2. DATOS Y SINTESIS DE DATOS

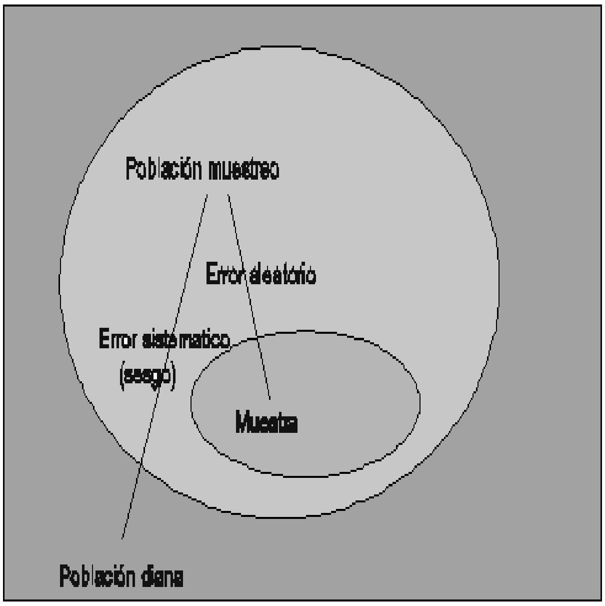
2.1. RECOPILACIÓN DE DATOS

La investigación cuya finalidad es: el análisis o experimentación de situaciones para el descubrimiento de nuevos hechos, la revisión o establecimiento de teorías y las aplicaciones prácticas de las mismas, se basa en los principios de Observación y Razonamiento y necesita en su carácter científico el análisis técnico de Datos para obtener de ellos información confiable y oportuna. Este análisis de Datos requiere de la Estadística como una de sus principales herramientas, por lo que los investigadores de profesión y las personas que de una y otra forma la realizan requieren además de los conocimientos especializados en su campo de actividades, del manejo eficiente de los conceptos, técnicas y procedimientos estadísticos

La estadística es la ciencia que estudia los métodos que permiten realizar este proceso para variables aleatorias. Estos métodos permiten resumir datos y acotar el papel de la casualidad (azar). Se divide en dos áreas:

estadistica descriptiva La descripción completa de una variable aleatoria está dada por su función densidad de probabilidad *(fdp)*. Afortunadamente una gran cantidad de variables de muy diversos campos están adecuadamente descritas por unas pocas familias de *fdps*: binomial, Poisson, normal, gamma, etc. Dentro de cada familia, cada *fdp* está caracterizada por unos pocos parámetros, típicamente dos: media y varianza. Por tanto la descripción de una variable indicará la familia a que pertenece la fdp y los parámetros correspondientes.

estadistica inferencial. Los dos tipos de problemas que resuelven las técnicas estadísticas son: estimación y contraste de hipótesis. En ambos casos se trata de generalizar la información obtenida en una muestra a una población. Estas técnicas exigen que la muestra sea aleatoria. En la práctica rara vez se dispone de muestras aleatorias, por la tanto la situación habitual es la que se esquematiza en la figura

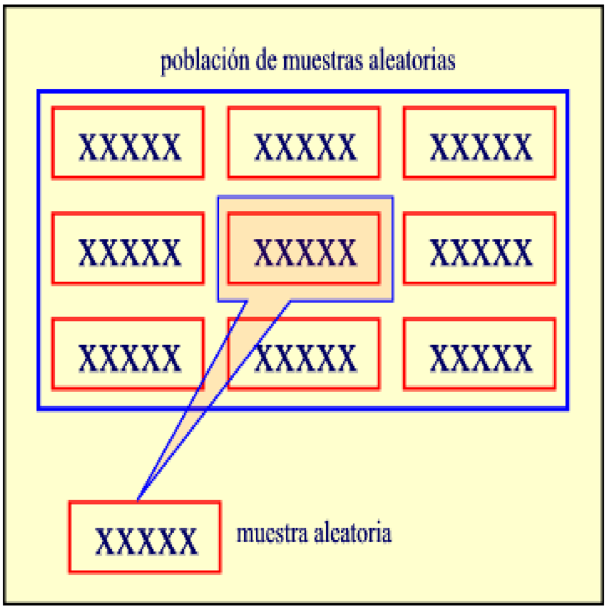


Entre la muestra con la que se trabaja y la población de interés, o población diana, aparece la denominada *población de muestreo*: población (la mayor parte de las veces no definida con precisión) de la cual nuestra muestra es una muestra aleatoria. En consecuencia la generalización está amenazada por dos posibles tipos de errores: *error aleatorio* que es el que las técnicas estadísticas permiten cuantificar y críticamente dependiente del tamaño muestral, pero también de la variabilidad de la variable a estudiar y el *error sistemático* que tiene que ver con la diferencia entre la población de muestreo y la población diana y que sólo puede ser controlado por el diseño del estudio.

RECOPILACION DE DATOS

Poblacion y muestra Al recoger datos relativos a las características de un grupo de individuos u objetos, sean alturas y pesos de estudiantes de una universidad o tuercas defectuosas producidas en una fábrica, suele ser imposible o nada práctico observar todo el grupo, en especial si es muy grande. En vez de examinar el grupo entero, llamado población o universo, se examina una pequeña parte del grupo, llamada muestra. Una población puede ser finita o infinita. Por ejemplo, la población consistente en todas las tuercas producidas por una fábrica un cierto día es finita, mientras que la determinada por todos los posibles resultados (caras, cruces) de sucesivas tiradas de una moneda, es infinita. Si una muestra es representativa de una población, es posible inferir importantes conclusiones sobre las poblaciones a partir del análisis de la muestra. La fase de la estadística que trata con las condiciones bajo las cuales tal diferencia es válida se llama estadística inductiva o inferencia estadística. Ya que dicha inferencia no es del todo exacta, el lenguaje de las probabilidades aparecerá al establecer nuestras conclusiones. La parte de la estadística que sólo se ocupa de describir y analizar un grupo dado, sin sacar conclusiones sobre un grupo mayor, se llama estadística descriptiva o deductiva.

Muestra aleatoria: muestra elegida independientemente de todas las demás, con la misma probabilidad que cualquier otra y cuyos elementos están elegidos independientemente unos de otros y con la misma probabilidad.



ANÁLISIS DESCRIPTIVO DE LA INFORMACION :

ayudará a observar el comportamiento de la muestra en estudio, a través de tablas, gráficos..**.**. Los resultados recogidos en la muestra se resumen en una matriz de datos **N x M** , en la cual **N** es el número de unidades de análisis utilizadas (número de casos) y **M** es el número de características de dichas unidades , unidades de las que tenemos información. **Transformación de los datos :** la transformación persigue la consecución de una distribución aproximada a la normal. Tipos de transformación :

Lineales : suma , resta , división , multiplicación , cambia los valores brutos (datos obtenidos) de la variable sin alterar nada mas.

No lineales monotónicas : cambian los valores originales y tambien sus distancias pero no el orden

No lineales no monotónicas : similar a la anterior pero no altera el orden..

DATOS NO AGRUPADOS

Tendencia central: la tendencia central se refiere al punto medio de una distribución. Las medidas de tendencia central se conocen como medidas de posición.

