



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

“TÉCNICAS DE OBTENCIÓN Y REPRESENTACIÓN DE DATOS. TABLAS ESTADÍSTICAS”

AUTORÍA SILVIA BORREGO DEL PINO
TEMÁTICA ESTADÍSTICA
ETAPA UNIVERSITARIA, POST-UNIVERSITARIA

Resumen

La Estadística puede definirse como el conjunto de métodos que tiene por objeto la obtención, tratamiento e interpretación de datos de observación relativos a un grupo de individuos o unidades. A continuación, comentaremos técnicas empleadas en Estadística Descriptiva para la obtención de datos y recogida de la información, para su estructuración y tratamiento mediante tablas.

Palabras clave

Estadística.

Obtención de datos estadísticos.

Representación de datos estadísticos.

Tablas estadísticas.

1. INTRODUCCIÓN

La Estadística actual es el resultado de la unión de dos disciplinas que evolucionan de manera independiente hasta confluir en el siglo XIX: el Cálculo de Probabilidades, que nace en el siglo XVII como teoría matemática de los juegos de azar, y la Estadística, ciencia del estado, que estudia la recogida y descripción de datos y que es de raíces bastante más antiguas.

Esto hace que una de las acepciones más aceptadas de Estadística sea la que la define como el conjunto de métodos que tiene por objeto la obtención, tratamiento y la interpretación de un conjunto de datos de observación relativos a un grupo de individuos o unidades. La Estadística actúa como disciplina puente entre los modelos matemáticos y los fenómenos reales.



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

Es difícil establecer una cronología exacta de los orígenes de la Estadística. Desde la antigüedad, los estados han recogido información sobre la población y riquezas que existía en sus dominios (censos, inventarios...). Por otra parte, desde el siglo XVII se ha tratado de interpretar fenómenos biológicos y sociales de poblaciones a partir de datos numéricos (tablas de mortalidad, contrastación de la teoría de Darwin, estudio de la herencia humana...) mediante procesos deductivos (Estadística Descriptiva). Desde finales del siglo XIX, aplicando métodos inductivos (Estadística Inferencial), la Estadística ha visto ampliado su campo de aplicación a prácticamente todos los sectores (Ingeniería, Física, Medicina...).

Así, la Estadística se ocupa de la descripción de datos (procedimientos para resumir la información), del análisis de muestras (elegir muestras representativas y hacer inferencia), de la contrastación de hipótesis (comparar predicciones con datos observados), de la medición de relaciones (relación estadística), de la predicción (mediante el estudio del historial de las variables), etc.

El método científico se basa en dos tipos de razonamientos: el deductivo (de lo general a lo particular) y el inductivo (de lo particular a lo general). Ambos tipos de razonamientos darán lugar respectivamente a la Estadística Descriptiva y a la Inferencial. Mientras que la Estadística Descriptiva trata del recuento, ordenación y clasificación de los datos obtenidos por las observaciones y mediciones, la Estadística Inferencial, haciendo uso del Cálculo de Probabilidades, describe, predice, compara y generaliza resultados a una población estadística a partir de la información que obtiene de una parte de la población.

Cabe destacar a Florence Nightingale, considerada madre de la enfermería moderna y creadora del primer mapa conceptual de enfermería. Florence, en la guerra de Crimen, postuló tres categorías que explicaban el origen de las bajas: muertes por enfermedades como la cólera, el tifus, etc., muertes por heridas propias de guerra, y muertes por otras causas. Presentó un novedoso diagrama de frecuencias, inventado por ella, y llamado gráfico polar. En dicho gráfico se miraba de forma trágica y patente las diferencias significativas de los sectores de frecuencias, donde se demostraba que, largamente, era mayor la frecuencia de la primera causa sobre las otras dos, con el agregado que el segundo sector era el de menor frecuencia. Aparte de otros aportes, que hizo a la Estadística, introdujo el método científico en las teorías de decisión para la guerra, en lo que se refiere a ayuda asistencial médica.

