemente a los alumnos por ojos claros/ oscuros y pelo claro/ oscuro. El 61 % (37 alumnos tenía los ojos oscuros) y el 58% tenía los ojos oscuros, luego el alumno típico es una chica de pelo y ojos oscuros que practica deporte en forma moderada.

5. ¿Cuál es el número de calzado típico?

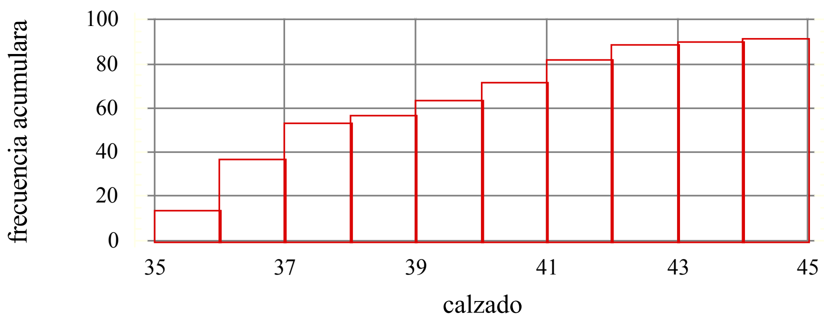
Los alumnos analizarán la distribución del número de calzado (Figura 5.4.1.), cuya moda es el 37 (el calzado más frecuente). Puesto que la variable se mide ahora en escala de razón, se podría plantear el cálculo de la media, cuyo valor es igual *38.8.* Se puede plantear al alumno la pregunta de ¿Qué significa que el número promedio de calzado es 38.8, cuando tal número realmente no existe y también, por qué hay tanta diferencia entre la media y la moda en este caso.

Figura 5.4.1. Distribución del número de calzado



Observamos que unos pocos alumnos tienen un pie muy grande y hacen subir artificialmente el valor de la media, que es muy sensible a los casos atípicos. La mediana es un estadístico más robusto, y sugerimos a los alumnos calcular la mediana a partir del diagrama de frecuencias acumuladas (Figura 5.4.2.)

Figura 5.4.2. Diagrama de frecuencias acumuladas



Vemos de este diagrama que el 50 % de los alumnos tienen un número de calzado igual o menor a 37 y el resto igual o mayor (puesto que en 37 la frecuencia acumulada salta del 43% al 60%. El valor mediano del número de calzado es 37 y coincide con la moda. Tenemos que añadir, como característica típica del alumno típico el número de calzado 37.

6. ¿Cuáles son el peso, la talla y la longitud típica de brazos?

Al trabajar con variables continuas o variables en que el número de valores diferentes es grande, se hace necesaria la agrupación. Como paso previo a la construcción de una tabla de frecuencias o un gráfico, se puede pedir a los alumnos que construyan un dagrama de tallo y hojas (Figura 5.4.3). En este diagrama se visualiza la frecuencia en intervalos de amplitud 10 o 5 y se conservan los valores numéricos de los datos. Es sencillo de construir con una hoja de papel cuadriculado.

Figura 5.4.3. Gráfico de tallo y hojas. Altura de los alumnos

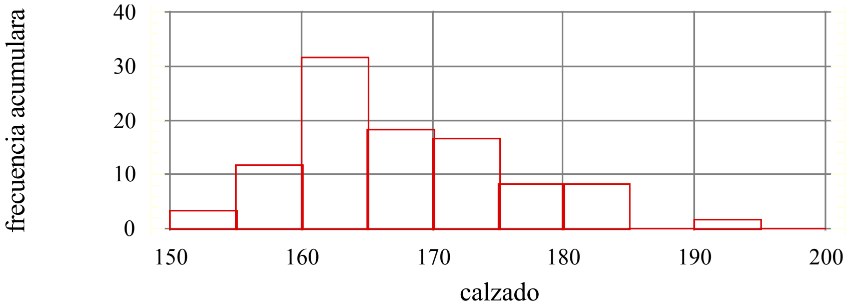
15|55689
16|0000111122223333444
16|555566777899
17|000222344
17|55556889
18|0344
18|55
19|1

A partir de este gráfico los alumnos prepararán tablas de frecuencia para las variables, similares a la tabla 5.5.1. Un punto importante es que no hay una regla fija respecto a la elección de los intervalos de clase y el número y límite de intervalos determinará la forma del histograma (Figuras 5.4.4 y 5.4.5.). Mientras que con 9 intervalos se visualiza un valor atípico, este aparece oculto con 7 intervalos. Un criterio a seguir es que los extremos de los intervalos sean números enteros y también que el número de intervalos sea, aproximadamente, la raiz cuadrada del número de datos (aunque no exacta; en nuestro caso, con 60 datos, un número razonable de intervalos sería 7, pero hemos tomado nueve, para que los extremos sean múltiplos de 5.

Tabla 5.4.1. Distribución de la talla (cm.) de los alumnos

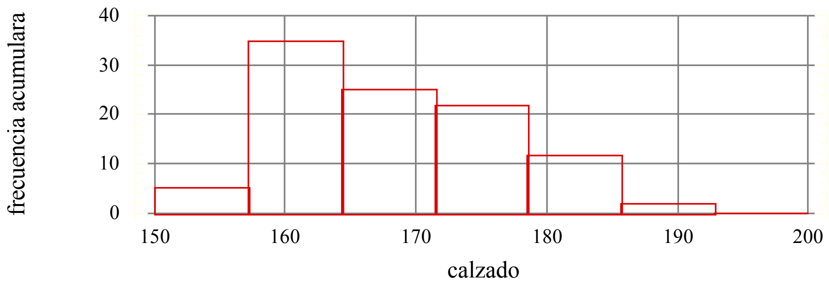
Intervalo	Marca de clase	Frecuencia aboluta	Frecuencia relativa	Frecuencia acumulada	Acumulada relativa
150-155	152.5	2	0.0333	2	0.033
155-160	157.5	7	0.2267	9	0-1500
160-165	162.5	19	0.3167	28	0.4667
165-170	167.5	11	0.1833	39	0.6500
170-175	172.5	10	0.1677	49	0.8167
175-180	177.5	5	0.0833	54	0.9000
180-185	182.5	5	0.0833	59	0.9833
185-190	187.5	0	0.0000	59	0.9833
190-195	192.5	1	0.0167	60	1.0000

Figura 5.4.4. Distribución de la talla de los alumnos (9 intervalos)



La agrupación en intervalos introduce una reducción en los datos; por ejemplo, si calculamos la media y otros estadísticos directamente de la tabla, los valores obtenidos son aproximados, por eso también el número de intervalos afectará a estos estadísticos. Este problema no se presenta al calcular directamente los datos con la calculadora o el ordenador.

Figura 5.4.5. Distribución de la talla de los alumnos (7 intervalos)



Tomando la mediana como medida de posición central llegamos a la conclusión de que el alumno típico es una chica de pelo y ojos oscuros que practica deporte moderadamente, calza el 37, mide 166,5 cm., pesa 62 kilos y la longitud de sus brazos es 165 cm. También llegaremos a la conclusión de que esta chica realmente no existe: Ninguna de las alumnas corresponde exactamente a esta descripción !

Este proyecto puede ser más o menos complejo, en función del número y tipo de variables incluídas. En la Tabla 5.5.2. incluimos datos obtenidos sobre características físicas y práctica de deporte en una clase de 60 estudiantes del curso preuniversitario. Dependiendo de la edad de los alumnos y el tiempo disponible estas variables podrían reducirse o ampliarse. Por ejemplo, el estudio podría llevarse a cabo sólo con las variables cualitativas (sexo, deporte, ojos, pelo, y número de calzado) o añadir otras como perímetro de cintura, anchura de hombros, etc. Para relacionar posteriormente las variables será bueno elaborar una hoja de recogida de datos como la que se muestra en la Tabla 5.5.2.

7. ¿Cuáles son las principales diferencias entre sexos?¿Cómo sería el chico/ chica típico?

En este fichero hemos mezclado datos de dos poblaciones diferenciadas en cuanto a sus características físicas. Ello podría explicar el hecho de que no hubiésemos encontrado en la clase un representante del alumno típico. El trabajo continua analizando las características que diferencian al alumno/ alumna típicos.

Tabla 5.5.2. Datos obtenidos en el estudio sobre el alumno típico en 60 estudiantes

Sexo	Deporte	Ojos	Pelo	N. calzado	Peso (kgs.)	Talla (cm.)	L. brazos (cm.)
M	2	C	C	37	59	161	160
V	1	O	O	41	62	178	181
M	2	O	O	36	50	159	153
V	2	O	O	42	69	176	179
V	2	O	O	43	74	175	179
M	3	C	C	37	62	169	165
M	2	O	O	36	56	162	158
M	2	O	O	37	58	162	163
M	1	O	O	38	52	170	171
V	1	O	O	42	68	170	172
V	3	O	O	43	72	184	185
V	2	C	C	42	74	180	182
V	2	C	C	41	66	175	177
M	2	O	O	38	60	170	168
M	1	C	C	38	60	165	161
M	3	O	O	36	55	163	160
M	2	O	O	37	60	167	165
M	2	C	O	37	50	167	165
M	2	C	C	35	52	160	157
M	1	O	O	37	53	164	160
M	2	O	C	38	58	163	166
M	2	O	O	40	74	175	178
M	2	O	O	39	63	173	180
M	2	O	C	38	60	161	164
M	2	O	O	37	53	162	162
V	3	C	C	41	82	174	180
V	2	O	C	42	68	178	180
M	1	0	C	37	64	172	175
M	2	O	O	40	65	165	165
M	1	O	O	37	46	160	158
M	2	C	C	38	58	164	166
M	3	C	O	46	86	191	180
M	1	O	O	44	70	161	185
M	1	O	C	40	64	166	171
M	3	O	O	38	64	166	155
M	1	O	O	35	70	156	152
M	2	C	C	37	51	165	160
M	2	C	O	38	62	167	159
M	2	O	C	37	58	160	160
V	3	C	C	43	71	185	187
V	3	C	C	42	68	175	172
V	3	O	C	45	74	183	178
M	3	O	O	37	55	160	154
V	3	O	O	42	68	185	185
M	1	O	O	37	57	161	155
M	3	C	O	38	57	169	164
M	1	O	O	36	68	158	150
V	2	O	O	41	69	172	172
M	3	C	O	37	50	155	155
M	1	O	O	38	58	163	162
M	1	C	O	39	66	168	168
M	2	C	C	36	50	163	161
V	2	O	O	43	81	184	188
M	1	C	C	36	60	165	160
M	2	C	C	35	50	155	155
V	2	C	C	35	65	179	171
V	2	O	O	36	65	164	158
M	2	C	C	40	62	174	179
M	2	O	C	36	58	162	160
M	2	C	C	41	63	172	171

Para ello los alumnos pueden comparar las distribuciones de las variables en las dos muestras (muestras independientes). Por ejemplo, en las figuras 5.4.6, 5.4.7, 5.4.8. y 5.4.9 incluímos algunas de las gráficas que se podrían usar para comparar las variables en chicos y chicas. Mientras el chico típico calzaría el 42, pesaría 69 kilos, mediría 177 cm., y tendría una longitud de brazos de 179.5 cm.. la chica típica calzaría el 38, pesaría 58 kilos, mediría 163 cm y tendría una longitud de brazos de 161 cm.

Los gráficos de caja visualizan las medias y medianas, cuartiles, recorrido intercuartílico y, en caso de haberlos, los valores atípicos. Observamos que estos valores son siempre menores en las chicas, así como la dispersión de los datos. Los histogramas, por su lado resaltan las modas y la frecuencia de casos en cada intervalo.

Los gráficos de cuantiles ponen de relieve, que, para cualquier rango de percentil (por ejemplo el 30 o 60 % la altura de las chicas es siempre menor que la de los chicos. Por último las curvas empíricas de distribución, obtenidas de la suavización del polígono de frecuencias, indican que las distribuciones son más o menos simétricas, concentradas en el centro del rango de valores. Si el número de datos hubiese sido mayor, observaríamos la forma característica de la distribución normal.

Figura 5.4.6. Gráficos de cajas paralelos. Número de calzado de chicos y chicas

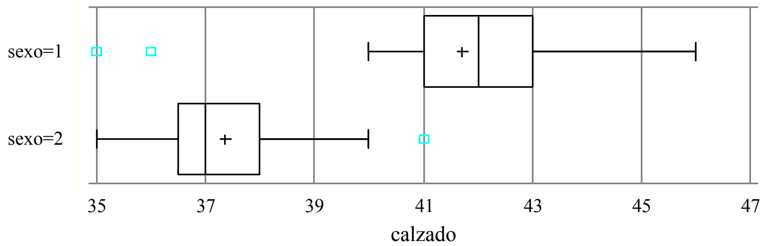


Figura 5.4.7. Histogramas. Longitud de brazos de chicos y chicas

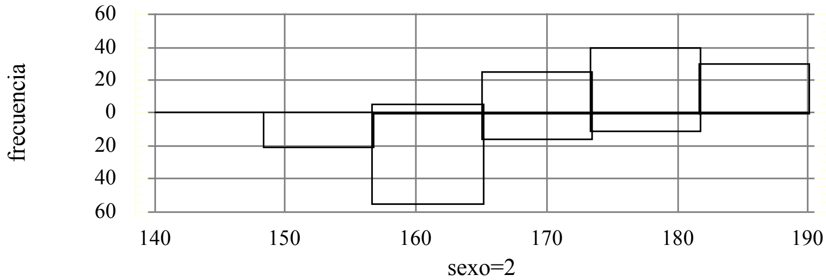


Figura 5.4.8. Gráfico de cuantiles. Alturas de chicos y chicas

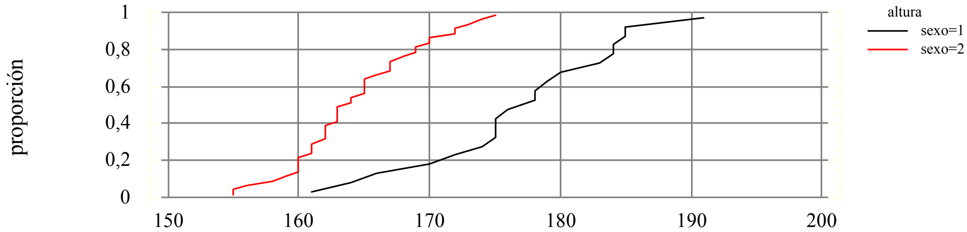
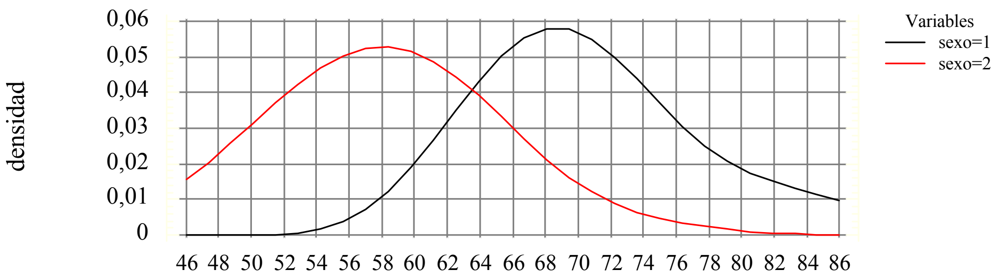


Figura 5.4.9. Curvas de densidad empíricas. Alturas de chicos y chicas



5.4.4. Actividades de ampliación

En este proyecto puede introducirse en forma muy intuitiva el análisis de tablas de contingencia y de la asociación entre las variables de la misma, ya que tratamos de determinar si existen también diferencias en el color de pelo, ojos y práctica de deporte entre chicos y chicas. La idea de asociación puede introducirse usando el mismo tipo de tarea empleado en las investigaciones de Piaget e Inhelder: el estudio de la asociación entre el color de ojos y pelo. Para ello, la primera actividad será pedir a los alumnos que clasifiquen los datos respecto al color de pelo y ojos, obteniendo la tabla 5.5.3.

Tabla 5.5.3. Clasificación cruzada de alumnos según color de pelo y ojos

	Ojos claros	Ojos oscuros	Total
Pelo claro	16	11	27
Pelo oscuro	5	28	33
Total	21	39	60

A partir de ella se les puede plantear preguntas tales como:

8. ¿Cuántos alumnos tienen pelo claro? ¿Ojos oscuros? ¿Cuál es la proporción de alumnos con ojos claros entre los que tienen pelo claro? ¿ Y de alumnos con pelo claro entre los que tienen ojos claros?¿Hay la misma proporción de alumnos con ojos claros si se tiene el pelo claro que si se tien oscuro?¿Piensas que hay relación entre el color de pelo y ojos?

Tabla 5.5.4. Distribución condicional de color de ojos según color de pelo

	Ojos claros	Ojos oscuros	Total
Pelo claro	16 59 %	11 41%	27
Pelo oscuro	5 15%	28 85%	33
Total	21 35%	39 65%	60

En la tabla 5.5.4 podemos observar que mientras el 35% de los alumnos tiene ojos claros y el 65 % azules, estas proporciones son el 58% y 41% en alumnos rubios y el 15% y 85% en morenos. Dicho de otro modo hay doble número de alumnos con ojos negros si se es moreno y casi cuatro veces más alumnos con ojos azules si se es rubio.

Tabla 5.5.5. Distribución condicional de color de pelo según color de ojos

	Ojos claros	Ojos oscuros	Total
Pelo claro	16	11	27
	76%	28%	45%
Pelo oscuro	5	28	33
	24%	72%	55%
Total	21	39	60

En la tabla 5.5.5 podemos observar que, mientras la proporción de alumnos con pelo claro y oscuro es casi la misma (45% y 55%), 3 de cada 4 alumnos con ojos claros son rubios y 3 de cada 4 alumnos con ojos oscuros es moreno.