Dispersión: se refiere a la extensión de los datos en una distribución, es decir, al grado en que las observaciones se distribuyen.

DATOS AGRUPADOS

Medidas de Dispersión Se llaman medidas de dispersión aquellas que permiten retratar la distancia de los valores de la variable a un cierto valor central, o que permiten identificar la concentración de los datos en un cierto sector del recorrido de la variable. Se trata de coeficiente para variables cuantitativas.

Medidas de Tendencia central La estadística busca entre otras cosas, describir las características típicas de conjuntos de datos y, como hay varias formas de hacerlo, existen y se utilizan varios tipos de promedios. Se les llama medidas de tendencia central porque general mente la acumulación más alta de datos se encuentra en los valores intermedios.

Las medidas de tendencia central comúnmente empleadas son :

- Media aritmética
- Mediana
- Moda
- Media geométrica
- Media armónica
- Los cuantiaos

GRAFICOS DE ESTADISTICA DESCRIPTIVA

Los gráficos se han de explicar enteramente por sí mismos. El contenido de un gráfico deberá ser tan completo como sea posible. Las escalas vertical y horizontal estarán rotuladas con claridad dando las unidades pertinentes. La mayorías de los gráficos presentan información numérica con escalas, que deben rotularse para describir completamente la variable presentada en la escala y para variables de medida se dirán las unidades de medición. No se debe tratar de abarcar demasiada información en un solo gráfico. Es mejor hacer varios gráficos que comprimir toda la información en uno solo. Una regla práctica segura es evitar gráficos que contengan más de 3 curvas. Los gráficos tienen que dar una visión general y no una imagen detallada de un conjunto de datos. Las presentaciones detalladas se deben reservar para las tablas. Las tablas se explicarán por sí mismas enteramente, como los gráficos, se ha de dar suficiente información en el título y en los encabezamientos de columnas y filas de la tabla para permitir que el lector identifique fácilmente su contenido. Como el título será por lo general lo primero que se lee en detalle, deberá suministrar toda la información esencial sobre el contenido de la tabla y deberá especificar el tiempo, lugar, material ó estudio experimental y relaciones que se presenten en la tabla.

Para cada variable numérica se han de dar las unidades. La función del rayado es dar claridad de interpretación. Las anotaciones de numéricas del cero se han de escribir explícitamente. Una anotación numérica no debe comenzar con una punto decimal. Los números que indican valores de la misma característica se han de dar con el mismo número de decimales.

PARA LOS SUIGUITES DATOS VAN A SER VASADOS EN LA TABLA 1(TANTO AGRUPADOS COMO NO AGRUPADOS)

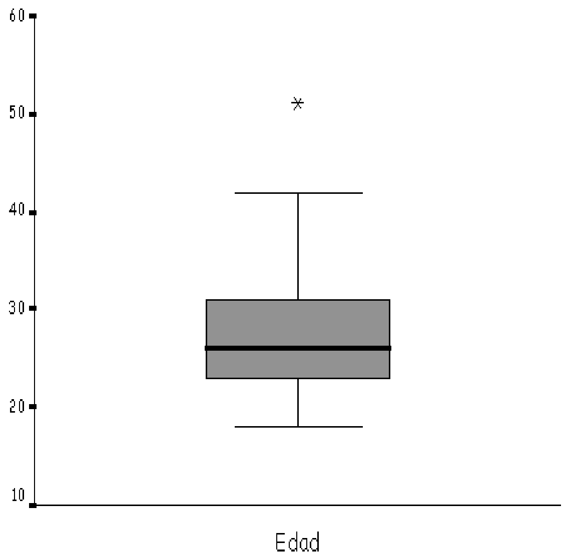
Tabla I. Distribución de frecuencias de la edad en 100 pacientes.

Edad	Nº de pacientes
18	1
19	3
20	4
21	7
22	5
23	8
24	10
25	8
26	9
27	6
28	6
29	4
30	3
31	4
32	5
33	3
34	2
35	3
36	1
37	2
38	3
39	1
41	1

42	1

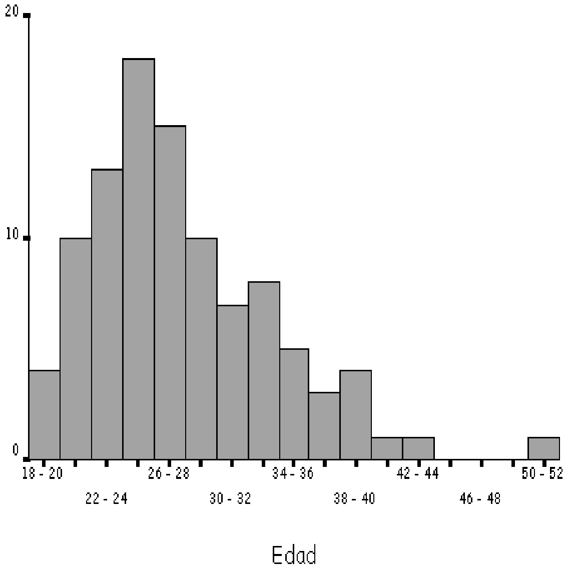
NO AGRUPADOS

Otro modo habitual, y muy útil, de resumir una variable de tipo numérico es utilizando el concepto de percentiles, mediante diagramas de cajas. La Figura muestra un gráfico de cajas correspondiente a los datos de la Tabla I. La caja central indica el rango en el que se concentra el 50% central de los datos. Sus extremos son, por lo tanto, el 1er y 3er cuartil de la distribución. La línea central en la caja es la mediana. De este modo, si la variable es simétrica, dicha línea se encontrará en el centro de la caja. Los extremos de los "bigotes" que salen de la caja son los valores que delimitan el 95% central de los datos, aunque en ocasiones coinciden con los valores extremos de la distribución. Se suelen también representar aquellas observaciones que caen fuera de este rango (outliers o valores extremos). Esto resulta especialmente útil para comprobar, gráficamente, posibles errores en nuestros datos. En general, los diagramas de cajas resultan más apropiados para representar variables que presenten una gran desviación de la distribución normal.



DATOS AGRUPADOS

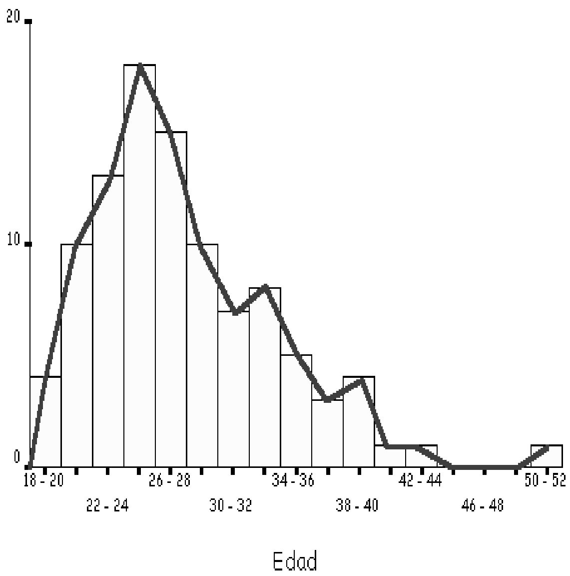
Histograma: Esta formado por rectángulos cuya base es la amplitud del intervalo y tiene la característica que la superficie que corresponde a las barras es representativa de la cantidad de casos o frecuencia de cada tramo de valores, puede construirse con clases que tienen el mismo tamaño o diferente (intervalo variable). La utilización de los intervalos de amplitud variable se recomienda cuando en alguno de los intervalos , de amplitud constante, se presente la frecuencia cero o la frecuencia de alguno o algunos de los intervalos sea mucho mayor que la de los demás, logrando así que las observaciones se hallen mejor repartidas dentro del intervalo.