2. EL MÉTODO ESTADÍSTICO

La investigación científica es un proceso de aprendizaje dirigido. El objeto de los métodos estadístico es hacer que dicho proceso sea lo más eficiente posible. Ese aprendizaje parte de una hipótesis inicial y, por un proceso de deducción, conduce a ciertas consecuencias que pueden ser comparadas con los datos. Cuando esas consecuencias y los datos no coinciden, la discrepancia puede conducir a la modificación de las hipótesis. De este modo, se inicia un segundo ciclo de iteración, deduciendo las consecuencias de las hipótesis modificadas y comparando de nuevo con los datos, pudiendo llevar a nuevas modificaciones y ganancia del conocimiento.



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

El método estadístico es el procedimiento mediante el cual se sistematiza y organiza el proceso de aprendizaje iterativo deductivo-inductivo. Las etapas básicas de aplicación de los métodos estadísticos son:

1. Planteamiento del problema.

Hay que limitar el problema en cuestión a términos abordables. Para ello hay que definir su ámbito de aplicación, así como las características a tener en cuenta. Las conclusiones que se obtengan sólo podrán aplicarse a la población definida, dependiendo su validez de las variables medidas. El resultado de esta fase es una variable respuesta y un conjunto de variables explicativas.

2. Diseño del experimento.

Debemos determinar un modelo matemático-estadístico que se aproxime a la realidad objeto de estudio. Cuando la información disponible corresponde a una única variable se denominan modelos extrapolativos y cuando incluyen además los valores de una o más variables explicativas se denominan modelos explicativos. Por otro lado, si el objeto es representar un instante temporal dado único, se denominan modelos estáticos, mientras que si se desea representar una evolución a lo largo del tiempo, se llaman modelos dinámicos. En cualquier caso, pueden representarse por $y = \alpha + x$, donde y es la observación, α es la parte sistemática y x es la parte aleatoria.

3. Obtención de los datos.

A continuación, hay que medir los valores de la variable de interés. La información puede recogerse por muestreo, observando pasivamente una muestra y anotando los valores de las variables, o por diseño de experimentos, fijando los valores de ciertas variables y observando la respuesta de otras. Algunos autores clasifican los datos en primarios y secundarios, según hayan sido éstos recogidos y registrados por el investigador por primera vez o no.

4. Depuración de los datos de la muestra.

Tras la recogida de datos, el estadístico debe proceder a su depuración para detectar posibles errores (de medición, de transcripción,...). Suelen usarse técnicas en el cálculo de outliers.

5. Estimación de parámetros.

Utilizando la información disponible de la muestra puede estimarse el valor o valores de ciertos parámetros, así como el posible error de estimación.

6. Simplificación.

Determinar si es posible conseguir un modelo más simple, determinando si todos los parámetros definidos previamente son o no necesarios.

7. Crítica del modelo. Formulación de la respuesta.

Se investiga la compatibilidad entre la información empírica y el modelo estadístico, Tras esta fase podemos aceptar el modelo como correcto y hacer uso de él, o en caso contrario, volver a la fase 2 y reformular otro modelo.



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

Puede notarse la metodología iterativa señalada al comienzo de este apartado. Así, la Estadística Descriptiva recoge, organiza y obtiene los parámetros de una serie de datos, mientras que la Estadística Inferencial describe, predice, compara y generaliza los resultados obtenidos de una muestra a toda la población.

3. RECOGIDA DE LA INFORMACIÓN

El primer paso en toda investigación estadística consiste en fijar el conjunto de elementos que queremos estudiar, que llamaremos población o universo. Cada elemento de la población se denomina individuo o unidad estadística. La población puede ser el conjunto de personas de una localidad, las llamadas telefónicas a una central... Llamaremos muestra a un subconjunto limitado extraído de la población, con objeto de reducir el número de experiencias.

Cuando una investigación se realiza a toda la población decimos que estamos realizando un censo (Estadística Descriptiva), pero si recogemos datos de observación relativos sólo a una muestra diremos que estamos realizando una encuesta o sondeo (Inferencia Estadística).

El muestreo es el primer paso en el proceso de inferencia estadística y se acude a él por resultar más rápido, y económico, o porque se usan procedimientos destructivos, o por necesidad de pruebas testigos, etc.