5.4. 5. Algunas dificultades y errores previsibles

Los alumnos pueden tener dificultad al construir las tablas de datos agrupados, si no se especifica con claridad el convenio de agrupación. Por ejemplo, se puede tener duda si el valor 150 se incluye en el intervalo (140-150) o (150-160). Una costumbre es usar intervalos semiabiertos por la derecha, pero este tipo de intervalo puede ser poco familiar a los alumnos. El profesor puede discutir con los alumnos el convenio de agrupación y llegar con ellos a un acuerdo acerca de la forma de construir los intervalos.

El cálculo de promedios y medidas de dispersión es bastante complicado a partir de datos agrupados. Los alumnos pueden olvidar la necesidad de ponderar por la frecuencia en cada intervalo y los cálculos son engorrosos. Sugerimos obviar este punto recurriendo a las calculadoras que dan directamente estos estadísticos, sin más que introducir los datos originales.

El análisis de una tabla de contingencia es un punto difícil porque de una sola frecuencia absoluta (número de alumnos con pelo y ojos claros) se pueden deducir diferentes frecuencias relativas:

- Frecuencia relativa doble respecto al total de datos: frecuencia relativa de alumnos con pelo y ojos claros en la muestra;
- Frecuencia relativa condicional respecto a su fila: frecuencia de alumnos con ojos claros si se tiene el pelo claro;
- Frecuencia relativa condicional respecto a su columna: frecuencia de alumnos con pelo claro si se tiene el pelo claro.

Además aparecen las dos frecuencias relativas marginales (frecuencias relativas de alumnos con pelo claro y ojos claros) y las relaciones entre todas estas frecuencias. El problema de la búsqueda de asociación en una tabla de contingencia 2x2 se reduce, a este nivel intuitivo, a un problema de comparación de proporciones:

- Comparar las frecuencias condicionales por filas: proporción de color de ojos según color de pelo;

- Comparar las frecuencias condicionales por columnas: proporción de color de pelo, según color de ojos;
- Comparar alguna de estas con sus correspondientes frecuencias marginales.

Sin embargo, en el estudio de Piaget e Inhelder con chicos a partir de 13 14 años y en trabajos posteriores con adultos se ha visto que la interpretación de las tablas de contingencia no es intuitiva y algunos alumnos tienden a usar frecuencias absolutas o usar sólo una parte de la información en la tabla para resolver este tipo de problemas.

5.3.6. Análisis del contenido estadístico

En este proyecto podemos identificar, explícita o implícitamente los siguientes contenidos:

1. *Aplicaciones de la estadística:*
- Diseño de un experimento;
- Análisis de datos experimentales; comparación de datos experimentales con patrones teóricos.

2. C*onceptos y propiedades*:
- Datos: codificación de datos: dificultad de la categorización;
- Variable estadística: Variable nominal, discreta y continua, frecuencia absoluta, relativa y acumulada; tabla de frecuencias; distribución de frecuencias, agrupación; intervalos, extremos y marcas de clase;
- Posición central: moda, media, mediana, percentiles, rangos de percentiles;
- Dispersión: rango, máximo, mínimo, cuartiles; recorrido intercuartílico;
- Asociación: tablas de contingencia; frecuencias dobles, marginales y condicionadas; asociación en tablas de contingencia.

3. Notaciones y representaciones:
- Palabras como extremos de clase, marcas de clase, cuartiles, recorrido intercuartílico;
- Símbolos correspondientes a las frecuencias acumuladas, marginales y condicionadas;
- Tablas de frecuencia; gráficos de tallo y hoja., caja, histograma, curvas empírica de distribución, gráfico de cuantiles, tablas de contingencia.

4.*Técnicas y procedimientos:*
- Elaboaración de un cuestionario; observación, medida, codificación de datos;
- Elaboración de tablas de frecuencia simples y cruzadas; recuento y cálculo de frecuencias agrupadas y frecuencias marginales y condicionadas en tablas de contingencia;
- Elaboración de gráficos de tallo y hoja., caja, histograma, curvas empírica de distribución, gráfico de cuantiles;
- Interpretación de tablas y gráficos; elaboración de conclusiones a partir del

análisis de tablas y gráficos; estudio de asociación en tablas de contingencia;

- Elaboración de argumentos y conclusiones a partir del análisis de datos obtenidos en observación, encuesta y medida
- Uso de calculadora gráfica, hojas de cálculo o software estadístico.

5. A*ctitudes:*

- Reflexión sobre la dificultad de codificación y cómo ésta introduce siempre una simplificación en la realidad;
- Valoración de la utilidad de la estadística para analizar datos obtenidos mediante encuesta observación y medida;
- Valoración de la utilidad de la estadística para identificar relaciones de asociación entre variables;
- Valoración de la estética y la claridad en la construcción de tablas y gráficos estadísticos;
- Reflexión sobre las tendencias y dispersiones en los datos; sobre el excesivo énfasis en los prototipos y el hecho de que éstos con frecuencia son modelos que no se dan en la realidad.

5.5. Proyecto 3. Análisis demográfico

5.6. 1. Objetivos

Se trata de analizar una serie de variables demográficas, entender su utilidad y el motivo por el cual son recogidas, analizar las relaciones entre las diferentes variables y estudiar las diferencias en sus distribuciones entre países según su nivel de desarrollo. Un objetivo importante es mostrar la utilidad de la estadística en el estudio de interelaciones entre variables en estudios transversales. Los alumnos, alternativamente, podrían tomar otros ficheros de datos de Internet o de anuarios estadísticos y desarrollar proyectos sobre otros temas en diferentes áreas de aplicación. El análisis es descriptivo, con finalidad exploratoria y no se plantean problemas de inferencia.

Rouncenfield (1995) presenta este proyecto sobre el cual los alumnos pueden trabajar en grupos, comparando las variables en los diferentes grupos de países y formulando por si mismos preguntas de su interés. Mostraremos algunos ejemplos de las posibles actividades a desarrollar en un curso de análisis exploratorio de datos.

Alumnos: Este proyecto podría ser desarrollado con alumnos de los últimos cursos de secundaria (16-17 años) o en el curso introductorio en la Universidad. En este último caso, se podrían plantear preguntas de tipo inferencial, lo que requeriría el uso de procedimientos estadísticos más avanzados.

5.6.2. Los datos

La actividad se desarrolla en torno a un proyecto a partir de un fichero que contiene datos de 97 países y que ha sido adaptado del preparado por Rouncenfield (1995), quien usó como fuentes Day (1992) y U.N.E.S.C.O. (1990). Este fichero ha sido tomado de Internet, del servidor de Journal of Statistical Education (http://www.amstat.org/publications/jse/). Contiene las siguientes variables, que se refieren a 1990:

- *Tasa de natalidad*: Niños nacidos vivos en el año por cada 1000 habitantes;
- *Tasa de mortalidad*: Número de muertes en el año por cada 1000 habitantes; *Mortalidad infantil*: Número de muertes en el por cada 1000 niños de menos de 1 año;
- *Esperanza de vida* al nacer para hombres y mujeres;
- *PNB*. Producto Nacional Bruto per cápita en dólares (USA);
- *Grupo*: Clasificación de países en función de la zona geográfica y situación económica, en las siguientes categorías: 1 = Europa Oriental, 2 = Iberoamérica, 3 = Europa Occidental, Norte América, Japón, Australia, Nueva Zelanda, 4 = Oriente Medio, 5 = Asia, 6 = Africa.

Hemos añadido el número de habitantes en 1990 en miles de personas (*Población*), tomado del anuario publicado por el periódico español "El País". En la tabla 5.6.1 listamos los datos del proyecto.

Tabla 5.6.1. Fichero de datos del proyecto "Análisis demográfico"

País	Grupo	Tasa natalidad	Tasa mortalidad	Mortalidad infantil	Esperanza vida hombre	Esperanza vida mujer	PNB	Población (miles)
Afganistán	5	40.4	18.7	181.6	41.0	42.0	168	16000
Albania	1	24.7	5.7	30.8	69.6	75.5	600	3204
Alemania (Oeste)	3	11.4	11.2	7.4	71.8	78.4	22320	16691
Alemania Este	1	12.0	12.4	7.6	69.8	75.9	.	61337
Algeria	6	35.5	8.3	74.0	61.6	63.3	2060	24453
Angola	6	47.2	20.2	137.0	42.9	46.1	610	9694
Arabia Saudí	4	42.1	7.6	71.0	61.7	65.2	7050	13562
Argentina	2	20.7	8.4	25.7	65.5	72.7	2370	31883
Austria	3	14.9	7.4	8.0	73.3	79.6	17000	7598
Bahrein	4	28.4	3.8	16.0	66.8	69.4	6340	459
Bangladesh	5	42.2	15.5	119.0	56.9	56.0	210	111590
Bélgica	3	12.0	10.6	7.9	70.0	76.8	15540	9886
Bielorusia	1	15.2	9.	13.1	66.4	75.9	1880	.
Bolivia	2	46.6	18.0	111.0	51.0	55.4	630	7110
Botswana	6	48.5	11.6	67.0	52.3	59.7	2040	1217
Brasil	2	28.6	7.9	63.0	62.3	67.6	2680	147294
Bulgaria	1	12.5	11.9	14.4	68.3	74.7	2250	9001
Camboya	5	41.4	16.6	130.0	47.0	49.9	.	8250
Canadá	3	14.5	7.3	7.2	73.0	79.8	20470	26302
Colombia	2	27.4	6.1	40.0	63.4	69.2	1260	32335
Congo	6	46.1	14.6	73.0	50.1	55.3	1010	2208
Corea (Norte)	.5	23.5	18.1	25.0	66.2	72.7	400	21143
Checoslovaquia	1	13.4	11.7	11.3	71.8	77.7	2980	15641
Chile	2	23.4	5.8	17.1	68.1	75.1	1940	12980
China	5	21.2	6.7	32.0	68.0	70.9	380	1105067
Dinamarca	3	12.4	11.9	7.5	71.8	77.7	22080	5132
Ecuador	2	32.9	7.4	63.0	63.4	67.6	980	10329
Egipto	6	38.8	9.5	49.4	57.8	60.3	600	51390
Emiratos Arabes	4	22.8	3.8	26.0	68.6	72.9	19860	1544
España	3	10.7	8.2	8.1	72.5	78.6	11020	39161
Etiopía	6	48.6	20.7	137.0	42.4	45.6	120	48861
Filipinas	5	33.2	7.7	45.0	62.5	66.1	730	61224
Finlandia	3	13.2	10.1	5.8	70.7	78.7	26040	4974
Francia	3	13.6	9.4	7.4	72.3	80.5	19490	56119
Gabón	6	39.4	16.8	103.0	49.9	53.2	390	1105
Gambia	6	47.4	21.4	143.0	41.4	44.6	260	848
Ghana	6	44.4	13.1	90.0	52.2	55.8	390	14425

Grecia	3	10.1	9.2	11.0	65.4	74.0	5990	10039
Guayana	2	28.3	7.3	56.0	60.4	66.1	330	95
Holanda	3	13.2	8.6	7.10	73.3	79.9	17320	14828
Hong_Kong	5	11.7	4.9	6.10	74.3	80.1	14210	5735
Hungría	1	11.6	13.4	14.8	65.4	73.8	2780	10587
India	5	30.5	10.2	91.0	52.5	52.1	350	832535
Indonesia	5	28.6	9.4	75.0	58.5	62.0	570	178211
Irán	4	42.5	11.5	108.1	55.8	55.0	2490	50204
Iraq	4	42.6	7.8	69.0	63.0	64.8	3020	18271
Irlanda	3	15.1	9.1	7.5	71.0	76.7	9550	3537
Israel	4	22.3	6.3	9.7	73.9	77.4	10920	4525
Italia	3	9.7	9.1	8.8	72.0	78.6	16830	57537
Japón	3	9.9	6.7	4.0	75.9	81.8	25430	123045
Jordania	4	38.9	6.4	44.0	64.2	67.8	1240	4041
Kenya	6	47.0	11.3	72.0	56.5	60.5	370	23277
Kuwait	4	26.8	2.	15.6	71.2	75.4	16150	2020
Líbano	4	31.7	8.7	48.0	63.1	67.0	.	2900
Libia	6	44.0	9.4	82.0	59.1	62.5	5310	4395
Malasia	5	31.6	5.6	24.0	67.5	71.6	2320	17340
Malawi	6	48.3	25.0	130.0	38.1	41.2	200	8230
Marruecos	6	35.5	9.8	82.0	59.1	62.5	960	24567
México	2	29.0	23.2	43.0	62.1	66.0	2490	85440
Mongolia	5	36.1	8.8	68.0	60.0	62.5	110	2128
Mozambique	6	45.0	18.5	141.0	44.9	48.1	80	15357
Namibia	6	44.0	12.1	135.0	55.0	57.5	1030	1300
Nepal	5	39.6	14.8	128.0	50.9	48.1	170	18431
Nigeria	6	48.5	15.6	105.0	48.8	52.2	360	113665
Noruega	3	14.3	10.7	7.8	67.2	75.7	23120	4215
Omán	4	45.6	7.8	40.0	62.2	65.8	5220	1486
Pakistán	5	30.3	8.1	107.7	59.0	59.2	380	109950
Paraguay	2	34.8	6.6	42.0	64.4	68.5	1110	4161
Perú	2	32.9	8.3	109.9	56.8	66.5	1160	21142
Polonia	1	14.3	10.2	16.0	67.2	75.7	1690	38061
Portugal	3	11.9	9.5	13.1	66.5	72.4	7600	10333
Rumania	1	13.6	10.7	26.9	66.5	72.4	1640	23148
Sierra Leona	6	48.2	23.4	154.0	39.4	42.6	240	4040
Singapur	5	17.8	5.2	7.5	68.7	74.0	11160	2664
Somalia	6	50.1	20.2	132.0	43.4	46.6	120	6089
Sri_Lanka	5	21.3	6.2	19.4	67.8	71.7	470	16779
Sudáfrica	6	32.1	9.9	72.0	57.5	63.5	2530	34925
Sudán	6	44.6	15.8	108.0	48.6	51.0	480	24423
Suecia	3	14.5	11.1	5.6	74.2	80.0	23660	8485
Suiza	3	12.5	9.5	7.1	73.9	80.0	34064	6541
Swazilandia	6	46.8	12.5	118.0	42.9	49.5	810	761
Tailandia	5	22.3	7.7	28.0	63.8	68.9	1420	55200
Tanzania	6	50.5	14.0	106.0	51.3	54.7	110	25627
Túnez	6	31.1	7.3	52.0	64.9	66.4	1440	7988
Turquía	4	29.2	8.4	76.0	62.5	65.8	1630	54899
U.K.	3	13.6	11.5	8.4	72.2	77.9	16100	57270
U.S.A.	3	16.7	8.1	9.1	71.5	78.3	21790	248243
Ucrania	1	13.4	11.6	13.0	66.4	74.8	1320	.
Uganda	6	52.2	15.6	103.0	49.9	52.7	220	16722
Uruguay	2	18.0	9.6	21.9	68.4	74.9	2560	3067
URSS	1	17.7	10.0	23.0	64.6	74.0	2242	287664
Venezuela	2	27.5	4.4	23.3	66.7	72.8	2560	19244
Vietnam	5	31.8	9.5	64.0	63.7	67.9	.	65758
Yugoslavia	1	14.0	9.0	20.2	68.6	74.5	.	23707
Zaire	6	45.6	14.2	83.0	50.3	53.7	220	34442
Zambia	6	51.1	13.7	80.0	50.4	52.5	420	7837
Zimbabwe	6	41.7	10.3	66.0	56.5	60.1	640	9567

5.6. 3. Preguntas, actividades y gestión de la clase

1. ¿Podrías explicar qué significa cada una de las variables del fichero? ¿Quién y cómo las calcula? ¿Cómo se recogen los datos? ¿Habría otro modo de calcular el índice de natalidad? ¿Para qué sirven? ¿Podrías encontrar alguna noticia en la prensa relacionada con estas *variables?*

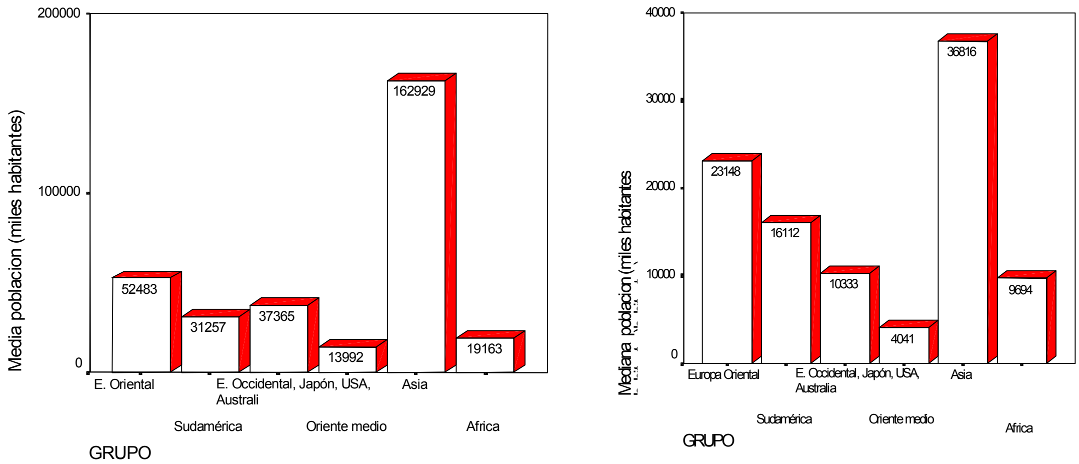
La actividad inicial consiste en discutir el significado de las variables de este fichero y analizar cómo se han calculado las diferentes tasas: natalidad, mortalidad, esperanza de vida, PNB. Los alumnos podrían investigar qué otros indicadores alternativos se emplean para obtener un indicador demográfico o económico de la riqueza de un país. El profesor podría pedir a los alumnos que busquen artículos en la prensa en que se hable de alguno de estos indicadores y que expliquen con sus propias palabras la utilidad que pueden tener y que averigüen quien y como los calcula.