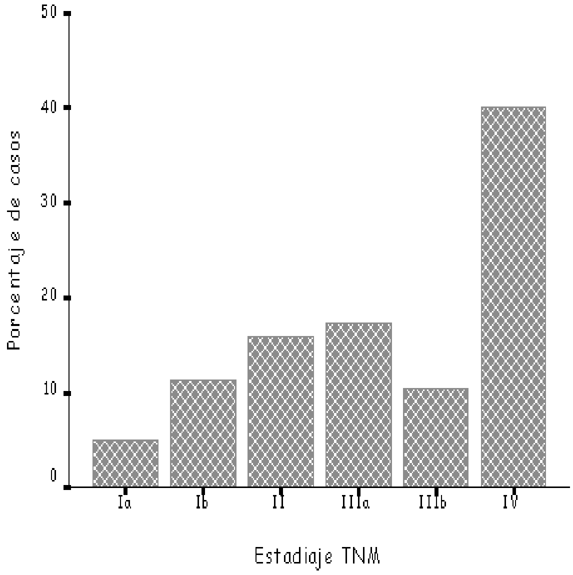
Ojivas: Cuando se trata de relacionar observaciones en un mismo aspecto para dos colectivos diferentes no es posible ejecutar comparaciones sobre la base de la frecuencia, es necesario tener una base estándar, la frecuencia relativa. La ojiva representa gráficamente la forma en que se acumulan los datos y permiten ver cuantas observaciones se hallan por arriba o debajo de ciertos valores. Es útil para obtener una medida de los cuartiles, deciles , percentiles.

Polígono de Frecuencias

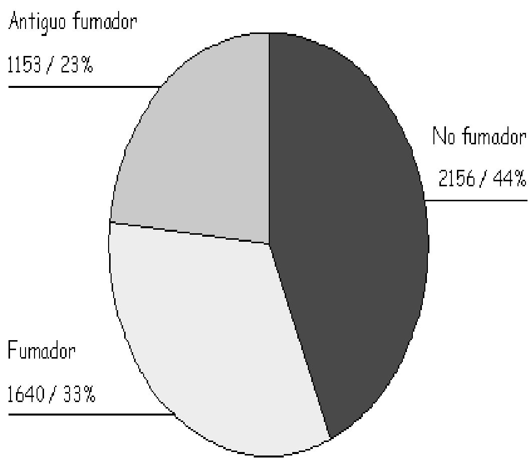
Se puede obtener uniendo cada punto medio (marca de clase) de los rectángulos del histograma con líneas rectas, teniendo cuidado de agregar al inicio y al final marcas de clase adicionales, con el objeto de asegurar la igualdad del áreas.



diagramas de barras son similares a los gráficos de sectores. Se representan tantas barras como categorías tiene la variable, de modo que la altura de cada una de ellas sea proporcional a la frecuencia o porcentaje de casos en cada clase . Estos mismos gráficos pueden utilizarse también para describir ***variables numéricas discretas*** que toman pocos valores



En los **gráficos de sectores**, también conocidos como diagramas de "tartas", se divide un círculo en tantas porciones como clases tenga la variable, de modo que a cada clase le corresponde un arco de círculo proporcional a su frecuencia absoluta o relativa. Un ejemplo se muestra en la . Como se puede observar, la información que se debe mostrar en cada sector hace referencia al número de casos dentro de cada categoría y al porcentaje del total que estos representan. Si el número de categorías es excesivamente grande, la imagen proporcionada por el gráfico de sectores no es lo suficientemente clara y por lo tanto la situación ideal es cuando hay alrededor de tres categorías. En este caso se pueden apreciar con claridad dichos subgrupos.



MEDIDAS DE ASIMETRIA Y APUNTAMIENTO

Sesgo: las curvas que representan los puntos de datos de un conjunto de datos pueden ser simétricas o sesgadas. Las curvas simétricas, tienen una forma tal que una línea vertical que pase por el punto más alto de la curva dividirá el área de ésta en dos partes iguales. Cada parte es una imagen espejo de la otra. En las curvas sesgadas, los valores de su distribución de frecuencias están concentrados en el extremo inferior o en el superior de la escala de medición del eje horizontal. Los valores no están igualmente distribuidos. Las curvas pueden estar sesgadas hacia la derecha (positivamente sesgadas) o sesgadas hacia la izquierda (negativamente sesgadas).

MEDIDAS DE CURTOSIS

Al comparar cuán aguda es una distribución en relación con la Distribución Normal, se pueden presentar diferentes grados de apuntalamiento.

1. Mesocúrtica, Normal

2. PlarticúrtiCa, Menor apuntalamiento

3. Leptocúrtica, Mayor apuntalamiento

2.2. NIVELES DE MEDICIÓN

La medición

Los datos se obtienen a través un proceso llamado medición. Desde este punto de vista, puede definirse **medición** como el proceso por el cual asignamos una categoría (o un valor) a una variable, para determinada unidad de análisis.

Ejemplo: cuando decimos que Martín es varón, estamos haciendo una medición, porque estamos asignando una categoría (varón) a una variable (sexo) para una unidad de análisis (Martín).

Se pueden hacer mediciones con mayor o menor grado de precisión.

Cuanto más precisa sea la medición, más información nos suministra sobre la variable y, por tanto, sobre la unidad de análisis. No es lo mismo decir que una persona es alta, a decir que mide 1,83 metros.

Los diferentes grados de precisión o de contenido informativo de una medición se suelen caracterizar como **niveles de medición**. Típicamente se definen cuatro niveles de medición, y en cada uno de ellos la obtención del dato o resultado de la medición será diferente:

Ejemplos de datos en diferentes niveles de medición

Nivel de medición	**Nivel nominal**	**Nivel ordinal**	**Nivel cuantitativo discreto**	**Nivel cuantitativo continuo**
DATO	Martín es electricista	Elena terminó la secundaria	Juan tiene 32 dientes	María tiene 70 pulsaciones por minuto
Unidad de análisis	Martín	Elena	Juan	María
Variable	Oficio	Nivel de instrucción	Cantidad de piezas dentarias	Frecuencia cardíaca
Categoría o valor	Electricista	Secundaria completa	32	70
Unidad de medida	-------------	------------	Diente	Pulsaciones por minuto

En el nivel nominal, medir significa simplemente asignar un atributo a una unidad de análisis (Martín es electricista).

En el nivel ordinal, medir significa asignar un atributo a una unidad de análisis cuyas categorías pueden ser ordenadas en una serie creciente o decreciente (la categoría ‘secundaria completa’ puede ordenarse en una serie, pues está entre ‘secundaria incompleta’ y ‘universitaria incompleta’).