Un punto muy importante es la discusión de las variables y el problema de la medición. Es preciso concienciar a los alumnos de la dificultad que reviste el proceso de categorización o de medición, porque la realidad es siempre más compleja que nuestros métodos para estudiarla. La toma de conciencia sobre la complejidad del proceso de elaboración de las estadísticas demográficas o económicas es un paso importante para valorar el trabajo del estadístico y fomentar la cooperación en censos y encuestas.

En este fichero se ha usado un código para agrupar los países en función de la zona geográfica y desarrollo económico Los alumnos podrían sugerir otras variables de clasificación de los países o añadir otras variables o países al fichero. El trabajo con un fichero completo, en lugar de centrarse en variables aisladas supone el inicio de una filosofía multivariante donde cada variable cobra su importancia o bien es explicada en función del resto y donde el alumno puede tratar de comprobar sus conjeturas con la incorporación de nuevas variables al estudio.

2. *En la Figura 5.6.1 hemos representado el número promedio de habitantes en cada país, según grupo, usando dos promedios diferentes: media y mediana. ¿Por qué las dos gráficas son tan diferentes? ¿Elegirías la media o la mediana para representar el número típico de habitantes en los países, según la zona geográfica.*

Figura 5.6. 1: Mediana y media del número de habitantes en los diferentes grupos de países



La elaboración de una tabla o un gráfico ya supone una primera reducción de los datos, pero a veces queremos hallar un único valor representativo de la distribución. Esta actividad puede realizarse a partir de las gráficas ya elaboradas o también pedir a los alumnos que las construyan previamente.

En el segundo caso la clase puede dividirse en grupos para calcular estos promedios, así como la moda, y para explicar lo que representa cada uno de estos promedios y elegir en cada grupo el que mejor lo representa, argumentando la elección. Se puede pedir a los alumnos que señalen las principales diferencias entre los dos gráficos y que decidan cuál de los dos promedios acentúa más las diferencias explicando la razón.

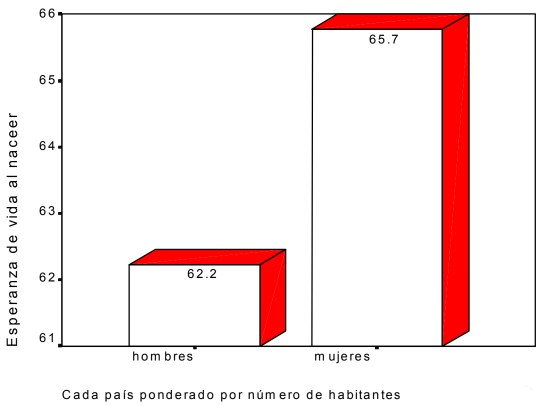
> *3. ¿Qué representa el valor obtenido al calcular la media de la esperanza media de vida al nacer en estos 97 países? ¿Cómo habría que hacer para calcular la esperanza media de vida al nacer en hombres y mujeres, si no tenemos en cuenta el país de nacimiento?.*

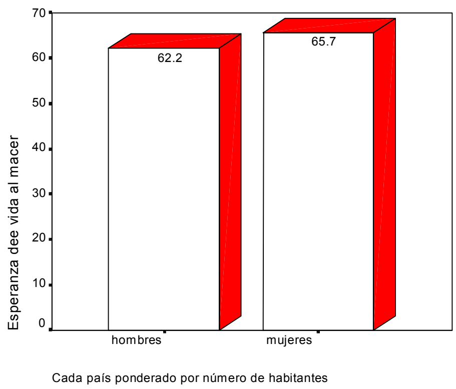
En este fichero las unidades estadísticas son agregados, lo que tiene repercusión en la interpretación de los promedios de las variables; por ejemplo la media (simple) de todas las esperanzas de vida al nacer en los distintos países no es igual a la esperanza global de vida, sino que ésta tiene que ser calculada como una media ponderada en función del número de hombres / mujeres de cada país. Esto puede servir de reflexión sobre las diferentes unidades estadísticas que pueden usarse en un estudio, la necesidad que a veces tenemos de trabajar con valores aproximados y la forma de combinar estudios parciales para obtener índices globales, así como para reflexionar sobre las propiedades de los promedios y la comprensión de las mismas por parte de los estudiantes (Tormo, 1995; Batanero, 2000). Esta misma observación debe hacerse en la interpretación de otros estadísticos, como los de dispersión o los coeficientes de correlación.

Con ejemplos sencillos se podría conducir a los alumnos a la idea de media ponderada y hacerles ver la necesidad de tener en cuenta el número de hombres y mujeres de cada país, para calcular el promedio.

5. *Nosotros hemos calculado la esperanza media de vida global en hombres y mujeres, ponderando los datos de cada país por su número de habitantes (suponiendo un número aproximadamente igual de hombres y mujeres). En la Figura 5.6.2 hemos representado la esperanza media de vida en hombres y mujeres con dos escalas diferentes. Compara estos dos gráficos e indica si te parecen o no adecuados para representar la diferencia entre la esperanza media de vida de mujeres y hombres. Uno de los dos gráficos ha sido obtenido directamente del ordenador, mientras que el otro ha sido manipulado. Averigua cuál ha sido manipulado.*

Figura 5.6.2. Esperanza de vida media en hombres y mujeres





5. ***La tasa de natalidad.*** *Construye ahora una tabla de frecuencias que muestre la distribución de las tasas de natalidad. ¿Por qué en este caso conviene agrupar en intervalos? ¿Qué representa la frecuencia dentro de un intervalos? ¿Cuántos intervalos conviene usar en la tabla de frecuencias? ¿Cómo representaría gráficamente estos datos? ¿Es simétrica la distribución? ¿Cómo cambia la forma al variar el número de intervalos? ¿Y si usamos intervalos de distinta amplitud? ¿Qué representa y cómo representaría la frecuencia acumulada? ¿Qué posición ocupa mi país respecto a la tasa de natalidad? ¿Hay algunos países atípicos respecto a la tasa de natalidad? Para contestar esta y otras preguntas similares puedes usar el gráfico del tallo y hojas (Figura 5.6.3).*

Figura 5.6.3. Gráfico de tallo y hojas: Tasa de natalidad

```
0|99
1|001111222223333333444444
1|556778
2|011222334
2|677888899
3|00111122234
3|5568899
4|011122224444
4|55566677788888
5|0012
```

Cuando la variable presenta un número grande de valores, surge la necesidad de agrupación. Ahora bien, no existe una regla fija sobre la forma de construir los intervalos de clase, aunque se recomienda un número de intervalos aproximadamente igual a la raíz cuadrada del número de datos. En esta actividad se muestra un ejemplo de como en estadística es posible tener más de una solución correcta al mismo problema y también como la forma del histograma varía en función de la elección de los intervalos.

Junto con las representaciones gráficas tradicionales, como diagramas de barras o histogramas, en análisis exploratorio de datos aparecen nuevas representaciones como el diagrama de tallo y hojas o los gráficos de cajas, cuya potencia exploratoria se acentúa con el paso de una a otra representación, así como la selección de partes del fichero para realizar estudios comparativos, por ejemplo al comparar las variables en los distintos grupos de países, que es el tema de la siguiente actividad. El interés no sólo se centra en las tendencias, sino en la variabilidad, así como el estudio de los valores atípicos. Vemos también como la idea de distribución siempre es relativa a un colectivo; por eso un valor puede ser atípico dentro de un subconjunto de datos y no serlo en el global.

6. Diferencias en la tasa de natalidad *Una vez estudiada globalmente la tasa de natalidad dividiremos la clase en grupos para estudiar si la tasa de natalidad es la misma en los diferentes grupos de países. Investigaremos qué grupos de países tienen valores atípicos para la tasa de natalidad y qué países tienen una tasa de natalidad atípica respecto a su grupo así como analizar por qué al realizar un gráfico global en todo el fichero no aparecen valores atípicos y sí aparecen dentro de los grupos. Se pide listar todas las diferencias observadas en los diversos grupos de países.*

Los alumnos elaborarán gráficos como los presentados en las figuras 5.6.4, 5.6.5 y 5.6.6. El gráfico de tallo y hojas ha sido útil para analizar las diferencias en la tasa de natalidad, pero también podemos usar otras representaciones, como diagramas de puntos y gráficos de caja. Divididos en grupos, los alumnos pueden tratar de hallar representaciones alternativas que pongan de manifiesto las diferencias en la tasa de natalidad. Por ejemplo, podemos observar las Figuras 5.6.5 y 5.6.6 y comparar con los gráficos del tronco de la Figura 5.6.4, señalando las ventajas tiene cada una de las representaciones gráficas y cómo podríamos cambiar cada gráfica para resaltar más (menos) las diferencias.

Figura 4 5.6. Distribución de la tasa de natalidad en los diferentes grupos de países

Gráfico del tallo y hojas: Países del Grupo 1

```
1|1
1|22333
1|445
1|7

HI|24.7
```

Gráfico del tallo y hojas: Países del Grupo 2

```
1|8
2|03
2|77889
3|224

HI|46.6
```

Gráfico del tallo y hojas: Países del Grupo 3

```
0|99
1|1111
1|222
1|3333
1|4444
1|5
1|6
```

Gráfico del tallo y hojas: Países del Grupo 4

```
2|22
2|689
3|1
3|8
4|222
4|5
```

Gráfico del tallo y hojas: Países del Grupo 5

```
1|1
1|7
2|1123
2|8
3|00113
3|69
4|012
```

Gráfico del tallo y hojas: Países del Grupo 6

```
LO|31.1

3|
3|
3|2
3|55
3|
3|89
4|1
4|
4|444455
4|66777
4|88888
5|001
5|2
```

Figura 5.6.5. Gráfico de puntos

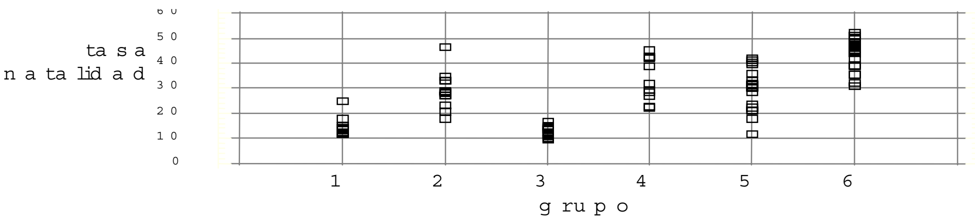
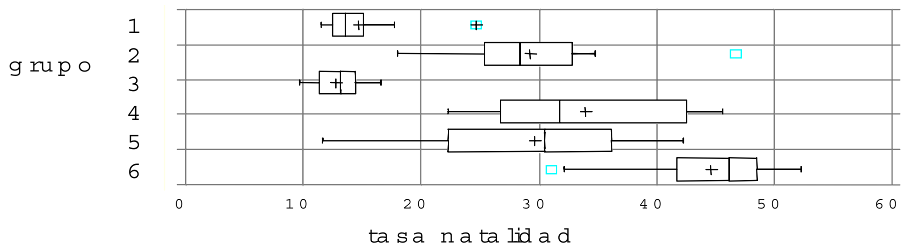


Figura 5.6. 6. Gráficos de caja



*7. **Esperanza de vida en hombres y mujeres.**: En la actividad anterior hemos comparado una variable en distintos subconjuntos de países, pero a veces tiene también sentido comparar dos variables diferentes en el total de los datos. Se dice, por ejemplo, que las mujeres forman el sexo débil. Sin embargo muchas chicas estarían en contra de esta opinión cuando comparen en algunos países la esperanza de vida al nacer en hombres y mujeres. Podríamos investigar si la distribución de la esperanza de vida en el total de países es igual en hombres y mujeres analizando para ello los gráficos de las figuras 5.6.7, 5.6.8 y 5.6.9 y contestar preguntas como las siguientes: ¿Hay mayor variabilidad entre países en la esperanza de vida en hombres o en mujeres? ¿En qué porcentaje de países la esperanza de vida de hombres (mujeres) es mayor de 58 años? ¿ y de 68 años? ¿Cuál es el valor de la esperanza de vida de modo que el 70 por ciento de países tiene una esperanza de vida mayor? ¿Es igual en hombres y mujeres? ¿Cómo podrías disimular la diferencia de esperanzas de vida en hombres y mujeres en cada gráfico?*

Figura 5.6.7: Gráficos de caja. Distribución de la esperanza de vida en hombres y mujeres

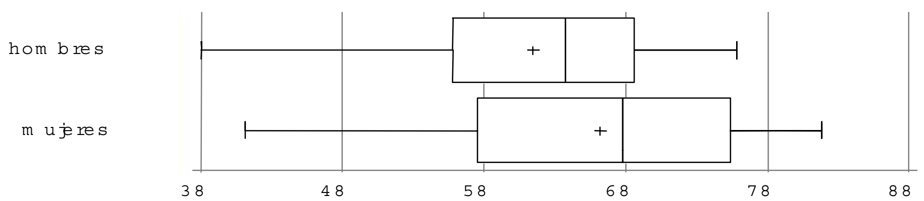
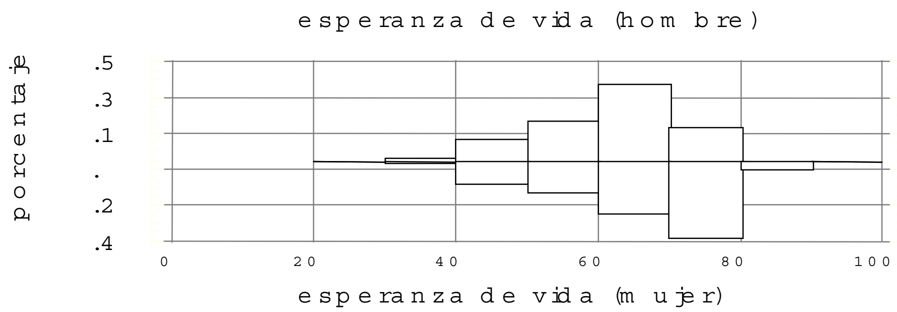
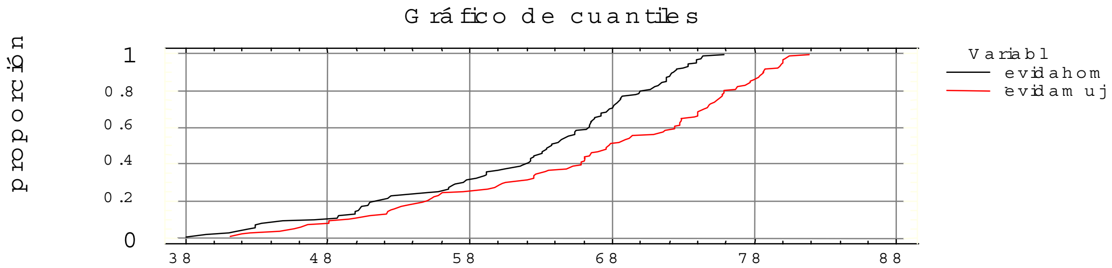


Figura 5.6.8: Histogramas. Distribución de la esperanza de vida en hombres y mujeres



En esta actividad se vuelven a aplicar los conceptos aprendidos al estudio de las diferencias de dos variables en una misma muestra. Se introduce también el estudio de las frecuencias acumuladas, percentiles y sus rangos. Hacemos notar que el cálculo de percentiles y sus rangos puede hacerse directamente a partir de la gráfica sin tener que recurrir a las fórmulas de cálculo que son complejas y para las que existe una variedad de casos, según las variables estén o no agrupadas en intervalos.

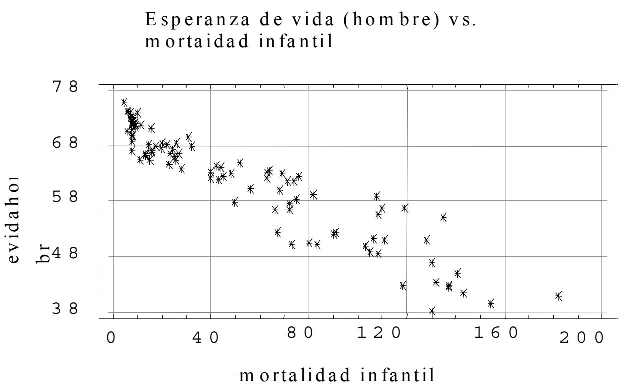
Figura 5.6.9. Distribución acumulativa de la esperanza de vida en hombres y mujeres

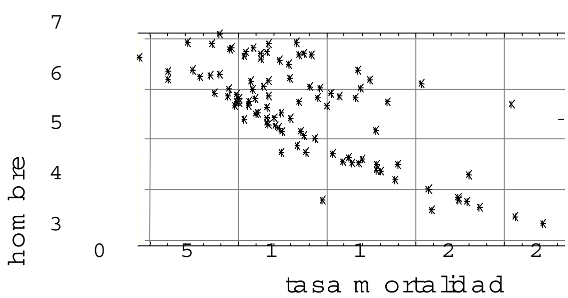


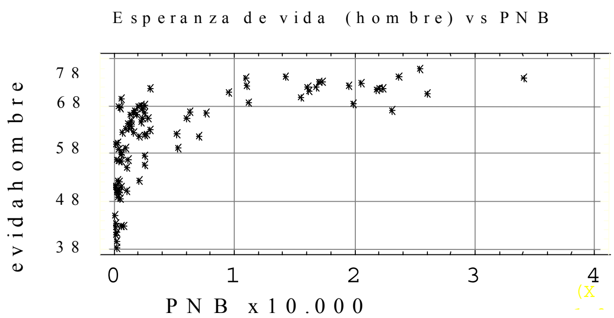
5.6.4. Actividades de ampliación

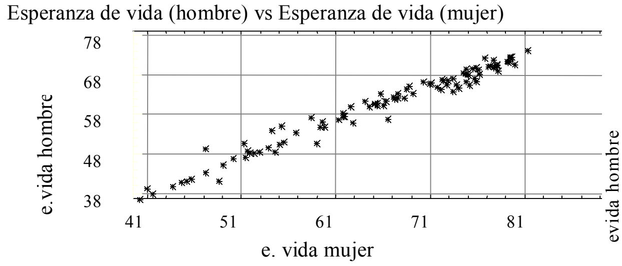
8. ***Relación de la esperanza de vida con otras variables.*** *En la Figura 5.6.10 hemos representado los diagramas de dispersión de la esperanza de vida del hombre en cada país, en función de diversas variables. ¿Cuál de las variables está relacionada con la esperanza de vida del hombre? ¿En cuáles la relación es directa /inversa? ¿Cuáles influyen o son influidas por la esperanza de vida? ¿Cuál de ella sirve mejor para predecir la esperanza de vida? ¿En qué casos la relación podría ser debida a otras variables? ¿Podríamos en alguno de los casos hallar una función matemática para predecir, aproximadamente la esperanza de vida del hombre a partir de la otra variable? ¿Qué tipo de función?*

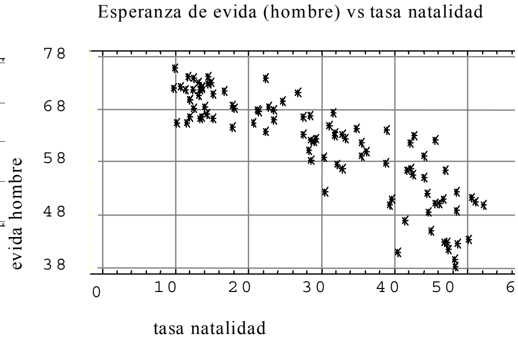
Figura 5.6. 10. Relación de la esperanza de vida del hombre con otras variables











Al analizar las relaciones entre dos variables numéricas los alumnos deben de extender la idea de dependencia funcional a dependencia aleatoria y diferenciar sus tipos (lineal o no; directa e inversa) así como graduar al menos intuitivamente la intensidad de la relación. Es importante también diferenciar correlación y causalidad y analizar los distintos tipos de relaciones que pueden llevar a la existencia de correlación: dependencia causal; interdependencia, dependencia indirecta, concordancia y correlación espúrea (Estepa, 1995). Una vez detectada la correlación, el interés se centra en la búsqueda de modelos que puedan predecir las variables explicadas en función de las variables explicativas, lo que de nuevo conecta con otro contenido del currículo de matemáticas: las funciones.

5.6.5. Algunas dificultades y errores previsibles

El primer paso en el análisis es el estudio de cada variable, la tabulación y representación gráfica. Algunos investigadores han analizado los diferentes niveles de comprensión de las gráficas (Curcio, 1989).y las dificultades de los alumnos en la elaboración de las mismas o la selección de un gráfico adecuado, debido a la diferente información que aportan las diversas gráficas estadísticas (Li y Shen, 1992). Deberíamos también fomentar en los alumnos un sentido gráfico que les haga ser críticos frente a los posibles gráficos tendenciosos que con frecuencia encontramos en los medios de comunicación.

El trabajo con análisis exploratorio de datos refuerza algunos objetivos sugeridos para la educación matemática en los nuevos currículos de secundaria, como el trabajo con problemas abiertos, el uso de sistemas múltiples de representación, la introducción al trabajo con ordenadores o calculadoras gráficas y la conexión de las matemáticas con otras áreas del currículo (Shaughnessy, Garfield y Greer, 1997).

Pero el razonamiento estadístico va más allá del conocimiento matemático y de la comprensión de los conceptos y procedimientos. La modelización, la valoración de la bondad del ajuste de los modelos a la realidad, la formulación de cuestiones, la interpretación y síntesis de los resultados, la elaboración de informes son también componentes esenciales de las capacidades que queremos desarrollar en nuestros alumnos. Los alumnos, acostumbrados a que en la clase de matemáticas cada problema tiene una única solución, podrían encontrar complejo el trabajo con los proyectos y la existencia de diferentes procedimientos y soluciones correctas. Es labor del profesor acostumbrarles al método y razonamiento estadístico.

Los estadísticos de orden pueden plantear problemas, ya que la función de distribución de la variable discreta tiene discontinuidades de salto y los alumnos no están acostumbrados a este tipo de funciones. Asimismo, alguno de ellos pudiera tener deficiencias en el razonamiento proporcional. Respecto a la correlación y regresión es una buena ocasión de aplicar los conocimientos previos de los alumnos sobre la función lineal y otras funciones sencillas. El profesor debe recordar que no siempre se puede despejar la variable independiente para encontrar la línea de regresión de X sobre Y, porque el criterio de ajuste es diferente del seguido en la construcción de la línea de regresión de Y sobre X.

5.6.6. Análisis del contenido estadístico

En este proyecto podemos identificar, explícita o implícitamente los siguientes contenidos:

1. *Aplicaciones de la estadística:*
- Estudios demográficos y socioeconómicos;
- Estadísticas oficiales; organismos y procedimientos en la elaboración de estadísticas oficiales;
- Fuentes de datos estadísticos: anuarios estadísticos; fuentes de datos en Internet.

2. C*onceptos y propiedades*:
- Agrupación de variables en intervalos; efectos de la agrupación;
- Medidas de posición: medias ponderadas**,** percentiles, rangos de percentiles;
- Valores atípicos y su efecto sobre los promedios;
- Asociación y sus tipos: directa/ inversa; lineal no lineal; asociación y causalidad;
- Modelos; ajuste de modelos sencillos a datos bivariantes; uso de modelos en la predicción.

3. Notaciones y representaciones:
- Palabras como, intervalos, extremos y marcas de clase, percentiles y sus rangos, cuartiles, valores atípicos, asociación, correlación, regresión;
- Símbolos usados para los diferentes estadísticos y otros como el sumatorio;
- Diagrama de tallo y hojas, gráficos de caja, gráficos de cuantiles, diagramas de dispersión.

4.*Técnicas y procedimientos:*
- Búsqueda de datos a partir de anuarios estadísticos o de la Internet;
- Elaboración de tablas de frecuencia con datos agrupados;
- Elaboración de gráficos de tallo y hojas, gráficos de caja, gráficos de cuantiles y diagramas de dispersión;
- Interpretación de tablas y gráficos; elaboración de conclusiones a partir del análisis de tablas y gráficos;
- Elaboración de argumentos y conclusiones a partir del análisis de datos;
- Uso de calculadora gráfica, hojas de cálculo o software estadístico.

5. A*ctitudes:*
- Valorar la utilidad y complejidad de la elaboración de las estadísticas oficiales y concienciarse de la importancia de su colaboración en encuestas y censos para obtener datos fiables;
- Concienciar al alumno sobre la importancia de las representaciones gráficas y la posibilidad de que se transmita información sesgada en una gráfica mal construida;
- Fomentar un espíritu crítico en el uso de paquetes estadísticos y sus opciones por defecto.

5.7.. Proyecto 4. Las matemáticas de la catadora de té

5.7.1. Objetivos

Este proyecto está basado en el artículo de Sir Ronald Fisher reproducido en Sigma: El mundo de las matemáticas. El propósito es plantear una situación que lleve a analizar una hipótesis y estudiar la forma en que se podría organizar un experimento, recoger los datos y analizarlos para llegar a una decisión acerca de si aceptamos o no la hipótesis como provisionalmente cierta.

Mediante el proyecto se trata de introducir algunos conceptos básicos de inferencia, y mostrar a los alumnos la utilidad de la inferencia estadística para poner a prueba hipótesis experimentales. Asimismo se pretende analizar el razonamiento de las pruebas de significación estadística introducidas por Fisher y comparar con los razonamientos deductivos que empleamos en otras ramas de las matemáticas. También se quiere concienciar a los alumnos sobre posibles interpretaciones erróneas de los resultados de las pruebas estadísticas de hipótesis.

Alumnos: El proyecto está pensado para alumnos de Bachillerato, aunque podría ser también útil en la formación universitaria para alumnos de ciencias sociales o humanas.

5.6.2. Los datos

Se utilizan dos tipos de datos:

- Datos hipotéticos sobre posibles resultados de un experimento;
- Datos obtenidos mediante simulación de experimentos aleatorios, con la finalidad de obtener en forma empírica las distribución en el muestreo de una proporción; Las simulaciones se realizarán con material manipulativo, tablas de números aleatorios y calculadoras u ordenadores.

5.6.3. Preguntas, actividades y gestión de la clase

Se comienza la clase relatando al alumno el episodio de la catadora de te descrito por Fischer o bien se da a los alumnos el día anterior al comienzo de la actividad la fotocopia del artículo reproducido en la enciclopedia de las matemáticas. Se plantea la siguiente pregunta:

1. Una señora afirma que al probar una taza de te con leche puede distinguir qué fue lo primero que se echó en la taza: el té o la leche. ¿Cómo podríamos hacer un experimento para comprobar si la señora tiene razón?

Se recogen sugerencias de la clase, discutiendo con los alumnos si el hecho de acertar una sóla vez significaría que la señora tiene razón. Se buscan ejemplos de situaciones en que podemos adivinar un resultado, simplemente por azar. Se continúa con la siguiente discusión:

2. Se organiza el siguiente experimento: Preparamos ocho tazas de te con leche. Antes de preparar cada una, y sin que lo vea la señora lanzamos una moneda: si sale cara echamos primero el té y en caso contrario la leche. Supongamos que la señora acierta el orden en las *tres primeras tazas. ¿Piensas que tiene razón, cuando dice que puede distinguir lo primero que se puso en la taza? ¿O podría haber acertado, justo por azar?*

Pedimos a la clase calcular la probabilidad de que la señora acierte una, dos y tres tazas por azar. Puesto que sólo hay dos posibilidades en cada taza sólo puede producirse un acierto o un fallo. Las diferentes posibilidades que pueden presentarse con una, dos y tres tazas se presentan en la tabla 5.7.1.

Tabla 5.7.1. Diferentes posibilidades de acierto en una, dos y tres tazas si la señora miente

A	F						
AA	AF	FA	FF				
AAA	AAF	AFA	AFF	FAA	FAF	FFA	FFF

En el caso que la señora mienta, podemos considerar que tiene las mismas posibilidades de acertar o fallar en cada taza, esto es la probabilidades de acierto y fallo son igual a 1/2. Para dos tazas, al haber cuatro posibilidades la probabilidad de acertar en los dos casos es 1/4 y la probabilidad de acertar en tres tazas 1/8. Puesto que un suceso que ocurre una vez cada ocho veces no es demasiado raro, el caso de que la señora acertara las tres primeras tazas podría darse por puro azar. Sin embargo, a medida que aumentamos el número de tazas, la probabilidad de acertarlas todas o de acertar la mayoría es cada vez más pequeña, *en el caso de que la señora mienta.* Planteemos la siguiente actividad para calcular la probabilidad de acertar 1, 2,...8 tazas en caso de que la señora mienta.

3. Supongamos que la señora miente y trata de adivinar al azar el orden en que se puso el te y la leche en 8 tazas de te con leche. Podemos sustituir el experimento original por el de lanzar 8 monedas al aire (cada moneda representa una taza). Si sale cara, suponemos que hemos acertado el orden en que se puso el te y la leche en la taza (cuya probabilidad es 1/2) y si sale cruz hemos fallado. Supongamos que hago la prueba y obtengo los siguientes resultados: AAAFAFFA

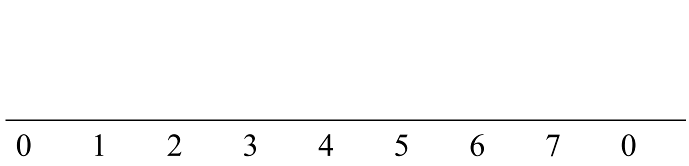
Esto significa que la señora ha acertado 5 veces de las ocho. Anoto un punto el el número de aciertos igual a 5. ¿Quiere decir esto que la señora miente? ¿Y si hubiera acertado todas, significa que dice la verdad? Realiza el experimento 10 veces y representa en el siguiente diagrama de puntos (gráfico 5.7.1) el número de aciertos obtenidos.

Figura 5.7.1. Número de aciertos por azar en 8 tazas de te (datos del alumno)

0 1 2 3 4 5 6 7 0

El profesor pregunta a los chicos los datos que han obtenido y si les parece fácil o difícil acertar 8 tazas por azar. A continuación se recogen los datos de toda la clase con un gráfico de puntos (Figura 5.7.2). En una clase de 25-30 alumnos se obtendrá una distribución acampanada, con la máxima frecuencia en los valores centrales (3-5 tazas) y con una frecuencia muy baja en los extremos (1-2 o 7-8). A partir de la distribución obtenida se pueden plantear las siguientes preguntas:

Figura 5.7.2. Número de aciertos por azar en 8 tazas de te (datos de la clase)



4 ¿Cuál/ cuáles es el número más probable de aciertos por azar en 8 tazas? ¿Cuál es la probabilidad de acertar 5 tazas o mas? ¿Y de acertar 7 o más? ¿A partir de qué número de tazas admitirías que la señora tiene razón?

A continuación se inicia una discusión sobre lo que se ha hecho, en términos parecidos a los siguientes:

5. En la actividad anterior hemos tomados datos experimentales para poner a prueba la hipótesis de que una señora adivina el orden en que se pone la leche y el te en una taza de te por puro azar. Hemos seguido el siguiente procedimiento de contraste de hipótesis:

a) Hemos considerado nuestra hipótesis H_0: "la señora hace sus apuestas sobre el orden del te y la leche al azar". Puesto que se trata de un fenómeno aleatorio no podemos probar deductivamente nuestra hipótesis en el mismo modo que hacemos una demostración, por ejemplo, en Geometría.

b) Hemos recogido unos datos para poner a prueba nuestra hipótesis H_0 (que llamaremos hipótesis nula): la serie de resultados AAAFAFFA. Esta serie es una muestra de todos los posibles resultados aleatorios que podría obtener si la señora hace sus apuestas al azar en una serie de ocho tazas de te.

c) Puedo tomar muestras más o menos grandes (series mas o menos largas de ensayos). Cuanto mayor sea la muestra más seguro estaré de la prueba de mi hipótesis. Por ejemplo, estoy más seguro si hago la prueba con 8 tazas que con 3 y estaría todavía más seguro si hiciera la prueba con 20 o 30 tazas.

d) Hemos supuesto que nuestra hipótesis H_0 (que llamaremos hipótesis nula) es cierta. Es decir, suponemos que la señora miente y hace sus apuestas al azar. Por tanto tiene las mismas posibilidades de acertar o no cada una de las tazas. A continuación calculamos la probabilidad que tienen los datos que he obtenido (cinco aciertos en ocho tazas) si la hipótesis nula fuese cierta. También hemos calculado la probabilidad de 7 u 8 aciertos en 8 tazas

e) He visto que mis datos serían muy poco probables si la hipótesis nula fuese cierta. De hecho teóricamente podemos probar que la probabilidad de obtener 7 u 8 aciertos en 8 ensayos, si fuese cierta la hipótesis nula, sería igual a 0.035. Como consecuencia, debo tomar la siguiente decisión :

- O admito que mi hipótesis H_0 es correcta (la señora miente y dice los resultados al azar) y que, por buena suerte ha obtenido un suceso muy poco probable.

- Mi hipótesis era falsa; la señora efectivamente es capaz de distinguir en qué orden se pone el te y la leche en la taza. Como consecuencia la rechazo, porque los datos experimentales

contradicen la hipótesis nula.

Podemos realizar nuevos estudios de la distribución del número de aciertos en n tazaz variando n, usando los programas de simulación. Por ejemplo, en Statgraphics el programa Probability Distribution permite simular y obtener valores de las tablas de una serie de distribuciones, incluyendo la binomial y cambiar los parámetros de las mismas (Figura 5.7.3.)

Cambiando los parámetros por defecto, podemos simular y grabar en una variable de un fichero de datos una muestra aleatoria de valores de la distribución binomial B(8, 0.5) y representar gráficamente sus resultados en un diagrama de barras mediante la opción DESCRIBE, como hemos hecho nosotros y cuyos resultados se presentan en la figura 5.7.4.