En el nivel cuantitativo, medir significa además asignar un atributo a una unidad de análisis de modo tal que la categoría asignada permita saber ‘cuánto’ mayor o menor es respecto de otra categoría, es decir, especifica la distancia o intervalo entre categorías (la categoría 70 es el doble de la categoría 35).

Las variables medibles en el nivel cuantitativo pueden ser discretas o continuas.

Una variable discreta es aquella en la cual, dados dos valores consecutivos, no puede adoptar ningún valor intermedio (por ejemplo entre 32 y 33 dientes, no puede hablarse de 32.5 dientes).

En cambio, una variable es continua cuando, dados dos valores consecutivos, la variable puede adoptar muchos valores intermedios (por ejemplo entre 1 y 2 metros, puede haber muchas longitudes posibles).

2.2.1. DATOS DE NIVEL NOMINAL

El término nivel nominal es normalmente usado para referirse a datos que solamente pueden clasificarse en categorías. Sin embargo, no hay mediciones y no hay escalas involucradas, solo hay conteo. En este tipo de nivel de medición el orden en que están acomodadas la categorías es totalmente arbitrario.

Religiones en México (población con 5 años o más, censo del 2000)	
Religión Católica	74 612 373
Religión no católica	3 483 593
Sin religión	2 982 929
total	81 078 895

2.2.2. DATOS DE NIVEL ORDINAL

Este tipo de nivel de medición tiene características similares al nivel nominal con la diferencia de que en el nivel ordinal las categorías indican que unas son más que las otras.

Evaluación de la atención médica	
buena / muy buena	75,7
regular	17,1
mala / muy mala	2
ns / nc	5,2
Total	100

2.2.3. DATOS DE NIVEL DE INTERVALO

En este nivel de medición, las categorías están definidas por intervalos de valores, y están acomodadas en orden a la magnitud de los valores. El tamaño de los intervalos es el mismo.

Calificaciones de los aspirantes a la academia militar	
Calificación	Número de aspirantes
90 – 99	42
80 – 89	19
70 – 79	7
60 – 69	4

2.2.4. DATOS DE NIVEL DE RAZÓN

Escala cuantitativa Racional

En este nivel al igual que en el nivel intervalar, las categorías son del mismo tamaño. La diferencia es que este nivel tiene un punto cero significativo y el valor de los categorías es en relación a ese punto.

Ingreso de los empleados de la compañía en relación a la media de la industria	
2000	12
1000	25
0	32
-1000	17
-2000	10

2.3. ERRORES EN LA ADQUISICIÓN DE DATOS

La enseñanza de la Estadística ha cobrado gran desarrollo en los últimos años, debido a su importancia, ampliamente reconocida, en la formación general del ciudadano. Algunos países han dedicado grandes esfuerzos a diseñar curriculos y materiales específicos, como los elaborados en Inglaterra para el Schools Council Project on Statistical Education por Holmes y cols. (1980), el Quantitative Literacy Project en Estados Unidos (Landewehr y Watkins, 1986; Landewehr y cols., 1987; Gnanadesikan y cols., 1987) y Azar y Probabilidad en España (Godino y cols., 1987). El interés creciente hacia la enseñanza de la Estadística se manifiesta, asimismo, por la existencia de revistas especificas (Teaching Statistics; Induzioni; Stochastik in der Schule); por las conferencias internacionales sobre la Enseñanza de la Estadística (ICOTS I en 1982 en Sheffield ; ICOTS II en 1986 en Victoria e ICOTS III en 1990 en Otago); por la serie de Mesas redondas promovidas por el I.S.I. (la más reciente tuvo lugar en Lennoxville en 1992) y por la formación en 1992 de una asociación internacional IASE (International Association for Statistical Education). Este interés también se demuestra mediante el establecimiento de Centros para la Educación Estadística en Inglaterra, Italia y Estados Unidos; por la Newsletter del International Study Group for Research on Learning Probability and Statistics de la universidad de Granada y la revista The Journal of Statistics Education editados mediante correo electrónico.

El mayor énfasis dado a la Estadística en los diferente curriculos, como los Standards del N.C.T.M. (1989), el Curriculum Nacional para Inglaterra y Gales (D.E.S., 1991) y los nuevos Diseños Curriculares en España (M.E.C., 1988a y 1988b) 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. Esta preparación debería incluir también el conocimiento de las dificultades y errores que los alumnos encuentran en el aprendizaje de la Estadística.

El propósito de este artículo es contribuir a la difusión de los resultados de la investigación sobre estas dificultades y errores, que no son suficientemente conocidos por los profesores. Existen ya algunos “estados de la cuestión” sobre la investigación en educación estadística (Hawkins y Kapadia, 1984; Garfield y Alhgren, 1988; Scholz, 1991 y Shaughnessy, 1992), pero estos trabajos están dirigidos a investigadores más que a profesores y su fin principal ha sido identificar nuevas cuestiones de investigación. Además, se han enfocado especialmente hacia la probabilidad, porque la investigación en este área es mucho más extensa que la relacionada con los conceptos estadísticos, aunque una excepción relevante es el nuevo libro para profesores de estadística de Hawkins, Joliffe y Glickman (1992).

En este artículo analizamos las investigaciones sobre los principales conceptos estadísticos elementales que han sido incluidos en muchos diseños curriculares recientes en los niveles no universitarios. Este análisis muestra la complejidad de algunos de estos tópicos y puede proporcionar al profesor una comprensión mayor del razonamiento estocástico de sus alumnos. Consideramos necesario comenzar esta exposición resaltando la importancia de la investigación sobre errores y dificultades de los alumnos y definiendo algunos conceptos teóricos relacionados con la misma. Advertimos, sin embargo, al lector que:

a) la estadística ha recibido hasta la fecha menos atención que otras ramas de las matemáticas;
b) la mayor parte de la investigación se ha llevado a cabo en situaciones experimentales, en lugar de en situaciones escolares;
c) muchos estudios se centran en niños muy pequeños o en estudiantes de universidad, siendo escasa la investigación en las edades 11 a 16 años;
d) las primeras investigaciones en el campo han sido efectuadas por psicólogos en lugar de por educadores matemáticos, aunque este aspecto está empezando a cambiar.

2. 1NVESTIGACIÓN SOBRE ERRORES, CONCEPCIONES Y OBSTÁCULOS EN DIDÁCTICA. ALGUNOS CONCEPTOS TEORICOS

Gran parte de la investigación teórica y experimental, que se está llevando a cabo actualmente en Didáctica de la Matemática, surge del hecho observable de que el alumno se equivoca cuando se le pide realizar ciertas tareas. El alumno proporciona respuestas erróneas, con respecto a una patrón de evaluación, o simplemente no es capaz de dar ninguna respuesta. En los casos en que no se trata de mera distracción se dice que tal tarea resulta demasiado difícil para el alumno en cuestión. Pero los errores y dificultades no se presentan de un modo aleatorio, imprevisible. Con frecuencia es posible encontrar regularidades, ciertas asociaciones con variables propias de las tareas propuestas, de los sujetos o de las circunstancias presentes o pasadas. La investigación didáctica trata de caracterizar estas regularidades y de construir modelos explicativos, en términos de relaciones entre las variables intervinientes. Algunos autores, como Radatz (1980), consideran el análisis de errores como “una estrategia de investigación prometedora para clarificar cuestiones fundamentales del aprendizaje matemático” (pag. 16). Asimismo, Borassi (1987) presenta el análisis de errores en educación matemática “como un recurso motivacional y como un punto de partida para la exploración matemática creativa, implicando valiosas actividades de planteamiento y resolución de problemas” (pag. 7).