Figura 5.7.3. Opciones de simulación y tablas de distribuciones en Statgraphics

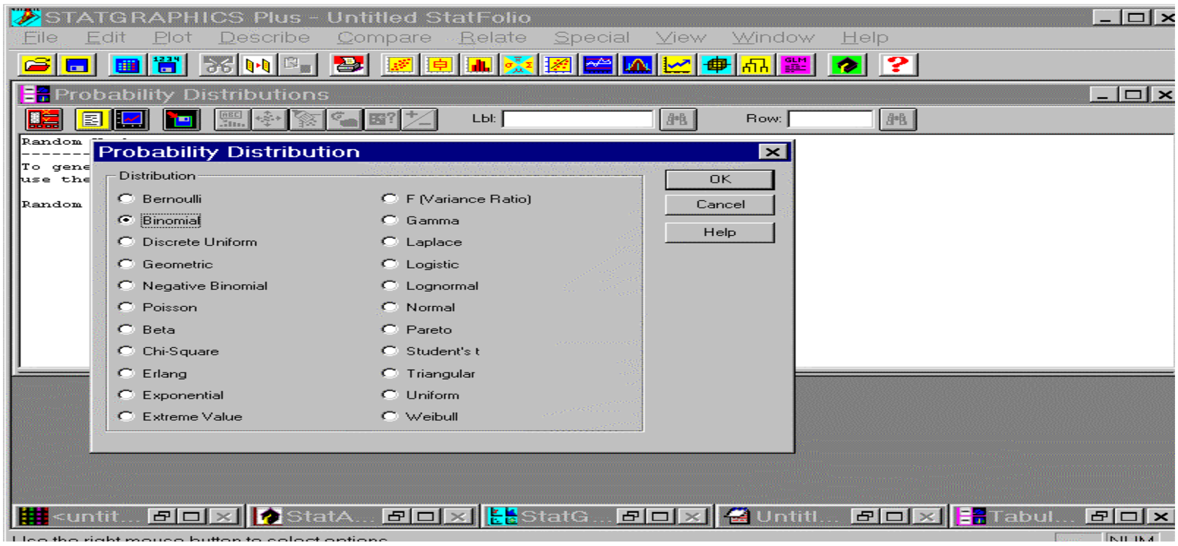
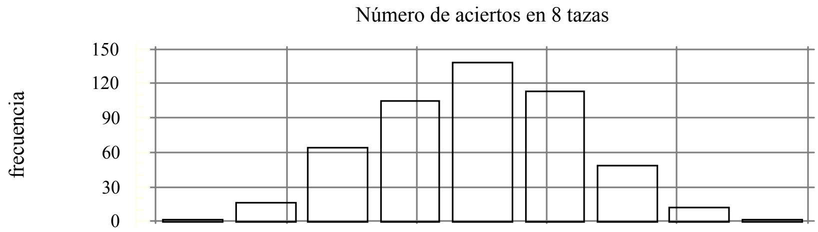
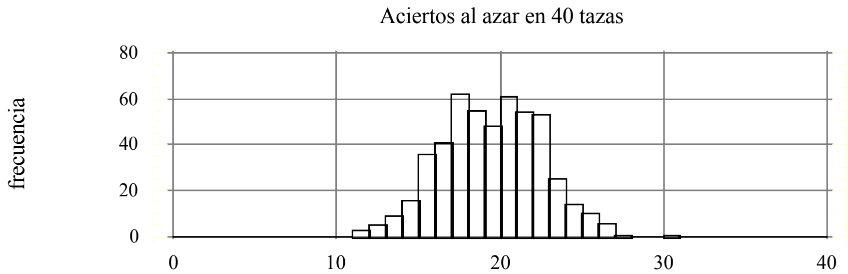


Figura 5.7.4. Resultados de 500 simulaciones de aciertos al azar en 8 tazas de te



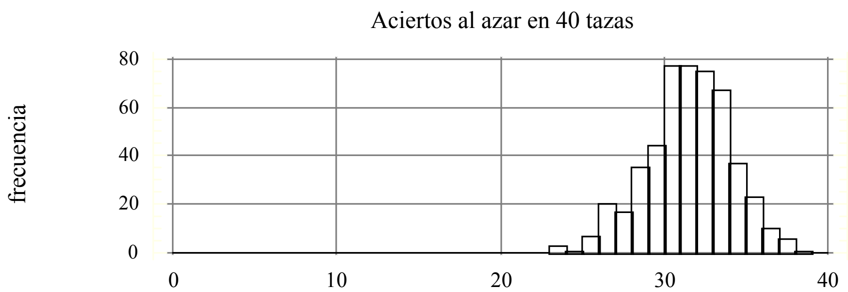
Los alumnos pueden experimentar con las opciones del programa cambiando por ejemplo, el número de tazas (ensayos en la distribución binomial) y conservando la probabilidad de acierto igual a a1/2. Como se muestra en la figura 5.7.5. el tamaño de la muestra reduce la dispersión relativa de los resultados. Mientras que acertar 2 de las 8 tazas (la cuarta parte de resultados) era relativamente fácil acertar por azar la misma proporción con 40 tazas es prácticamente imposible.

Figura 5.7.5. Resultados de 500 simulaciones de aciertos al azar en 40 tazas de te



Podemos discutir con los alumnos la importancia de la hipótesis sobre la distribución obtenida en el muestreo. Si suponemos que la señora tiene una mayor probabilidad de acertar que de fallar (por ejemplo si la probabilidad de acierto es 0.8) y repetimos la simulación, observamos como ahora los valores centrales se aproximan al valor esperado. Aunque las muestras que tomamos de una población de valores son aleatorias, y, por tanto sus resultados varían de una a otra muestra, en un conjunto muy grande de muestras la proporción muestral se distribuye alrededor del valor verdadero de la proporción en la población. Por otro lado, a mayor tamaño de muestra es menor la dispersión de los diferentes valores de las proporciones muestrales. Estas dos propiedades son las que nos permiten extraer conclusiones sobre toda una población a partir de los datos de la muestra.

Figura 5.7.6. Resultados de 500 simulaciones de aciertos en 8 tazas de te cuando la propobabilidad de acierto es 0.8



Finalmente se hace un resumen de lo analizado en la clase:

6. En las actividades anteriores hemos descrito el procedimiento que, generalmente, se sigue en estadística para poner a prueba una hipótesis. También hemos visto que nunca tenemos total seguridad en nuestra decisión. Existe una probabilidad 0.0035 de obtener siete u ocho aciertos al azar en una serie de 8 tazas de tem aunque no sepamos distinguir si se puso primero el te o la leche, es decir si admitimos como cierta la hipótesis nula. Esta probabilidad la asumo como riesgo sobre mi decisión al decidir que la hipótesis es falsa.

Hemos decidido rechazar la hipótesis. Volvamos a las tres primeros tazas. Hagamos la prueba de la hipótesis sólo con estas tres tazas acertadas. La probabilidad de obtener tres aciertos seguidos simplemente por azar es 1/8= 0.125. Esto quiere decir que no sería demasiado raro nuestro resultado que ocurre 1 de cada ocho veces. Como consecuencia vemos que con una muestra demasiado pequeña es difícil llegar a ninguna conclusión.

5.7.4. Actividades de ampliación

7. Supongamos que se obtuvieron los siguientes resultados en las pasadas elecciones: El 40% del total de los votantes, votaron al PP, el 38% votó al PSOE y 9% votó a IU. Si en esta ciudad tomamos una muestra aleatoria de 100 votantes y les preguntamos a quien votaron (imaginamos que las personas a las que preguntamos son sinceros): Podemos decir que necesariamente de estos 100 votantes, 40 votaron al PP, 38 al PSOE y 9 a IU?

8. Supongamos que tomamos una varias muestras aleatorias de 100 votantes. ¿Encontraremos siempre la misma proporción de votantes a cada partido en cada muestra? ¿Podrías adivinara, aproximadamente el porcentaje aproximado de personas que en cada muestra habrían votado al PP? Supongamos ahora que tomamos una muestra de 100 votantes en el País Vasco. ¿Crees que variarían los resultados?

9. Si quieres comprar un coche nuevo y quieres decidir entre la marca A y B. En una revista de automóviles encuentras una estudio estadístico sobre reparaciones efectuadas el último años que muestra que la marca A tiene menos averías que la B. Sin embargo, te encuentras un amigo tuyo que te dice que compró el año pasado un coche B y no ha tenido más que problemas: primero se le estropeó la inyección de gasolina y gastó 25.000 pts, luego tuvo que cambiar el eje trasero y al final, ha vendido el coche porque se le fue la transmisión. ¿Que decisión tomarías, comprar un coche A o B?

10. Discute en cuál de los siguientes estudios por muestreo habrá más variabilidad y en cuál habrá más representatividad: a)Tomar al azar muestras de 10 votantes para estimar la proporción de personas que votaron al PSOE ; b)Tomar al azar muestras de 1000 votantes para estimar la proporción de personas que votaron al PSOE;c)Tomar al azar muestras de 1000 votantes para estimar la proporción de personas que votaron a IU;d) Tomar al azar muestras de 10 votantes para estimar la proporción de personas que votaron a IU; e) Tomar muestras de 1000 jubilados para estimar la proporción de personas que votaron al PSOE; f)Tomar muestras de 10 personas al azar para estimar la proporción de mujeres.

11. La siguiente ficha técnica aparece publicada en un artículo de ELPAIS (26 de Abril de 1998) refiréndose a una encuesta efectuada por Demoscopia:
FICHA TECNICA: Encuesta realizada por Demoscopia SA. Ambito: nacional. Universo: población de 18 años y Más. Tamaño de la muestra: 1200 entrevistas afijadas mediante muestreo estratificado por región y tamaño de habitat proporcional a la distribución de la población y con aplicación de cuotas de sexo y edad. Error de muestreo: asumiendo los criterios de muestreo simple, para un nivle de confianza del 95,5%(dos sigmas) y para la hipótesis más desfavorable(p=q=50) el error para el total de la muestra sería de 62.9%. Método de recogida de la información: entrevista telefónica asistida por ordenador.
Elabora una lista de todos los términos que no comprendes de esta ficha técnica y busca su definición en un libro de estadística o preguntando a tu profesor. Elabora una lista de los posibles factores que pueden influir sobre la fiabilidad de esta encuesta.

5.7.5. Algunas dificultades y errores previsibles

En este proyecto tratamos de introducir algunas ideas sobre inferencia, las más importantes de las cuales es que una muestra proporciona "alguna" información sobre la población y de este modo aumenta nuestro conocimiento sobre la misma y que la estadística proporciona una colección de métodos para aprender de la experiencia.

La comprensión de esta idea básica implica el equilibrio adecuado entre dos ideas aparentemente antagónicas: la representatividad muestral y la variabilidad muestral. La primera de estas ideas nos sugiere que la muestra tendrá a menudo características similares a las de la población, si ha sido elegida con las precauciones adecuadas. La segunda, el hecho de que no todas las muestras son iguales entre sí.

Las investigaciones psicológicas han mostrado, sin embargo que ponemos un énfasis excesivo en la representatividad, olvidando la variabilidad y el efecto del tamaño de la muestra sobre la misma.

El contraste de hipótesis es un procedimiento muy complejo, ya que lleva asociado una serie de conceptos que el alumno debe diferenciar. En contra de lo que se cree, existen varias escuelas estadísticas sobre el contraste de hipótesis. Nosotros en este proyecto hemos seguido el enfoque de Fisher, quien no introduce formalmente (aunque usa implícitamente) la idea de hipótesis alternativa. Hemos evitado también algunos conceptos que no son necesarios a este nivel intuitivo como nivel de significación y regiones críticas/ de aceptación.

Una dificultad importante al iniciar el estudio de las distribuciones en el muestreo es que se presentan diferentes niveles para los mismos conceptos. En este tema particular estamos manejando dos tipos de proporciones: la proporción en la población (probabilidad de acierto) y la proporción en cada muestra (proporción de aciertos en la muestra particular). Mientras que la primera es una constante desconocida (no sabemos cuál es la verdadera proporción de aciertos de la señora) la segunda es una variable aleatoria, porque varía de una muestra a otra.

El principal fin de este proyecto es observar esta variabilidad y la distribución de la proporción muestral, así como sus características. La capacidad de simulación y representación gráfica de los ordenadores se pone al servicio de la visualización de estos conceptos. No obstante, pensamos que es necesario realizar algunas actividades de simulación con dispositivos concretos, como monedas.

Otra dificultad importante del tema es el uso de la probabilidad condicional. Todos nuestros cálculos, razonamientos y simulaciones se apoyan en una condición (si la señora miente). Pero, a veces, los alumnos confunden los dos términos de una probabilidad condicional. El profesor tendrá que insistir en la diferencia entre las preguntas:

- ¿Cuál es la probabilidad de que la señora acierte 8 tazas si hace sus apuestas al azar? Esta probabilidad es $1/2^8$
- ¿Cuál es la probabilidad de que la señora apueste al azar si ha acertado 8 tazas? La estadística clásica no nos da una respuesta a esta pregunta; tendríamos que ir a la inferencia bayesiana para poder responderla.

5.7.6. Análisis del contenido estadístico

En este proyecto podemos identificar explícita o implícitamente los siguientes contenidos:

1. Aplicaciones de la estadística:

- Prueba de una hipótesis;
- Diseño de experimentos;
- Votaciones, resultados de elecciones; encuestas de intención de voto;
- Consumo; orientación al consumidor.

2. Conceptos y propiedades:

- Experimento aleatorio; secuencia de resultados aleatorios, sucesos equiprobables, independencia de ensayos;
- variable estadística y variable aleatoria, dstribución de frecuencias y distribución de probabilidad;
- moda, mediana en una distribución de frecuencias y una distribución de probabilidad, rangos de percentiles;
- dispersión, rango, máximo, mínimo, casos centrales;
- hipótesis; contraste de una hipótesis, riesgo en una decisión, uso de modelos en el contraste de una hipótesis;
- Población, censo, muestra, representatividad y sesgo en el muestreo, variabilidad en las muestras; efecto del tamaño sobre la variabilidad.

3. Notaciones y representaciones:

- Palabras como hipótesis, contraste de hipótesis, riesgo, representatividad y variabilidad muestral;
- Diagrama de puntos, diagrama de barras, histogramas;
- Representación de experimentos aleatorios mediante simulación con materiales concretos y con ordenador.

4. Técnicas y procedimientos:

- Recogida y registro de datos experimentales;
- Recuento y cálculo de frecuencias;
- Representación gráfica de datos e interpretación de tablas y gráficos;
- Uso de ordenadores para simulación y representación gráfica.

5. Actitudes:

- Reflexión sobre las propias intuiciones incorrectas respecto al muestreo y la inferencia;
- Valoración de la utilidad de la estadística para poner a prueba hipótesis, organizar experimentos y analizar los datos de los mismos;
- Valoración de la estética y claridad en la elaboración de tablas y gráficos.

5.8. Proyecto 5. La estadística como herramienta de clasificación

5.8.1. Objetivos

Se trata de buscar criterios de clasificación de sujetos en función de sus atributos físicos para poder asignar en el futuro a un sujeto al grupo respecto al cual tiene un mayor parecido global. En concreto utilizaremos cuatro medidas de las hojas de tres especies diferentes de Iris (Setosa, Virgínica y Versicolor) para tratar de determinar una función de estas cuatro medidas o función discriminante que permita a una persona sin conocimientos de botánica clasificar una planta Iris en función de estas medidas en una de las tres especies dadas, de modo que se minimice el número de errores cometidos.

La finalidad principal es mostrar la utilidad de la estadística en la construcción de modelos predictivos y clasificatorios que tienen una gran aplicabilicad tanto en la taxonomía en ciencias como botánica o zoología, el diagnóstico médico, psicología y otras disciplinas.

Un ultimo objetivo es una iniciación intuitiva a las técnicas estadísticas multivariantes.

Alumnos: Este tema está pensado para trabajar con alumnos de Bachillerato o primeros cursos de Universidad, preferentemente con algunos conocimientos de informática, aunque el tema puede usarse para introducirlos al uso de software estadístico.

5.8.2. Los datos

Los datos provienen del trabajo original de Fisher sobre análisis discriminante y fueron publicados en Fisher, R. A. (1936). The Use of Multiple Measurements in Axonomic Problems. Annals of Eugenics 7, 179-188. Nosotros lo hemos tomado del servidor de *Journal of Statistics Education* en Internet y, en lugar de usar el fichero completo (50 observaciones de cada especie) lo hemos limitado a 30 observaciones de cada especie. En la Tabla 5.8.1. se presentan los datos y variables recogidas.

Tabla 5.8.1. Medidas de cuatro características en tres especies de Iris

	Código	Especie	Ancho pétalo	Longitud Pétalo	Ancho Sépalo	Longitud Sépalo
1	1	I. Setosa	02	14	33	50
2	1	I. Setosa	02	10	36	46
3	1	I. Setosa	02	16	31	48
4	1	I. Setosa	01	14	36	49
5	1	I. Setosa	02	13	32	44
6	1	I. Setosa	02	16	38	51
7	1	I. Setosa	02	16	30	50
8	1	I. Setosa	04	19	38	51
9	1	I. Setosa	02	14	30	49
10	1	I. Setosa	02	14	36	50
11	3	I. Verginica	24	56	31	67
12	3	I. Verginica	23	51	31	69
13	3	I. Verginica	20	52	30	65
14	3	I. Verginica	19	51	27	58
15	3	I. Verginica	17	45	25	49
16	3	I. Verginica	19	50	25	63

17	3	I. Verginica	18	49	27	63
18	3	I. Verginica	21	56	28	64
19	3	I. Verginica	19	51	27	58
20	3	I. Verginica	18	55	31	64
21	2	I.Versicolor	13	45	28	57
22	2	I.Versicolor	16	47	33	63
23	2	I.Versicolor	14	47	32	70
24	2	I.Versicolor	12	40	26	58
25	2	I.Versicolor	10	33	23	50
26	2	I.Versicolor	10	41	27	58
27	2	I.Versicolor	15	45	29	60
28	2	I.Versicolor	10	33	24	49
29	2	I.Versicolor	14	39	27	52
30	2	I.Versicolor	12	39	27	58

5.8.3. Preguntas, actividades y gestión de la clase

Una vez que los alumnos tienen la hoja de recogida de datos, el profesor les relata la investigación de Fisher sobre análisis discriminante, en donde el autor buscaba la manera de clasificar individuos en función de sus características físicas. Este tipo de técnica sería útil para predecir, por ejemplo, si un niño nacerá con peso por debajo del normal, en función de las constantes físicas de la madre en el embarazo; para diagnosticar una cierta enfermedad, en función de una serie de síntomas, para seleccionar un deportista de élite, en función de su rendimiento en una serie de pruebas, etc.

Este proyecto requiere de un mayor número de cálculos y gráficos. Por ello, o bien los alumnos trabajan en parejas con un ordenador, o se requiere un ordenador en la clase, cuyas salidas sean proyectables en una pantalla. Otra alternativa sería proporcionar a los alumnos fotocopias de las salidas de ordenador, a medida que se preparan.

Una vez descritas las variables, la primera actividad será preparar representaciones gráficas de cada una de las cuatro variables, clasificadas según especie. Los alumnos analizan las diferencias observables en cada variable para cada una de las especies. Por ejemplo, respecto al ancho del pétalo:

1. Representa gráficamente la anchura del pétalo, diferenciando las tres especies. ¿Qué diferencias observas? ¿Podrías dar un criterio para clasificar una planta en una de las tres especies, en función de la anchura del pétalo?

Los alumnos podrían producir algunas representaciones gráficas similares a la Figura 5.8.1 y estudiar, asimismo algunos estadisticos de la variable en los diferentes grupos, tales como los presentados en la Tabla 5.8.2.

Figura 5.8.1. Diagrama de puntos de anchura pétalo en tres especies de Iris

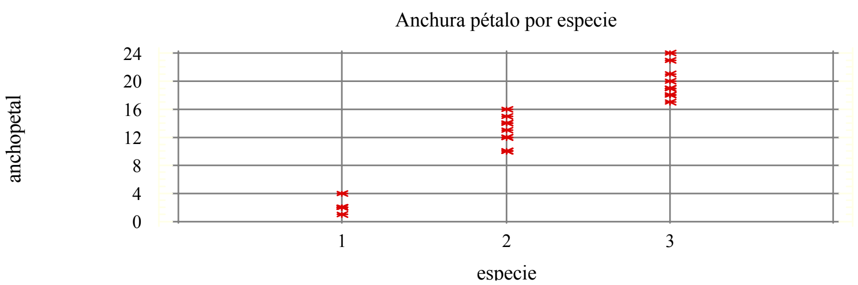


Tabla 5.8.2. Resúmenes estadísticos de la anchura de pétalo en tres especies de Iris

	Setosa (1)	Virgínica (2)	Versicolor (3)
Media	2.1	12.6	19.8
D. Típica	.74	2.17	2.25
Mínimo	1	10	17
Máximo	4	16	24

La clasificación de las plantas a partir de la anchura del pétalo parece fácil, porque no hay solapamiento del rango de valores de la variable.

2. Supongamos una nueva planta con anchura de pétalo 7 cm. ¿Dónde la clasificarías?

Es claro que esta planta se parece más a la especie Setosa que a las demás, porque el valor absoluto de la diferencia de su anchura de pétalos con la anchura media en esta especie (4.9 cm.) es menor que respecto a Virgínica (5.6) o Versicolor (12.8). Asignamos la nueva planta a Setosa porque la *similaridad* con esta especie es mayor, de modo que tratamos de aumentar la similaridad de individuos de la misma especie y también las *diferencias* entre especies.

Matemáticamente el criterio se reduce a calcular la *distancia* del sujeto a clasificar respecto a la media o *sujeto típico* en cada especie o lo que es lo mismo, respecto al *centro de gravedad* de todos los datos correspondientes a los sujetos de la especie, que podemos expresar como:

$d(7, \overline{x}_S) = 4.9$; $d(7, \overline{x}_{Ve}) = 5.6$, $d(7, \overline{x}_{Vi}) = 12.8$

3. Sólo tenemos una muestra de 10 plantas en cada especie y la diferencia entre Virgínica y Versicolor no es muy grande. Si apareciera, por ejemplo una planta con anchura de pétalo igual a 16.5 ¿Dónde la clasificaríamos? Supongamos que las medidas de esta planta son P(16.5, 45, 27, 57) para la anchura y longitud de pétalo y anchura y longitud de sépalo. Designemos la planta como P.

Una idea sería usar el resto de las variables. Si consideramos la distancia a las especies Virgínica y Versicolor, teniendo sólo en cuenta la anchura del pétalo, obtenemos:

$d(P, \overline{x}_S) = 14.4$; $d(P, \overline{x}_{Ve}) = 3.9$, $d(P, \overline{x}_{Vi}) = 3.3$.

Parece que el mejor criterio es clasificar la planta como versicolor, pero no hay mucha diferencia con Virgínica.

4. ¿Cómo puedes extender la distancia que hemos definido anteriormente al caso de dos variables? ¿Dónde clasificarías la planta en función de las dos primeras variables?

Trataremos de tener en cuenta, en primer lugar la longitud del pétalo (Figura 5.8.2 y Tabla 5.8.3). De nuevo la diferencia es mínima. Podemos tratar de representar los dos datos para cada una de las plantas en un diagrama bidimensional y ver si se aprecia algún tipo de relación con la especie (Figura 5.8.3).

Figura 5.8.2. Diagrama de cajas de longitud de pétalo en tres especies de Iris

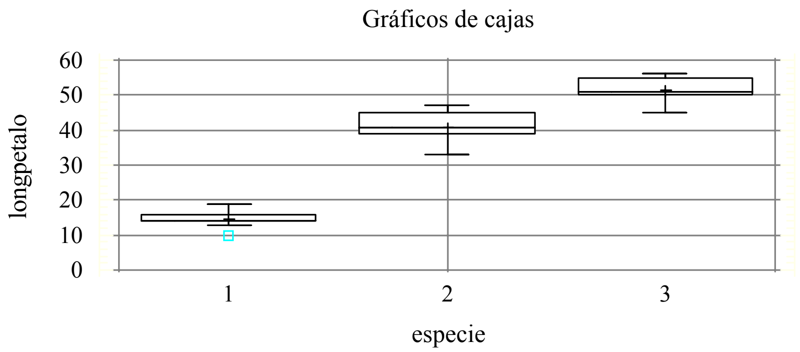
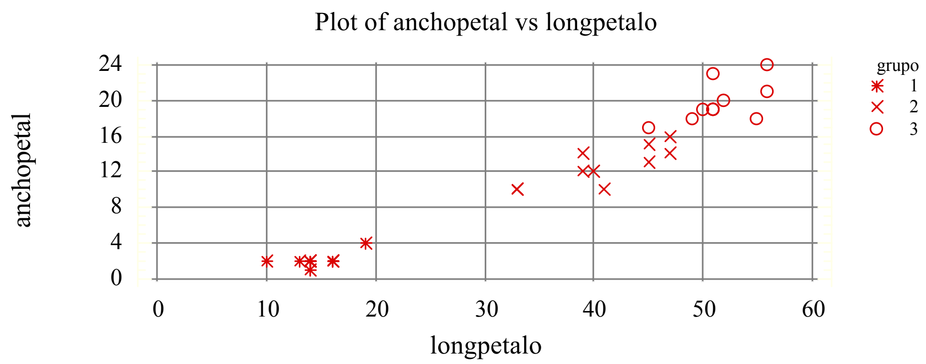


Tabla 5.8.3. Resúmenes estadísticos de la longitud de pétalo en tres especies de Iris

	Setosa (1)	Virgínica (2)	Versicolor (3)
Media	14.6	40.9	51.6
D. Típica	2.36	5.17	3,4
Mínimo	10	33	45
Máximo	19	47	56

Figura 5.8.3. Representación conjunta de longitud y ancho de pétalo en tres especies de Iris



Vemos en la gráfica que los individuos de las tres especies diferentes aparecen en regiones casi separadas del plano, de forma que podríamos dividir el plano en tres regiones por medio de dos líneas rectas y la separación sería casi perfecta. Podemos pedir a los alumnos que dibujen dos rectas que aproximadamente dividan el plano en tres regiones, de modo que la mayoría de los sujetos de cada especie estén en la misma región. Estas son las *funciones discriminantes.* Se necesitan dos funciones discriminantes para separar tres grupos.

Para clasificar la planta P podemos ver en qué región del plano queda. Alternativamente, podemos estudiar si la distancia al centro de gravedad (individuo típico) es mayor en Versicolor o Virgínica. Usaremos como distancia la suma de valores abolutos de la diferencia a las medias de las dos variables:

d (P, $\overline{x}_{Ve}$) = 3.9 + 4.1= 8, d (P, $\overline{x}_{Vi}$) = 3.3 + 6.6=9.9; la planta P se aproxima más a Versicolor que a Virgínica, pero la diferencia no es muy grande.

5. ¿Cómo puedes mejorar la clasificación usando todas las variables? ¿Dónde clasificarías la planta en función de las cuatro variables?

Podemos representar gráficamente las otras dos variables (Figura 5-8-4 y 5.8.5) para analizar cuál es la que nos conviene tomar a continuación.

Figura 5.8.4. Diagrama de cajas de ancho sépalo en tres especies de Iris

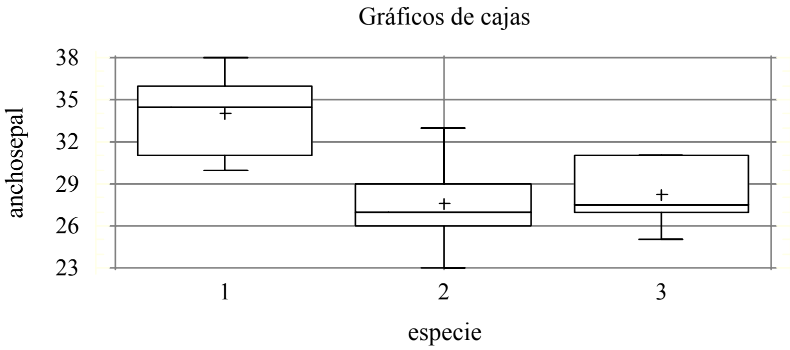
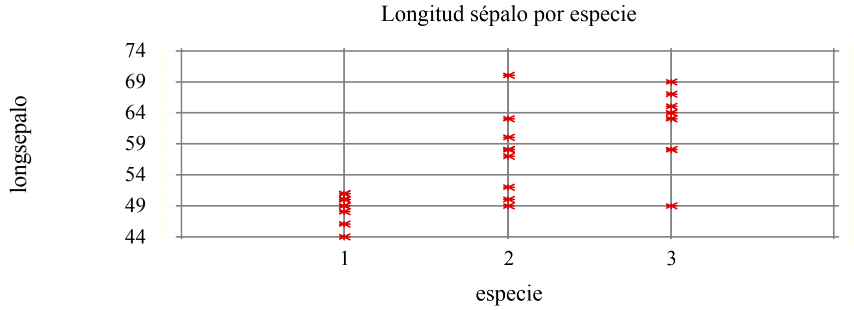


Figura 5.8.5. Diagrama de puntos de longitud sépalo en tres especies de Iris



Observamos que tampoco las diferencias son grandes. Calculando los valores medios de la anchura sépalo para cada especie $\bar{x}_{Ve}$= 27.6 y $\bar{x}_{Vir}$=28.2 y de la longitud del sépalo $\bar{x}_{Ve}$= 57.5 y $\bar{x}_{Vir}$=62, podemos extender la distancia usada para dos variables al caso de cuatro, simplemente añadiendo nuevos sumandos y tenemos:

$$d\,(P, \bar{x}_{Ve}) = 3.9 + 4.1 + 0.6 + 0.5 = 9.1$$

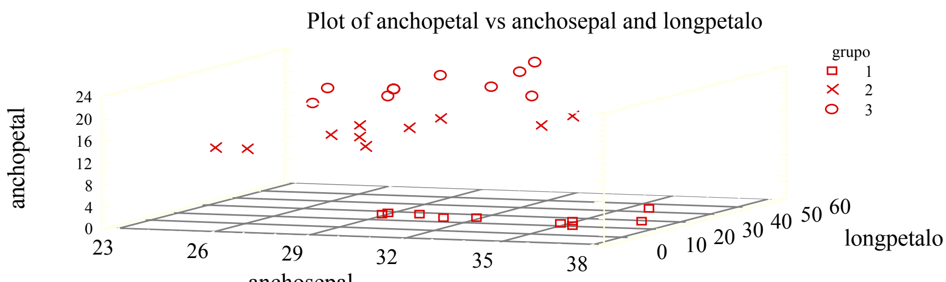
$$d\,(P, \bar{x}_{Vi}) = 3.3 + 6.6 + 1.2 + 5 = 16$$

Vemos ahora más claramente que la planta pertenede a la especie Iris Versicolor. Observamos también como al incluir nuevas variables hemos mejorado la clasificación debido a la nueva información.

Los valores de las cuatro variables para cada unas de las plantas de la muestra pueden considerarse como las coordenadas de un punto. Un conjunto de datos puede asemejarse, por tanto a un conjunto de puntos en un espacio. Cuando tenemos sólo dos variables, estamos en un plano de coordenadas (siendo las variables X e Y las dos variables), como se muestra en la Figura 5.8.3. para los datos de nuestra muestra.

pero nos lo podemos imaginar.

Figura 5.8.6. Representación conjunta de longitud y ancho de pétalo y ancho de sépalo en especies Iris



Con tres variables estaríamos en el espacio de tres dimensiones. Por ejemplo, en la Fidgura 5.8.6. hemos representado conjuntamente tres de las variables de nuestra muestra. No podemos representar un espacio de cuatro o más dimensiones, Cuando tenemos un conjunto de puntos en un espacio, puede ser que, como en las Figuras 5.8.3 y 5.8.6 aparezcan agrupaciones. Entonces podemos clasificar los puntos en grupos o categorías (en este caso las especies) usando las funciones discriminantes.

En el caso del plano, las funciones discriminantes son rectas (que es una función de una sola variable); para separar puntos en un espacio de tres dimensiones, necesitamos planos (que es una función de dos variables), para cuatro, cinco o n variables las funciones discriminantes son funciones lineales con n-1 variables. Para clasificar un individuo en un grupo, calculamos las distancias a los centros de gravedad de los diferentes grupos y lo asignamos al grupo con el que tiene menor distancia. Esta clasificación no siempre es perfecta, pero funciona en la mayoría de los casos.

Tabla 5.8.4. Distancias entre individuos en especies Iris

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.
1.																													
2.	41																												
3.	12	65																											
4.	11	26	31																										
5.	38	29	26	43																									
6.	30	65	58	13	94																								
7.	13	88	5	42	49	65																							
8.	55	114	71	42	125	13	78																						
9.	10	61	6	37	30	72	5	97																					
10.	9	32	33	2	53	9	40	34	37																				
11.	1156	1531	1052	1233	1330	1098	1015	893	1150	1195																			
12.	1454	1863	1386	1519	1714	1326	1335	1097	1490	1463	74																		
13.	1634	2105	1590	1715	1976	1502	1509	1281	1678	1649	190	54																	
14.	889	1244	801	978	1061	869	756	698	873	940	31	139	233																
15.	525	778	421	612	581	579	402	458	475	594	227	501	693	126															
16.	893	1250	805	972	1069	859	762	690	883	938	27	133	221	6	144														
17.	1246	1639	1158	1327	1458	1172	1111	959	1252	1279	14	30	114	47	305	49													
18.	507	746	403	586	553	553	390	432	461	570	233	509	717	138	2	154	315												
19.	809	1102	705	884	909	795	686	622	787	854	63	225	413	42	72	56	105	70											
20.	825	1166	745	908	997	799	702	634	815	870	39	141	237	2	120	8	53	130	40										
21.	2541	3066	2445	2642	2863	2389	2374	2074	2573	2562	351	165	191	506	1078	518	255	1098	630	530									
22.	2175	2676	2107	2278	2511	2039	2028	1758	2211	2196	289	105	99	388	918	406	185	942	530	402	30								
23.	2002	2485	1910	2097	2290	1880	1845	1605	2024	2029	166	54	90	273	735	279	100	753	383	291	37	27							
24.	1758	2195	1630	1855	1954	1684	1587	1419	1748	1803	74	86	210	171	485	181	60	495	205	193	147	153	60						
25.	1251	1580	1103	1338	1323	1239	1092	1018	1211	1308	89	265	503	132	198	150	141	194	58	146	530	508	339	125					
26.	1818	2299	1706	1937	2068	1758	1639	1499	1806	1875	106	82	132	175	543	191	66	567	271	199	113	89	34	30	225				
27.	1686	2147	1586	1791	1938	1610	1523	1361	1686	1731	78	44	94	143	505	153	38	525	237	161	117	81	26	30	217	6			
28.	2346	2865	2226	2453	2626	2230	2161	1927	2354	2385	234	132	182	377	871	383	174	891	483	407	27	63	22	66	371	50	60		
29.	1758	2195	1630	1855	1954	1684	1587	1419	1748	1803	74	86	210	171	485	181	60	495	205	193	147	153	60		125	30	30	66	
30.	2137	2630	2033	2220	2421	1995	1974	1710	2163	2158	183	73	117	322	808	312	129	822	432	344	46	66	15	69	362	63	53	19	69

5.8.4. Actividades de ampliación

Partiendo de la idea de que los valores de las variables para una unidad estadística son las coordenadas de un punto en el espacio, podemos tratar ver si somos capaces de identificar puntos cercanos y lejanos y si corresponden a sujetos de la misma o diferente especie. Para ello planteamos la siguiente actividad:

6. Considerando los valores de las cuatro variables como las coordenadas de un punto en un *espacio de cuatro dimensiones, ¿Puedes encontrar algunos individuos que se encuentren próximos? ¿que se encuentren lejos?*

Algunos ejemplos de individuos cercanos son el 1 y 4 o el 1 y 10 0 el 7 y 9, todos ellos de la especie Setosa. Ejemplos de individuos lejanos son el 4 con el 11 o el 7 con el 22 (de diferentes especies). El profesor puede dar a los chicos para ayudarse la tabla de distancias 5.8.4.

7. Las técnicas de clasificación automática clasifican los datos estadísticos, en función de la distancia entre los mismos, considerando cada fila del conjunto de datos como un punto en un espacio de n dimensiones. Como resultado producen un dendograma o gráfico en forma de árbol, como en la Figura 5.8.9. Interpreta el dendograma y comprueba con la ayuda de la tabla de distancias y del conjunto de datos que los individuos más semejantes aparecen agrupados en el dendograma.