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. Averígüese esto y enséñese consecuentemente”. El interés reciente de los estudios de didáctica por las concepciones de los estudiantes (Confrey, 1990) sería una consecuencia del mencionado principio psicológico.

La problemática que se plantea para la didáctica es que algunas de estas concepciones, que permiten resolver un conjunto de tareas en términos adecuados, se muestran limitadas, inapropiadas cuando se aplican a casos más generales, y que el sujeto muestra una resistencia a su sustitución. En estas circunstancias se habla de la existencia de un obstáculo cognitivo que puede explicar la existencia de errores y dificultades especiales. 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 universal exigirá un punto de vista diferente.

*****El alumno resiste a las contradicciones que el obstáculo le produce y al establecimiento de un conocimiento mejor. Es indispensable identificarlo e incorporar su rechazo en el nuevo saber.

* Después de haber notado su inexactitud, continúa manifestándolo, de forma esporádica. Brousseau ha identificado tres tipos de obstáculos:

a) 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 el razonamiento proporcional.

b) *Obstáculos didácticos:* resultan de las elecciones didácticas hechas para establecer la situación de enseñanza. Por ejemplo, la introducción de un nuevo simbolismo tal como:

$$(\sum x_i)/n$$

cuando los estudiantes necesitan trabajar con ejemplos concretos.

c) 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 mediante un análisis histórico, y superarlos parece ser una condición necesaria para la construcción de una concepción adecuada. Finalmente, hacemos notar que 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. El propósito de la caracterización de concepciones y obstáculos es que ello permite delimitar los distintos componentes implicados en la comprensión de un concepto. La investigación reciente como, por ejemplo, el trabajo de Sierpinska (1991) sobre los “actos de comprensión” de la noción de limite de una sucesión numérica muestra la complejidad del significado de los objetos matemáticos

3. REPRESENTACIÓN GRÁFICA Y TABULACIÓN DE DATOS

Comenzamos nuestra exposición sobre errores y dificultades en el aprendizaje de la Estadística con los que se refieren al uso de representaciones gráficas y tablas de frecuencias.

La destreza en la lectura crítica de datos es un componente de la alfabetización cuantitativa y una necesidad en nuestra sociedad tecnológica. Curcio (1989) describe tres niveles distintos de comprensión de los gráficos:

(a) “Leer los datos”: este nivel de comprensión requiere una lectura literal del gráfico; no se realiza interpretación de la información contenida en el mismo.

(b) “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.

(c) “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.

Por ejemplo, si analizamos las tareas que se requieren en la interpretación de una nube de puntos, “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 intensidad de la covariación, sobre si la relación podría ser representada o no mediante una función lineal o sobre si la dependencia es directa o inversa. Finalmente la predicción del valor de la coordenada y, para un valor de la coordenada x requeriría el trabajo en el nivel de “leer más allá de los datos”.

Curcio (1987) estudió, con alumnos de 4º a 7º, 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; - 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.).

Encontró que las principales dificultades aparecen en los dos niveles superiores (“leer dentro de los datos” y “leer más allá de los datos”). También mostró el efecto de la edad y el curso escolar sobre la comprensión de los gráficos.

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.

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. 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.

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. Li y Shen indican que es de sentido común no comparar 30 sillas y 50 kg. de carne. Sin embargo, presentan un ejemplo de proyecto realizado por los alumnos sobre la industria textil en que se comparan cantidades heterogéneas en un mismo gráfico.

4. CARACTERÍSTICAS ESTADÍSTICAS

4.1. LA MEDIA

Además de ser uno de los principales conceptos estadísticos, la media tiene muchas aplicaciones en cuestiones prácticas de la vida diaria. Este concepto es aparentemente simple, pero Pollatsek y cols. (1981) describen el error consistente en emplear la fórmula de cálculo (120+180)/2 = 150 para resolver la cuestión siguiente:

Hay 10 personas en un ascensor, 4 mujeres y 6 hombres. El peso medio de las mujeres es de 100 libras. y el de los hombres de 180. ¿Cuál es el peso medio de las 10 personas del ascensor?

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.

Otro ítem propuesto por Pollasek y cols. (1981), trata 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:

La media en fluidez verbal de una clase de un colegio es de 400. Si extraemos una muestra aleatoria de 5 estudiantes y resulta que la puntuación de los 4 primeros es de 380. 420, 600, 400. ¿Cuál seria aproximadamente la puntuación esperada para el quinto estudiante?

La respuesta correcta a este ítem es 400, el valor esperado en la población. Sin embargo, algunos alumnos pensaban erróneamente que la puntuación del quinto sujeto debería ser tal que, sumada a las cuatro anteriores, diera una media de 400.

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.

Las investigaciones que hemos descrito se refieren a los aspectos computacionales de la media. Respecto a la comprensión de los aspectos conceptuales, 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:

a) La media es un valor comprendido entre los extremos de la distribución.

b) La suma de las desviaciones de los datos respecto de la media es 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 puede ser una fracción (ello puede no tener sentido para la variable considerada).

f) Hay que tener en cuenta los valores nulos en el cálculo de la media.

g) 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, 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).

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.

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. Este tipo de tarea ha sido también propuesta por Goodchild (1988), quien proporcionó a los estudiantes cajas de cerillas en las que se había impreso la frase “contenido medio 35 cerillas”. Una de sus preguntas, presentada mediante entrevista a 8 alumnos, requería que el alumno construyese una distribución hipotética del contenido de 100 cajas. El hecho más notable de estas distribuciones fue su falta de forma, ya que el gráfico no tenía en absoluto forma acampanada (como la distribución normal). Goodchild sugirió que ello se debe a la falta de comprensión de la media como medida de posición de la distribución obtenida a partir de un proceso estocástico.

Russell y Mokros también encontraron cuatro categorías generales en las que clasificaron las concepciones de los estudiantes sobre los promedios:

a) el “valor más frecuente” o moda;
b) el “valor razonable”;
c) el “punto medio”;
d) una “relación algorítmica”, es decir, una fórmula de cálculo.

Cada uno de estos aspectos puede ser cierto en un caso dado, pero puede ser inapropiado en otro. Finalizan el artículo señalando la necesidad de usar diferentes contextos y representaciones en la enseñanza de un concepto matemático. 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.

4.2. 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. Para Campbell (1974) 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 cuál 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.

Mevarech (1983) encontró en alumnos universitarios las mismas dificultades en el cálculo de la varianza que en el cálculo de la media. En particular, los estudiantes suponen que el conjunto de datos junto con la operación de cálculo de la varianza tiene una estructura de grupo.

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, Huck y cols. (1986) han señalado 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 abcisa 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 $=(10-5)/\sqrt{2.5}=3.16$ que es un limite finito pero mayor que 3.

4.3. ESTADÍSTICOS DE ORDEN

En la actualidad, el estudio de los estadísticos de orden toma una gran importancia por dos motivos:

- El análisis exploratorio de datos, surgido a partir de los estudios de Tukey (1977), 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, tanto a nivel procedimental como a nivel conceptual. En primer lugar, 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. Como indica Schuyten (1991), 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.