Figura 5.8.9. Clasificación automática de individuos en especies Iris

```
Distancia entre Grupos

      C A S O         0         5        10        15        20        25
  Especie        Num  +---------+---------+---------+---------+---------+

  I. Verginica    24   ─┐
  I. Verginica    29   ─┼─┐
  I. Verginica    26   ─┤ │
  I. Verginica    27   ─┘ ├─┐
  I. Verginica    23   ─┐ │ │
  I. Verginica    30   ─┤ │ │
  I. Verginica    28   ─┼─┘ ├─────┐
  I. Verginica    21   ─┤   │     │
  I. Verginica    22   ─┘   │     │
  I. Versicolo    12   ─┬───┘     ├─────────────────────────────────────┐
  I. Versicolo    13   ─┘         │                                     │
  I. Versicolo    15   ─┬───┐     │                                     │
  I. Versicolo    18   ─┘   │     │                                     │
  I. Versicolo    14   ─┐   ├─────┘                                     │
  I. Versicolo    20   ─┤   │                                           │
  I. Versicolo    16   ─┼─┐ │                                           │
  I. Versicolo    11   ─┤ ├─┘                                           │
  I. Versicolo    17   ─┘ │                                             │
  I. Versicolo    19   ─┬─┘                                             │
  I. Verginica    25   ─┘                                               │
  I. Setosa        6   ─┬─┐                                             │
  I. Setosa        8   ─┘ │                                             │
  I. Setosa        7   ─┐ │                                             │
  I. Setosa        9   ─┤ ├─────────────────────────────────────────────┘
  I. Setosa        3   ─┤ │
  I. Setosa        4   ─┤ │
  I. Setosa       10   ─┼─┘
  I. Setosa        1   ─┤
  I. Setosa        2   ─┤
  I. Setosa        5   ─┘
```

5.8.5. Algunas dificultades y errores previsibles

En este tema tratamos de iniciar algunas ideas de análisis multivariante, lo que supone un cambio hacia la perspectiva geométrica del análisis de datos.

Se ofrece la oportunidad de aplicar la idea de distancia en un contexto no habitual. Los alumnos podrían tener dificultades para considerar cada individuo (unidad estadística) como un punto en un espacio vectorial y usar esta "metáfora" para poder analizar las diferencias/similaridades entre individuos en términos de distancia. Es decir pasar de una métrica multivariante (conjunto de valores de las variables) a otra univariante (distancia al origen o distancia entre individuos).

Puesto que el espacio vectorial considerado es de dimensión mayor que tres, debemos usar representaciones de proyecciones parciales en dos o tres dimensiones para analizar los datos.

Un último punto es realizar la cadena de interpretaciones que va de los datos al espacio de puntos, el trabajo dentro del espacio vectorial para sacar conclusiones y la interpretación de estas conclusiones respecto al conjunto de datos originales.

5.8.6. Análisis del contenido estadístico

En este proyecto podemos identificar explícita o implícitamente los siguientes contenidos:

1. Aplicaciones de la estadística:
- Botánica, taxonomía, clasificación automática.

2. Conceptos y propiedades:
- Distancia, similaridad y disimilaridad; centro de gravedad, distancia al centro de gravedad;
- Representación geométrica de datos; coordenadas; las unidades estadísticas como puntos;
- discriminación; función discriminante; clasificación; criterios de clasificación.

3. Notaciones y representaciones:
- Tablas; listado de datos;
- Gráficos de puntos; gráficos de cajas; diagramas de dispersión en dos y tres dimensiones;
- Dendograma.

4. Técnicas y procedimientos:
- Cálculo de distancias;
- Representación de datos.

6. Actitudes:
- Valoración de la estadística como herramienta de análisis de datos complejos y clasificación;
- Valoración de la estética y claridad en la elaboración de tablas y gráficos;
- Valoración del álgebra lineal y geometría como elemento modelizador en el análisis de datos.

5.8. Ideas para nuevos proyectos

5.8.1. Actitudes hacia la estadística

Utilizando el cuestionario sobre actitudes que hemos incluido en la sección XX se recogen datos de los alumnos al comenzar y finalizar la asignatura. Se trata de analizar los puntos siguientes:

- ¿Cómo eran las actitudes iniciales de los alumnos?
- ¿Qué componentes se pueden diferenciar y en cuales de ellos tenía la clase una actitud positiva/negativa?
- ¿Han cambiado las actitudes al finalizar el curso? ¿Qué componente ha cambiado más/menos? ¿En qué sentido?
- ¿Depende la actitud final de la inicial?

5.8.2.¿Existe discriminación laboral respecto a la mujer?

Los alumnos recopilarán datos de la prensa y de anuarios estadísticos que reflejen la situación laboral de hombres y mujeres en España durante el pasado año y hace 10 años. La clase se puede dividir para localizar los datos y contestar preguntas como las siguientes:

- ¿Qué proporción hay de mujeres activas?
- ¿Cuál es la proporción de paro entre la población activa femenina?
- ¿Cúal es esta proporción según nivel de educación (básica, media, superior)?
- ¿Qué proporción de mujeres universitarias ocupan cargos ejecutivos o de dirección?
- ¿Cuál es la tasa de paro femenino por comunidades autónomas?
- ¿Cómo se comparan estos datos con los correspondientes a hombres?
- ¿Cómo ha cambiado la situación en los últimos diez años?

5.8.3. España en la comunidad Europea

Los alumnos recopilarán datos de Internet, accediendo al servidor de Eurostat, la agencia estadística de la comunidad europea y recopilarán indicadores socioeconómicos de los países miembros de la comunidad europea. Se trata de describir la distribución de las diferentes variables entre los países miembros y estudiar el lugar que España ocupa en cada una de las variables. Los alumnos pueden elegir variables de su interés. Algunas sugerencias son:

- Indicadores económicos: Renta per cápita en euros; Producto Nacional Bruto, consumo, tasa de empleo y paro;
- Datos socio-demográficos: Población, densidad de población, población joven (hasta 25 años); tasa de natalidad y mortalidad; esperanza de vida;
- Transporte: Kms de autopista/ extensión; distribución del transporte en autopistas, ferrocarril y otros;
- Consumo de energía; % de energía importada;
- Turismo: número de visitantes; salidas al extranjero; ingresos por turismo
- Educación: % población escolarizada, % con estudios universitarios

5.8.4. Intención de voto en las próximas elecciones al consejo escolar

Se propone diseñar, llevar a cabo y analizar los datos de una encuesta en el centro para estudiar la intención de voto en el próximo consejo escolar, una vez que se conocen los candidatos a representantes de los alumnos. Los alumnos deben diseñar el cuestionario, seleccionar una muestra representativa de alumnos del centro, distribuir el cuestionario y analizar los datos. Algunas cuestiones relacionadas son:

- ¿Qué preguntas debemos incluir en el cuestionario? ¿Están claras las preguntas? ¿Qué variables identificativas del alumno podrían influir en su intención de voto?
- ¿Cómo elegimos la muestra de alumnos? ¿Cuál es la población objetivo? ¿Cuál es la población que podemos alcanzar?
- ¿Sería la encuesta fiable si hay un porcentaje alto de no respuesta? ¿Cómo podemos motivar la participación y disminuir la no respuesta? ¿Cómo y cuando distribuimos el cuestionario y recogemos los datos?
- ¿Cómo extendemos las conclusiones de la muestra a todo el centro? ¿Entre qué límites cabe esperar que varíe la proporción de alumnos que votarán a uno u otro candidato? ¿Cómo puedo usar el cálculo de probabilidades para poder calcular estos límites con un cierto margen de confianza?
- ¿Serán diferentes los resultados de la votación en los distintos cursos? ¿En chicos y chicas?

5.8.5. ¿Tiene ventaja el equipo que juega en su propio campo?

Los alumnos recogerán de una hemeroteca los datos referentes a todos los equipos de fútbol que han jugado en la liga del año anterior para cada semana. Estos datos se recogen habitualmente en las revistas deportivas y también podrían recogerse de Internet. Los alumnos tratarán de ver si es cierta la creencia de que el jugar en su propio campo favorece al equipo, analizando, para cada uno de los equipos y semanas los siguientes datos:

- ¿Jugó el equipo en su campo?
- Resultado del partido: ganó, perdió o empató;
- Número de goles marcados;
- Número de goles que le marcaron;
- Puntos conseguidos

5.8.6. Entrenamiento deportivo: ¿Se mejora con la práctica?

Durante la clase de gimnasia se recogen datos de cada alumno el primer día de clase y una vez transcurrido 5 meses. Podrían analizarse, entre otras las siguientes variables, para ver si la práctica ayuda a mejorar, qué alumno mejoró más globalmente y si mejoran más las chicas o los chicos:

- Tiempo en segundos para recorrer 50 metros;
- Pulsaciones por minuto antes y después de correr los 50 metros;
- Altura máxima que se puede saltar;
- Longitud máxima que se puede saltar;
- Número de abdominales seguidos hasta cansarse;
- Número de canastas encestadas en 10 intentos.

5.8.7. ¿Cuántas lentejas tiene un kilo de lentejas?

Se trata de estimar el número aproximado de lentejas en un kilo, sin tener que contarlas todas. Puesto que el proceso de llenado de un paquete de lentejas tiene un componente aleatorio, este número variará de uno a otro paquete. Se plantea así un problema de estimación que es común a otros muchos contextos, por ejemplo, cuando se estima el número medio de glóbulos rojos en sangre de individuos adultos.

Los alumnos por equipos podrían tratar de estimar el número de lentejas en paquetes seleccionados de varias marcas comerciales. Se presentaría el problema de que hay que especificar con claridad la variedad, pues existen diversos tamaños. Una vez fijada una variedad y comprados paquetes de diversas marcas cada equipo trataría de estimar el número de lentejas de su paquete.

Para ello se pueden tomar datos del número de lentejas en varias muestras de unidades de capacidad pequeñas, como el centímetro cúbico y resolver primero el problema de la estimación del número de lentejas en un cm^3. Los alumnos recogerán datos de las muestras de cm^3 representándolos gráficamente, y estudiando su distribución que será, aproximadamente normal, determinando su media y desviación típica.

Calculado el volumen de los paquetes de kilo de lenteja, para calcular la distribución del número de lentejas en un paquete de kilo, se trata de hacer un cambio de variable en una distribución normal. Por tanto, la media y desviación típicas quedarán afectadas por el cambio de escala que pasa del cm^3 al volumen del paquete.

REFERENCIAS

American Psychological Association (1994*). Publication Manual of the American Psychological Association*. Washington, DC: American Psychological Association.

Ausubel, D. P., Novak, J. D. y Hanesian, H. (1983). *Psicología educativa. Un punto de vista cognoscitivo*. México: Trillas.

Barr, G. V. (1980). Some students ideas on the median and mode. *Teaching Statistics,* 2, 38-41.

Batanero, C. (1998). Recursos para la educación estadística en Internet. *UNO,* 15, 13-25.

Batanero, C. (2000a). Cap on va l'educació estadística?. *Blaix*, 15, 2-13.

Batanero, C. (2000b). Controversies around the role of statistical tests in experimental research. *Mathematical Thinking and Learning,* 2(1&2), 75-97.

Batanero, C. (2000c). Significado y comprensión de las medidas de tendencia central. *UNO,* 25, 41-58

Batanero, C., Estepa, A., Godino, J. D.y Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education,* 27(2), 151-169.

Batanero, C, Godino, J. D., y Navas, F. (1997). Some misconceptions about averaes in prospective primary teachers. En E. Pehknonen (Ed.), *Proceedings of 21 PME Conference* (v.1, pp. 276). University of Lahti.

Batanero, C. y Serrano, L.(1995). Aleatoriedad, sus significados e implicaciones educativas. *UNO*, 5, 15-28.

Begg, A. (1997). Some emerging influences underpining assessment in statistics. En I. Gal y J. Garfield (Eds.), *The assessment challenge in statistics education.* Amsterdam: IOS Press.

Biehler, R. (1997). Software for learning and for doing statistics. *International Statistical Review,* 65(2), 167-190.

Birnbaum, I. (1982). Interpreting statistical significance*. Teaching Statistics, 4,* 24–27.

Black, M. (1979). *Inducción y probabilidad.* Madrid: Cátedra.

Bollen K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Borassi, R. (1987). Exploring mathematics through the analysis of errors. *For the learning of mathematics,* 7(3), 2-8.

Bright, G., Curcio, F. y Friel, S. (En prensa). Making sense of graphs: Critical factors than influence coomprehension and instructional implications. *Journal for Research in Mathematics Education.*

Brousseau, G. (1983). Les obstacles epistémologiques et les problémes en mathématiques. *Recherches en Didactique des Mathematiques,* 4(2), 164-198.

Burks, A. W. (1977). *Chance, cause, reason: An inquiry into the nature of scientific evidence*. Chicago: University of Chicago Press.

Cabriá, S. (1994). *Filosofía de la estadística.* Servicio de Publicaciones de la Universidad de Valencia.

Cai, J. (1995). Beyod the computational algorithm: Students' understanding of the arithmetic average concept. En L. Meira (Ed.). *Proceedings of the 19th PME Conference* (v.3, pp. 145-151). Universidade Federal de Pernambuco, Recife, Brasil.

Cañizares, M. J. (1997). *Influencia del razonamiento proporcional y combinatorio y de creencias subjetivas en las intuiciones probabilísticas primarias.* . Departamento de Didáctica de la matemática. Universidad de Granada

Carvalho. C. (1998). Tarefas estatísticas e estratégias de resposta. Comunicación presentada en el *VI Encuentro en Educación Matemática de la Sociedad Portuguesa de Ciencias de la Educación.* Castelo de Vide, Portugal.

Centeno, L. (1988). *Números decimales. ¿Por qué? ¿Para qué?.* Madrid: Síntesis.

Corbalán, F. (1994). *Juegos matemáticos para secundaria y Bachillerato*. Madrid: Síntesis.

Chance, B. L., Garfield, J. B. y del Mas R. C. y (1999). The role of assessment in research on teaching and learning statistics. Trabajo presentado en la *AERA Annual Meeting.* Montreal

Chatfield C. (1988). *Problem solving: a statistician's guide. London:* Chapman & Hall.

Chapman, L. J. y Chapman, J. P. (1967). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology,* 74, 271-280.

Chow, L. S. (1996). *Statistical significance: Rationale, validity and utility*. London: Sage.

Cobo, B. y Batanero, C. (2000). La mediana en la educación secundaria ¿Un concepto sencillo? *UNO,* 24.

Cohen, S. y Chechile, R. (1997). Overview of ConStatS and the ConStats assessment. En Garfield, J. B. y Burrill, G. (Eds.). *Proceedings of the IASE 1996 Round Table Conference*. (pp. 109 – 118). Voorburg: International Statistical Institute.

Confrey, J. (1990). A review of the research in students conceptions in mathematics, science and programming. *Review of Research in Education,* 16, 3-35.

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education,* 18 (5), 382-393.

Curcio, F. R. (1989). *Developing graph comprehension.* Reston, VA: N.C.T.M.

Dahl, H. (1999). Teaching hypothesis testing. Can it still be useful? *Bulletin of the International Statistical Institute: Proceedings of the Fifty-second Session of the International Statistical Institute* (Tome 58, Book 2) (pp. 197-200). Helsinki, Finland: International Statistical Institute.

Davis (1985). *The loogic of causal order.* Londres: Sage.

delMas, R. C., Garfield, J. B.y Chance, B. (1999). *Exploring the role of computer simulations in developing understanding of sampling distributions.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Davis P. J. y Hersch, R. (1982). *The mathematical experience.* Boston: Birkhäuser. [Ed. castellana, MEC-Labor, Madrid. 1989].

Diaconis, P., & Freedman, D. (1981). The persistence of cognitive illusions. *Behavioral and Brain Sciences, 4*, 378-399.

Eisenbach, R. (1994). What does de mean mean? Comunicación presentada en el *Fourth International Conference on Teaching Statistics*. Marrakesh, Marruecos.

Ellerton, N. (1996). Statistical significance testing and this journal. *Mathematics Education Research Journal, 8*(2), 97–100.

Ernest, P. (1991). *The philosophy of mathematics education.* London: The Falmer Press.

Escofier B. y Pages J. (1988). *Analyses factorielles simples et multiples: objectifs, méthodes et interpretation*. Paris: Dunod.

Estepa, A. (1990). *Enseñanza de la Estadistica basada en el uso de ordenadores: Un Estudio exploratorio*. Memoria de Tercer Ciclo. Universidad de Granada: Departamento de Didactica de la Matematica.

Estepa A. (1994). *Concepciones sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores*. Tesis doctoral Departamento de Didáctica de la Matemática. Universidad de Granada.

Estepa, A. (1995). Las tablas de contingencia y su enseñanza. ¿Qué podemos aprender de las investigaciones realizadas? *UNO*, 3. 89-100.

Estepa, A. (1995). Consideraciones sobre la enseñanza de la asociación estadística. *UNO, 5, 69-79.*

Estepa, A. y Batanero, C. (1994). Estrategias en los juicios de asociación en tablas de contingencia. *VII Jornadas de la Sociedad Andaluza de Profesores de Matemáticas Thales.* (pp. 417-423). Sevilla: S.A.E.M.Thales.

Estepa, A.y Batanero, C. (1995). Concepciones iniciales sobre la asociación estadística. *Enseñanza de las Ciencias,* 13(2), 155-170.

Estepa, A., Batanero, C. y Sánchez, F. T. (1999). Judgments of association in the comparison of two samples: students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education,* 7, 17-30.

Estepa, A. y Batanero, C. (1996). Judgments of correlation in scatter plots: students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education,* 4, 21-41.

Falk, R. (1981). The perception of randomness. En C. Comiti y G. Vergnaud (Eds.), *Proceedings of the V International Conference for the Psychology of Mathematics Education* (pp. 222-229). Universidad de Grénoble.

Falk, R. (1986). Misconceptions of statistical significance. *Journal of Structural Learning,* 9, 83-96.