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 vanos puntos pueden tener la misma imagen.

Schuyten (1991) ha señalado también la diferencia entre el conocimiento conceptual de la mediana y el método de cálculo que se emplea para obtener su valor. 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.

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 sé:*

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 estudiante 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.

5. ASOCIACION EN TABLAS DE CONTINGENCIA

La idea de asociación estadística extiende la de dependencia funcional, y es fundamental en muchos métodos estadísticos que permiten modelizar numerosos fenómenos en las diversas ciencias. El término asociación se emplea para expresar la existencia de una dependencia estadística entre dos variables arbitrarias, tanto cualitativas como cuantitativas. La palabra correlación suele restringirse a las variables cuantitativas. Ambos términos, asociación y correlación, no implican necesariamente relación de causalidad sino meramente la existencia de covariación entre 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.

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

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

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 la 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 (Scholz, 1987), 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:

> 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	3	17
No la toma	7	1	3
	16	9	25

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

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.

El estudio del razonamiento sobre la asociación estadística fue 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, Inhelder y Piaget (1955) sólo estudiaron este tipo de problemas con sujetos que se encuentran en la etapa de operaciones formales IIIa y IIIb. El contexto empleado por estos autores es el problema de 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. El material se presenta al adolescente en dos formas: sin clasificar, dejando que el sujeto establezca la clasificación (que construya las cuatro casillas de la tabla de doble entrada), bien presentándole las cartas ya clasificadas. Aunque la tarea no es exactamente igual a la presentada en nuestro ejemplo es equivalente para un análisis formal de las estrategias de resolución empleadas por los sujetos, que pasamos a describir. 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 por Smendlund (1963) 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, como señalan Jenkins y Ward (1965), 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. Para el caso general, Jenkins y Ward han propuesto como estrategia correcta 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 (Jennings y cols., 1982). Finalmente, como señala Scholz (1987), los estudios posteriores han mostrado que para la misma estructura del problema de asociación los sujetos adoptan diversas estrategias e incluso una misma persona puede emplear diferente estrategia, dependiendo del contexto.

6. 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.

7. INFERENCIA

7.1. 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. Como Moses (1992) indica “se puede pensar en la inferencia estadística como una colección de métodos para aprender de la experiencia”. Rubin y cols. (1991) indican que, 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 si. 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. El término heurística es empleado en psicología, inteligencia artificial, y en resolución de problemas (Groner y cols., 1983). Aunque no hay un consenso general para el significado del término heurística, normalmente se emplea para referirse a procesos cognitivos que se utilizan para reducir la complejidad de un problema durante el proceso de su resolución.

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.

En la heurística de la representatividad se estima la probabilidad de obtención de una muestra por el parecido de ésta con la población de la que proviene. En consecuencia, aparece cierta insensibilidad al tamaño de la muestra y una confianza exagerada en las pequeñas muestras, fenómeno que se conoce con el nombre de "creencia en la ley de los pequeños números". Por ejemplo, consideremos la siguiente pregunta:

> Una cierta ciudad está atendida por dos hospitales. En el hospital más grande nacen aproximadamente 45 bebés cada día y en el hospital más pequeño nacen aproximadamente 15 bebés cada día. Como sabes, aproximadamente el 50 por ciento de todos los recién nacidos son varones, pero el porcentaje exacto varía de un ida a otro. A veces puede ser mayor que el 50 por ciento, a veces más bajo. Durante un periodo de un año, cada hospital registró los idas en que más del 60 por ciento de los recién nacidos fueron varones. ¿Cuál hospital crees que registró más de estos días:
>
> El hospital grande.
>
> El hospital pequeño.
>
> Aproximadamente igual (esto es, si la diferencia entre ambos es menor del *5* por ciento).

Muchas personas creen que la respuesta correcta debe ser la tercera, puesto que en ambos hospitales la proporción de varones es la misma (60 por ciento) y piensan que este es el único hecho de importancia para determinar la probabilidad de los sucesos requeridos. No conceden atención al tamaño de la muestra, aunque la teoría de la probabilidad nos enseña que hay mayores fluctuaciones del valor de la proporción en las muestras pequeñas que en las muestras grandes.

Según Kahneman y cols. (1982) esta confianza excesiva en las pequeñas muestras tiene graves consecuencias en las aplicaciones de la estadística, especialmente en la investigación. El "creyente en la ley de los pequeños números" tiende a estimar a la baja la amplitud de los intervalos de confianza obtenidos, a sobrestiman la significación de sus resultados estadísticos y a esperar que los resultados obtenidos en los primeros ensayos se le confirmen en el futuro.

Otra consecuencia de la aplicación de la heurística de la representatividad sería el error denominado "falacia del jugador". Por ejemplo, muchas personas creen que después de una racha larga de caras, es más probable obtener una cruz.

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 urna. 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. Puesto que estas personas pueden no haber tenido nunca la experiencia de extraer bolas de una urna, este modelo es de carácter teórico y no práctico. Como indica Steinbring (1986), la idea de independencia tiene también carácter teórico y es difícil estar seguro de su aplicabilidad en un contexto práctico. Por esta razón, la independencia es un buen ejemplo de la diferencia entre la comprensión conceptual de un concepto y la capacidad de aplicar este concepto en la resolución de problemas (Heitele, 1975).

Otro problema relacionado con el muestreo son los diferentes niveles de concreción de un mismo concepto en estadística descriptiva e inferencia (Schuyten, 1991). 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(X) 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 podían 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 nuestra 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 a1eatoria, que coincide con la media de la población en el muestreo aleatorio. Esto supone una gran dificultad conceptual.

7.2. CONTRASTE DE HIPOTESIS

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, como comenta Brewer (1986), 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 paramétricas, 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. En particular Peskun (1987) ha señalado las dificultades de los estudiantes 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.

Uno de los aspectos claves en la correcta aplicación de un contraste de hipótesis es la comprensión del concepto de nivel de significación, que se define como “la probabilidad de rechazar una hipótesis nula, en el caso de ser cierta”, definición que se expresa en la igualdad siguiente:

(1) $\alpha = P(\text{Rechazar } H_0 \mid H_0 \text{ cierta})$

Falk (1986) señala la confusión corrientemente encontrada entre los investigadores que consiste en intercambiar los sucesos condición y condicionado en la definición anterior, esto es,interpretar el nivel de significación en la forma siguiente:

(2) $\alpha = P(H_0 \text{ cierta} \mid \text{se ha rechazado } H_0)$

Sugiere como posible causa de este error el lenguaje empleado en la definición del nivel de significación, esto es la “probabilidad de error Tipo I”. En esta expresión no se indica explícitamente que estamos tratando con una probabilidad condicional, lo que lleva al estudiante a suponer que es posible definir un “suceso condicional”. En consecuencia, no se diferencia entre las dos probabilidades condicionales (1) y (2).

La definición correcta vendría dada por el siguiente enunciado:

i) Un nivel de significación del 5% supone que, en promedio, 5 de cada 100 veces que la hipótesis nula es cierta, la rechazaremos.

La investigación de Birnbaum (1982) y otros autores muestra que algunos estudiantes consideran correcta la siguiente definición (incorrecta) de α:

ii) Un nivel de significación del 5% implica que, en promedio, *5* de cada 100 veces que rechacemos la hipótesis nula, estaremos equivocados.

Las definiciones i) y ii) fueron propuestas por Vallecillos (1991) a estudiantes de universidad, a los que se preguntó para cada una de ellas si era cierta o falsa, analizando el razonamiento de los estudiantes. Vallecillos 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:

a) *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.

b)Las probabilidades de error y relación entre las mismas: Los dos tipos de error tienen probabilidades α (Tipo I) y β (Tipo II). Es necesario la comprensión de las probabilidades condicionales que intervienen en la definición de α y β, de la dependencia de β como función del parámetro desconocido y de las relaciones entre α y β. Aparte del error señalado por Falk (1986) 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 a como probabilidad de error (tanto tipo I como tipo II) en la decisión tomada.

c) Nivel de significación como riesgo del decisor Los valores de α y β, 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.

d) 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.

White (1980), en un trabajo sobre el empleo de los métodos estadísticos en la investigación educativa, ha mencionado la interpretación errónea de un resultado no estadísticamente significativo. También considera otro aspecto relacionado con la comprensión del nivel de significación, que es el problema de las comparaciones múltiples. Este problema se produce cuando se aplican muchos test de significación al mismo conjunto de datos. Por ejemplo, en una investigación epidemiológica, podrían medirse 150 variables en cada persona de un grupo de gente con buena salud y también medir las mismas variables en un grupo de personas con una cierta enfermedad. Si escogiéramos un nivel de significación de 0.05, entonces, puesto que 150x0.05=7.5, cabe esperar 7.5 resultados "estadísticamente significativos" en promedio, incluso si ninguna de las variables medidas está relacionada con la enfermedad estudiada (Moses, 1990)

8. CONSIDERACIONES FINALES

En un "survey" de la literatura de investigación Garfield y Alhgren (1988) señalan las siguientes razones para algunas de las dificultades que surgen en la enseñanza de la estadística:

- Algunos conceptos estocásticos, tales como el de probabilidad, correlación, necesitan del razonamiento proporcional, que ha demostrado ser un tópico difícil en diversas investigaciones.
- Existen falsas intuiciones que los alumnos llevan consigo al empezar la enseñanza. Aunque estas intuiciones son mejor conocidas para el caso de la probabilidad (Piaget e Inhelder, 1951; Fischbein, 1975) aún han sido poco estudiadas para los conceptos estadísticos.
- A veces los alumnos muestran una falta de interés hacia la estadística, porque se les ha enseñado en forma muy abstracta en edades tempranas.

Hay dos razones más que posiblemente influyan en la dificultad del tema: En primer lugar, la Probabilidad y la Estadística tienen un desarrollo reciente. Aunque en la actualidad existe una axiomática para el Cálculo de Probabilidades, comúnmente aceptada, a partir de los trabajos de Kolmogorov, no ha cesado aún la controversia sobre el significado último del término "probabilidad", existiendo diversas escuelas: empiricistas, subjetivistas, lógicas, etc. (Fine,1973). Esta controversia se repite en la inferencia estadística, con la polémica sobre si es posible o no el cálculo inductivo de la probabilidad de una hipótesis y si ello puede lograrse o no con la aproximación clásica o bayesiana de la inferencia (Rivadulla, 1991). Numerosas investigaciones muestran cómo las dificultades epistemológicas, que han debido ser superadas en el desarrollo histórico del conocimiento, se repiten con frecuencia en el aprendizaje del mismo.

Por otro lado, gran parte de los conceptos estadísticos han tenido su origen fuera del campo estricto de la matemática. La Estadística ha sido desde sus comienzos una ciencia interdisciplinar y las grandes etapas de su progreso han estado marcadas por aportaciones originadas a partir de la necesidad de resolver problemas en campos diversos. En la enseñanza los conceptos se presentan aislados de las aplicaciones originales. Pero cada una de estas aplicaciones aporta una parte del significado global de los mismos (Steimbring, 1990)). Así, el concepto de media toma un significado diferente cuando se aplica como centro de gravedad, esperanza de vida o número índice.

En resumen, y como señala Green (1992, pg. 12): “Los conceptos estadísticos proporcionan un área de exploración fascinante. Lo que parece tan obvio y sencillo a los estadísticos (términos como promedio, variabilidad, distribución, correlación, sesgo, aleatoriedad, ...) ha sido el producto de la experiencia de varias generaciones de las mentes más capaces. Es demasiado esperar que esta herencia nos pueda ser transmitida sin esfuerzo por nuestra parte”.

2.4. MEDIDAS DE TENDENCIA

Medidas de tendencia central

Las características globales de un conjunto de datos estadísticos pueden resumirse mediante una serie de cantidades numéricas representativas llamadas parámetros estadísticos. Entre ellas, las medidas de tendencia central, como la media aritmética, la moda o la mediana, ayudan a conocer de forma aproximada el comportamiento de una distribución estadística.

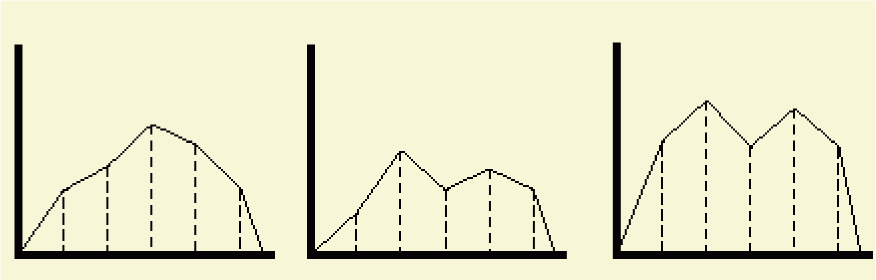
Medidas de centralización

Se llama **medidas de posición**, **tendencia central** o **centralización** a unos valores numéricos en torno a los cuales se agrupan, en mayor o menor medida, los valores de una variable estadística. Estas medidas se conocen también como **promedios**.

Para que un valor pueda ser considerado promedio, debe cumplirse que esté situado entre el menor y el mayor de la serie y que su cálculo y utilización resulten sencillos en términos matemáticos.

Se distinguen dos clases principales de valores promedio:

- Las **medidas de posición centrales**: medias (aritmética, geométrica, cuadrática, ponderada), mediana y moda.
- Las **medidas de posición no centrales**: entre las que destacan especialmente los cuantiles.



Las medidas de centralización son parámetros representativos de distribuciones de frecuencia como las que ilustra la imagen.

Media aritmética

Se define **media aritmética** de una serie de valores como el resultado producido al sumar todos ellos y dividir la suma por el número total de valores. La media aritmética se expresada como $\overline{x}$.

Dada una variable x que toma los valores x_1, x_2, ..., x_n, con **frecuencias** absolutas simbolizadas por f_1, f_2, ..., f_n, la media aritmética de todos estos valores vendrá dada por:

$$\overline{x} = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_n x_n}{f_1 + f_2 + \ldots + f_n} = \frac{\sum f_i x_i}{\sum f_i} \text{ con } i = 1, 2, \ldots, n$$

Media ponderada

En algunas series estadísticas, no todos los valores tienen la misma importancia. Entonces, para calcular la media se ponderan dichos valores según su peso, con lo que se obtiene una **media ponderada**.