Falk, R. y Greenbaum, C. W. (1995). Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory and Psychology, 5* (1), 75-98.

Fischbein (1975). *The intuitive sources of probabilistic thinking in children.* Dordrecht: Reidel.

Fisher, R. A. (1935). *The design of experiments*. Edimburgh: Oliver & Boyd.

Fisher, R. A. (1956). *Statistical methods and scientific inference*. Edinburgh: Oliver & Boyd.

Fisher, R. A. (1958). *Statistical methods for research workers* (13 edición). New York: Hafner.

Freudenthal H. (1991), *Revisiting mathematics education*. Dordrecht, Kluwer AC.

Gal y J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 17-26).

Galmacci, G. (1996). Statistics in the era of networking. *Annual Report on International Statistics,* 3, 1-7.

Garfield, J.B. (1995). La evaluación del aprendizaje de la estadística. *UNO*, 5, 5-14.

Garfield, J. y Alhgren, A.. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education,* 19 (1), 44-63.

Garfield, J. y Burrill, G. (1997). *Training Researchers in the Use of Statistics.* IASE.

Gattuso, L. y Mary, C. (1998). Development of the concept of weighted average among high-school students. En L. Pereira-Mendoza, L. Seu Keu, T. Wee Kee y W. K. Wong, *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 685-691). Singapur: International Association for Statistical Education.

Gerber, R., Boulton-Lewis, G. y Bruce, C. (1995). Children's understanding of graphic representation of quantitative data. *Learning and Instruction,* 5, 70-100.

Gingerenzer, G. (1993). The superego, the ego and the id in statistical reasoning. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methdological issues* (pp. 311-339). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gingerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J. y Kruger, L. (1989). *The empire of chance. How probability changed science and everyday life.* Cambridge: Cambridge University Press.

Godino, J. D. (1996). Significado y comprensión de los objetos matemáticos. En L. Puig y A. Gutiérrez (Eds.). *Proceedings of the 20 th PME Conference,* 2, 417 – 424. Valencia.

Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, 14 (3), 325 – 355.

Godino, J. D. y Batanero, C. (1997) Clarifying the meaning of mathematical objects as a priority area of research in Mathematics Education. En: A, Sierpinska (Ed). *Mathematics Education as a Research Domain: A Search for Identity* (pp. 177-195)..Dordrecht: Kluwer.

Godino, J.D., Batanero, C. y Cañizares, M. J. (1987). *Azar y probabilidad Fundamentos didácticos y propuestas curriculares.* Madrid: Síntesis.

Goodchild (1988). School pupils' understanding of average. *Teaching Statistics,* 10, 77-81.

Green, D. R. (1983 a). School pupils' probability concepts. *Teaching Statistics*, 5 (2), 34-42.

Green, D. R. (1983 b). A Survey of probabilistic concepts in 3000 pupils aged 11-16 years. En D. R. Grey et al. (Eds,. *Proceedings of the First International Conference on Teaching Statistics* (v.2, p.766 - 783). University of Sheffield.

Green, D. R. (1991). A longitudinal study of children's probability concepts. En D.Vere Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics (pp. 320 - 328)*. Dunedin: University of Otago

Hacking, I. (1975). *The logic of statistical inference*. Cambridge: Cambridge University Press.

Harlow, L. L. (1997). Significance testing: Introduction and overview. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds*.), What if there were no significance tests?* (pp. 1-20). Mahwah, NJ: Lawrence Erlbaum Associates.

Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (1997). *What if there were no significance tests?* Mahwah, NJ: Lawrence Erlbaum Associates.

Hawkins, A. (1997). How far have we come? Do we know where we are going? En E. M. Tiit (Ed.), *Computational statistics & statistical education* (pp. 100-122). Tartu: International Association for Statistical Education e International Association for Statistical Computing.

Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics*, 6, 187-205.

Holmes, P. (1980). Teaching Statistics 11 -16. Sloug: Foulsham Educational.

Inhelder, B. y Piaget, J. (1955). *De la logique de l'enfant a la logique de l'adolescent* (Paris: P.U.F.).

Ito, P. K. (1999, August). *Reaction to invited papers on statistical education and the significance tests controversy*. Invited paper presented at the Fifty-second International Statistical Institute Session, Helsinki, Finland.

Kahneman, D.; Slovic, P. y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. (New York: Cambridge University Press).

Kilpatrick, J. (1994). Historia de la investigación en educación matemática. En J. Kilpatrick, L. Rico y M. Sierra (Eds.), *Educación matemática e investigación* (pp. 17-98). Madrid: Síntesis.

Konold, C. (1991). Understanding students' beliefs about probability. En E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 139-156). Dordrecht: Kluwer.

Konold, C., Lohmeier, J., Pollatsek, A., & Well, A. (1993). Inconsistencies in students' reasoning about probability. *Journal for Research in Mathematics Education,* 24, 392-414.

Kyburg, H. E. (1974). *The logical foundations of statistical Inference.* Boston: Reidel.

Lecoutre, B. (1999). Beyond the significance test controversy: Prime time for Bayes? *Bulletin of the International Statistical Institute: Proceedings of the Fifty-second Session of the International Statistical Institute* (Tome 58, Book 2) (pp. 205-208). Helsinki, Finland: International Statistical Institute.

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics*, 23, 557 - 568.

Leon, M. R., y Zawokeswski, J. S. (1991). Use of the arithmetic mean: An investigation of four properties. Issues and preliminary resuls. En D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 302-306). Voorburg, Holanda: International Statistical Institute.

Levin, J. R. (1998 a). To test or not to test H_0*? Educational and Psychological Measurement, 58,* 313-333.

Levin, J. R. (1998 b). What if there were no more bickering about statistical significance tests? *Research in the Schools, 2,* 45-53.

Li, K. Y. y Shen, S. M. (1992). Students' weaknesses in statistical projects. *Teaching Statistics*, 14 (1), 2-8.

Lindley, D. V. (1993). The analysis of experimental data: The appreciation of tea and wine. *Teaching Statistics, 15*, 22-25.

Loosen, F.; Lioen, M. y Lacante, M. (1985). The standard deviation: some drawbacks of an intuitive approach. *Teaching Statistics,* 7 (1), 2-5.

Méndez, H. (1991). *Understanding the Central Limit Theorem*. Tesis Doctoral. Universidad de California. UMI 6369

Menon, R. (1993). Statistical significance testing should be discontinued in mathematics education research. *Mathematics Education Research Journal, 5*(1), 4–18.

Mevarech, Z. R. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics,* 14, 415-429.

Mokros, J. y Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education,* 26, 20-39.

Moore, D. S. (1991). Teaching Statistics as a respectable subject. En F. Gordon y S. Gordon (eds.), *Statistics for the Twenty-First Century,* (pp. 14-25). Mathematical Association of America.

Moore, D. S. (1995). *The basic practice of statistics*. New York: Freeman.

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, 65(2), 123–155.

Morrison, D. E. y Henkel, R. E. (1970). *The significance test controversy.* Chicago: Aldine.

Moses, L. E. (1992). The reasoning of statistical inference. In D. C. Hoaglin & D. S. Moore (Eds.), *Perspectives on contemporary statistics* (pp. 107-122). Washington, DC: Mathematical Association of America.

N. C. T. M. (1991). Estándares curriculares y de evaluación para la educación matemática. Sevilla: Sociedad Thales (traducción española del original publicado en 1989 por la NCTM).

Neyman, J. (1950). *First course in probability and statistics*. New York: Henry Holt.

Nisbett, R. E. y Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, New Jersey: Prentice-Hall.

Oakes, M. (1986). *Statistical inference: A commentary for the social and behavioural sciences*. Chichester, England: Wiley.

Ojeda, A. M. (1995). Dificultades del alumnado respecto a la probabilidad condicional. *UNO,* 5. 37-44.

Ojeda, A. M. (1996). Contextos, representaciones y la idea de probabilidad condicional. En F. Hitt (Ed.), *Investigaciones en matemáticas educativas* (pp.291-310). México: Grupo Editorial Iberoamericano.

Ortiz de Haro, J. J. (1999). *Significado de conceptos probabilísticos en los textos de Bachillerato.* Tesis Doctoral. Departamento de Didáctica de la Matemática. Universidad de Granada.

Pereira-Mendoza, L. S. Kea, T. W. Kee y W. K. Wong (1998) (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics*. Singapore: International Statistical Institute.

Pérez Echeverría, M. P. (1990). *Psicología del razonamiento probabilístico* . Madir: ICE de la Universidad Autónoma.

Piaget, J., e Inhelder, B. (1951). *La genése de l'idée de hasard chez l'enfant.* Paris: Presses Universitaires de France.

Plackett, R.L. (1970), The principle of the arithmetic mean. En E. S. Pearson y M. Kendall (Eds), *Studies in the history of statistics and probability* (v,1, pp. 121-126). London, Charles Griffin.

Pollard, P. y Richardson, J. T. E. (1987). On the probability of making Type I errors. *Psychological Bulletin,* 10, 159-163.

Pollasek, A., Lina, S. y Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics,* 12, 191-204.

Poincaré, H. (1936) El Azar. Artículo publicado originalmente en lengua inglesa en *Journal of the Atnerican Statistical Association,* 31, 10-30. Recogido en J. Newman (Ed.), *Sigma. El mundo de las Matemáticas,* 3, 68-82.

Poya, G. (1982). *Cómo plantear y resolver problemas*. México: Trillas.

Popper, K. R. (1967). *La lógica de la investigación científica.* Madrid: Tecnos.

Radatz, H. C. (1980). Students' errors in the mathematical learning: a survey. *For the learning of mathematics*, 1(1), 16-20.

Reading, C. y Pegg, J. (1996). Exploring understanding of data reduction. En L. Puig y A. Gutiérrez (Eds.), *Proceedings of the 20th Conference of the International group for the Psychology of Mathematics Education* (v.4, pp. 187-194). Universidad de Valencia, España.

Reeves, C. A. y Brewer, J. K. (1980). Hypothesis testing and proof by contradiction: An analogy. *Teaching Statistics,* 2 (2), 57-59.

Rico, L. (1991). La comunidad de educadores matemáticos. En A. Gutiérrez (Ed.), *Area de conocimiento: Didáctica de la matemática* (pp. 11-58). Madrid: Síntesis.

Rico, L. (1992). *Investigaciones sobre errores de aprendizaje en educación matemática*. Departamento de Didáctica de las Matemáticas. Universidad de Granada.

Rico, L. (1993). Errores y dificultades en el aprendizaje de las matemáticas. En P. Gómez, J. Kilpatrick y L. Rico (Eds.), *Educación matemática*. Bogotá: Grupo Editorial Iberoamericano.

Rico, L. y Sierra, M. (1994). Educación matemática en la España del siglo XX. En J. Kilpatrick, L. Rico y M. Sierra (Eds.), *Educación matemática e investigación* (pp. 99-202). Madrid: Síntesis.

Rico, L. (1997). Dimensiones y componentes de la noción de currículo. En L. Rico (Ed.), *Bases teóricas del currículo de matemática en educación secundaria.* Madrid: Síntesis.

Rivadulla, A. (1991). *Probabilidad e Inferencia cientifica* (Barcelona: Anthropos).

Robinson, D. H., & Levin, J. T. (1997). Reflections on statistical and substantive significance, with a slice of replication*. Educational Researcher, 26*(5), 21-26.

Royal, R. (1997). *Statistical evidence. A likelihood paradigm*. London: Chapman & Hall.

Rubin, A., Bruce, B., & Tenney, Y. (1991). Learning about sampling: Trouble at the core of statistics. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 314–319). Voorburg, The Netherlands: International Statistical Institute.

Rubin, A. y Rosebery, A. S. (1990). Teachers' misunderstandings in statistical reasoning; evidence from a field test of innovative materials. In A. Hawkins (Ed.) *Training teachers to teach Statistics*. (Voorburg, The Netherlands: ISI), 72-89.

Russell, S. J. y Mokros, J. R. (1991). What's Typical?: Children's Ideas about Average. In: D. Vere-Jones (Ed.) *Proceedings of the Third International Conference on Teaching Statistics* (Voorburg, The Netherlands: International Statistical Institute), 307-313.

Sánchez, E. (1996). Dificultades en la comprensión del concepto de eventos independientes. In F. Hitt (ed.), *Investigaciones en Mathemática Educativa,* pp. 389–404. México: Grupo Editorial Iberoamericano.

Sánchez-Cobo, F.T. (1996). *Análisis de la exposición teórica y de los ejercicios de correlación y regresión en los textos de Bachillerato.* Memoria de Tercer Ciclo, Universidad de Granada.

Seidenfeld, T. (1979). *Philosophical problems of staistical inference: Learning from R. A. Fisher.* Dordrecht, The Netherlands: Reidel.

Serrano, L., Batanero C. y Ortiz, J. J (1996). Interpretación de enunciados de probabilidad en términos frecuenciales por alumnos de Bachillerato. *SUMA*, 22, 43-50.

Serrano, L., Batanero, C., Ortiz, J.J.y Cañizares, M. J. (In press). Heurísticas y sesgos en el razonamiento estocástico de los estudiantes de secundaria. *Educación Matemática.*

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grows (Ed.), *Handbook of research in mathematics education.* MacMillan.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. En A. Bishop et al. (Eds.), *International handbook of mathematics education* (v.1, pp. 205-237). Dordrecht: Kluwer, A. P.

Stenhouse, L. (1984). *Investigación y desarrollo del currículo.* Madrid: Morata.

Strauss, S. y Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education,* 19 (1), 64-80.

Tanur, J. M. (Ed.) (1989). *La Estadística; una guía de lo desconocido.* Madrid: Alianza Editorial.

Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher, 25*(2), 26–30.

Toohey, P. G. (1995). *Adolescent perceptions of the concept of randomness.* Master Thesis. The University of Adelaide.

Tormo, C. (1995). Dificultades del alumnado respecto a la media aritmética. *UNO,* 5, 29-36.

Truran, K. M., & Truran, J. M. (1994). *Chance & data and the Profiles in Primary Schools.* In C. Beesy, & D. Rasmussen (eds), *Mathematics Without Limits,* pp. 126-131. Melbourne, Victoria: Mathematical Association of Victoria.

Vallecillos, A. (1994). *Estudio teórico - experimental de errores y concepciones sobre el contraste de hipótesis en estudiantes universitarios.* Tesis doctoral Departamento de Didáctica de la Matemática. Universidad de Granada.

Vallecillos, A. (1995). Consideraciones epistemológicas sobre la inferencia estadística: implicaciones para la práctica docente. *UNO,* 5, 80-90.

Vallecillos, A. (1996). *Inferencia estadística y enseñanza: un análisis didáctico del contraste de hipótesis*. Granada: Comares. ISBN: 84-8151-2443

Vallecillos, A. (1996). Comprensión de la lógica del contraste de hipótesis en estudiantes universitarios. *Recherches en Didactique des Mathematiques,* 15(3), 53-81.

Vallecillos, A. (1999). Some empirical evidence on learning difficulties about testing hypotheses. *Bulletin of the International Statistical Institute: Proceedings of the Fifty-second Session of the International Statistical Institute* (Tome 58, Book 2) (pp. 201-204). Helsinki, Finland: International Statistical Institute.

Vallecillos, A.y Batanero, C. (1997). Aprendizaje y enseñanza del contraste de hipótesis: Concepciones y errores. *Enseñanza de las Ciencias*.

Vallecillos, A.y Batanero, C. (1997). Conceptos activados en el contraste de hipótesis estadísticas y su comprensión por estudiantes universitarios. *Recherches en Didactique des Mathématiques,* 17(1), 29-48.

Vallecillos, A., Batanero, C.y Serrano, L. (1994). Nivel de significación en un test de hipótesis: un análisis de su interpretación por estudiantes universitarios. *Actas de las V Jornadas Nacionales de Enseñanza y Aprendizaje de las Matemáticas* (pp. 509-515). Badajoz: Sociedad Extremeña de Profesores de Matemáticas.

Vere-Jones, D. (1997). The coming age of statistical education. International Statistical Review, 63(1), 3-23.

Wainer, H. (1992). Understanding graphs and tables. *Educational Researher,*21, 14-23.

Watson, J. y Baxter, J. P. (1997). Learning the unlikely at distance as an information technology enterprise. En J. Garfield y G. Burrill (Eds). *Research on the Role of Technology in Teaching and learning Statistics. IASE 1996 Roundtable Conference* (pp. 285-300). International Association for Statistical Education.

Watson, J. M. y Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning,* 2(1&2), 11-50.

Well, A. D., Pollatsek, A., & Boyce, S. J. (1990). Understanding the effects of the sample size on the variability of the means. *Organizational Behavior and Human Decision Processes,* 47, 289–312.

White, A. L. (1980). Avoiding errors in educational research. In R. J. Shumway (Ed.): *Research in Mathematics Education* (Reston, Va: N.C.T.M.), 49-65.

Wilensky, U. (1991). Abstract mediations on the concrete and concrete implications for mathematics education. En Harel y Papert (Eds.), *Constructionism.* Norwood: Ablex Publishing Corp.

Wilensky, U. (1995). Learning probability through building computational models. En D. Carraher y L. Meira (Eds.), *Proceedings of the 19th PME Conference* (V. 3, pp. 152-159). Recife, Brazil: Organising Committee.

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. E*ducational Studies in Mathematics*, 33, 171-202.

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604.

Wood, G. R. (1998). Transforming first year university statistics teaching. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics* (Volume 1) (pp. 167-172). Singapore: International Statistical Institute.

Yates, F. (1951). The influence of "Statistical methods for research workers" on the development of the science of statistics. *Journal of the American Statistical Association, 46,* 19-34.

Zabell, S. L (1 992) Randomness and statistical applications. En F. Gordon and S. Gordon (Eds,), *Statistics for the XX1 Century.* The Mathematical Association of America.