Si se tiene una variable con valores x_1, x_2, ..., x_n, a los que se asigna un peso mediante valores numéricos p_1, p_2, ..., p_n, la media ponderada se calculará como sigue:

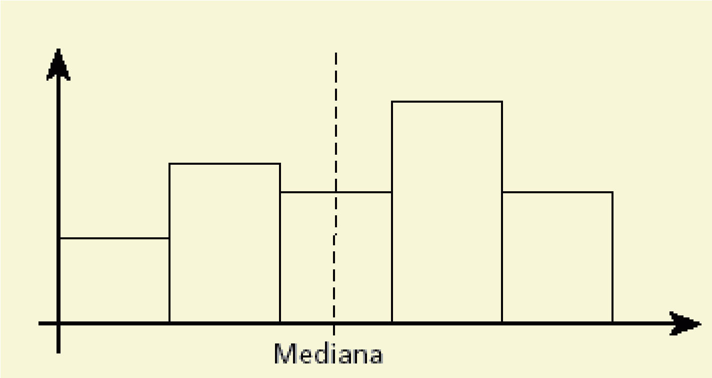
$$\overline{x} = \frac{p_1 x_1 + p_2 x_2 + \ldots + p_n x_n}{p_1 + p_2 + \ldots + p_n} = \frac{\sum p_i x_i}{\sum p_i} \text{ con } i = 1, 2, \ldots, n$$

Mediana

La media aritmética no siempre es representativa de una serie estadística. Para complementarla, se utiliza un valor numérico conocido como **mediana** o valor central.

Dado un conjunto de valores ordenados, su mediana se define como un valor numérico tal que se encuentra en el centro de la serie, con igual número de valores superiores a él que inferiores. Normalmente, la mediana se expresa como Me.

La mediana es única para cada grupo de valores. Cuando el número de valores ordenados (de mayor a menor, o de menor a mayor) de la serie es impar, la mediana corresponderá al valor que ocupe la posición (n + 1)/2 de la serie. Si el número de valores es par, ninguno de ellos ocupará la posición central. Entonces, se tomará como mediana la media aritmética entre los dos valores centrales.



Determinación de la mediana de una serie de valores.

Moda

En una serie de valores a los que se asocia una frecuencia, se define **moda** como el valor de la variable que posee una frecuencia mayor que los restantes. La moda se simboliza normalmente por Mo.

Un grupo de valores puede tener varias modas. Una serie de valores con sólo una moda se denomina unimodal; si tiene dos modas, es bimodal, y así sucesivamente.

Media geométrica

La **media geométrica** de una serie de valores x_1, x_2, ..., x_n, denotada por Mg, se define como la raíz n-sima del producto de todos estos valores. Esta medida central se utiliza principalmente para promediar índices, porcentajes y otros valores numéricos:

$$Mg = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n},\ .$$

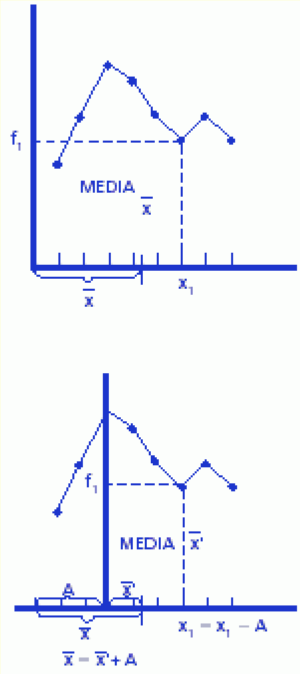
Media cuadrática

La **media cuadrática**, otra medida de tendencia central, se define como la raíz cuadrada de la media aritmética de los valores de la variable estadística considerada elevados al cuadrado.

$$Ma = \sqrt{\sum x_i^2 / n}$$

siendo $i = 1, 2, \ldots, n$

El cálculo de la media aritmética de una serie de valores



puede abreviarse si se resta a todos los valores un mismo número elegido convenientemente.

Cuantiles

Los **cuantiles** son medidas de tendencia no centrales, que permiten determinar la proporción de la población de una variable estadística cuyos valores estadísticos son menores o iguales que un valor tomado como referencia. Este valor puede determinarse dividiendo la población en diez partes (deciles), cien partes (percentiles), etcétera.

BIBLIOGRAFIA

Ausubel, D. P., Novak, J. D. y Hanesian, H. (1983). *Psicología educativa. Un punto de vista cognoscitivo.* (México: Trillas).

BATANERO, C. (2000): "Significado y comprensión de las medidas de tendencia central". en Uno. Revista de didáctica de las Matemáticas. Barcelona. Editorial Graó, n.º 35, pp. 41-58.

BATANERO, C.; GODINO, J. D.; GREEN, D. R.; HOLMES, P., y VALLECILLOS, A. (1994): "Errors and difficulties in understanding elementary statistical concepts", en Iternational Journal of Mathematics Education in Science and Technology, n.º 25 (4), pp. 527-547.

Estepa Castro, A. (1990). *Enseñanza de la estadística basada en el uso de ordenadores: Un Estudio exploratorio.* Memoria de Tercer Ciclo. (Universidad de Granada: Departamento de Didáctica de la Matemática).

Godino, J. D., Batanero, M. C. y Cañizares, M. J. (1987). *A zar y Probabilidad. Fundamentos didácticos y propuestas curriculares. (*Madrid: Síntesis).

Jennings, D. L.; Amabile, T.M. y Ross, L. (1982). Informal covariation assessment: data based versus theory based judgements. In D. Kahneman; P. Slovic and A. Tversky (Eds.): *Judgement under uncertainty: heuristics and biases* (New York: Cambridge University Press), 211-30.

M. E. C. (1988a.). *Diseño curricular base para la enseñanza primaria.* (Madrid: Ministerio de Educación y Ciencia).

M. E. C. (1 988b). *Diseño curricular base para la enseñanza secundaria obligatoria* (Madrid: Ministerio de Educación y Ciencia).

Piaget, J. e Inhelder, B. (1951). *La génese de l'idée de hasard chez l'enfant.* (Paris: Presses Universitaires de France).

Rivadulla, A. (1991). *Probabilidad e inferencia científica* (Barcelona: Anthropos).

Rubin, A. y Rosebety, 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.

Rubin, A.; Bruce, B. y Tenney, Y. (1991). Learning about sampling: Trouble at the core of

Scholz, R. (1987). *Decision making under uncertainly.* Amsterdam. North Holland.

Scholz, R. (1991). Psychological research on the probability concept and its aquisition. In R. Kapadia (Ed). *Chance Encounters: Probability in Education.* (Amsterdam: Reidel), 213-249.

Vallecillos, A. (1992). *Nivel de significacion en un contraste estadístico de hipotesis. Un estudio teorico-experimental de errores en estudiantes universitarios.* Memoria de Tercer Ciclo. Departamento de Didactica de la Matematica. Universidad de Granada.

White, A. L. (1980). Avoiding errors in educational research. En R. J. Shumway (Ed.): *Research in Mathematics Education* (Reston, Va: N.C.T.M.), 49-65.