Existen los siguientes tipos de elección de muestras:

- Muestreo aleatorio. Consiste en listar todos los elementos de la población y seleccionar aleatoriamente n elementos de la
- Muestreo sistemático. Consiste en seleccionar los n elementos de la muestra de k en k , siendo k el entero más próximo $a = \frac{\text{Número de elementos de la población}}{\text{Tamaño de la muestra}} = \frac{N}{n}$, y partiendo de un arranque aleatorio i , comprendido entre 1 y k .
- Muestreo por estratos. Consiste en dividir la población en estratos homogéneos y en cada uno de ellos tomar muestras aleatorias simples.
- Muestreo por conglomerados. Consiste en dividir la población en conglomerados. Se eligen al azar unos pocos de estos conglomerados y la muestra estará formada por todos los elementos de ellos o por muestras aleatorias simples de éstos.
- Muestreo anidado. Es una generalización del muestreo por conglomerados. En la primera etapa se seleccionan una serie de conglomerados o unidades muestrales primarias, en una segunda se seleccionan conglomerados más pequeños, pertenecientes a los anteriores, llamados unidades muestrales secundarias, y así sucesivamente cuantas etapas sean necesarias.

3.1. Censos



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

Lo que se denomina censo es una operación estadística, a gran escala, en la que se trata de investigar el 100 % de los objetos susceptibles de estudio. Habitualmente, los censos que se realizan corresponden a personas, viviendas, edificios, etc.

Aunque pueden pensarse que el censo es la mejor manera de estudiar las características de una población, éste plantea una serie de inconvenientes, entre los que se encuentran:

- El coste que conlleva su ejecución.
- El tiempo que se necesita para su realización.
- La aparición de errores intrínsecos a la naturaleza de las operaciones a gran escala.

Trataremos a continuación el tercer inconveniente.

La aparición de errores intrínsecos proviene principalmente de la recogida de datos y el proceso de información. Podemos distinguir entre errores de cobertura y errores de contenido.

El error de cobertura es el error en la enumeración de las personas o viviendas provocado por omisiones durante el recorrido censal, error por subenumeración o subcobertura censal; o bien por inclusiones erróneas en el censo, error de superenumeración o supercobertura. Estos errores suelen ser debidos a operaciones de campo defectuosas, falta de cuidado por los agentes enumeradores, falta de cooperación por parte de los respondientes o simplemente porque se extravían los cuestionarios o son destruidos durante la operación censal.

El error de contenido se define como el error en las características investigadas en el censo en aquellas unidades que fueron correctamente incluidas en el censo. Estos errores se producen por falta de respuesta, respuesta errónea o inconsistente en determinados ítems contenidos en el cuestionario, errores en la interpretación de las preguntas y errores en las codificaciones y proceso de datos.

Dado que el objetivo fundamental de un censo es el recuento exhaustivo de toda la población, los errores de cobertura adquieren especial relevancia.

Los instrumentos de medida de estos errores difieren fundamentalmente del nivel de sofisticación técnica y en los requerimientos de la información adicional al censo. Si sólo disponemos de la información que proporciona el propio censo, las posibilidades de evaluación son bastante limitadas y están circunscritas a la realización de análisis demográficos y sólo sirven para proporcionar indicadores de posibles errores; por ejemplo, la falta de población indica errores de cobertura. Otras posibilidades pasan por el cálculo de índices sensibles a estos errores, como puede ser el número de personas por hogar y su distribución por áreas geográficas; también pueden utilizarse ratios y distribuciones censales, por ejemplo, la proporción de personas censadas por edad y sexo, o el número de niños nacidos según la edad de la madre, estos ratios deben proporcionar series de progresión suave, cualquier salto brusco puede indicar un error.

Si disponemos de fuentes ajenas al propio censo se pueden realizar análisis más sofisticados. Si disponemos de la población en un censo anterior, podemos elaborar, a partir de tasas de intercambio durante el periodo intercensal, una "población esperada" que nos sirva de referencia para la que obtengamos; a partir del dato del censo anterior, la suma algebraica de nacimientos y defunciones,



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

inmigrantes y emigrantes, obtenemos esa estimación. Como es obvio, la bondad de la estimación depende de la fiabilidad de los registros utilizados.

También pueden construirse modelos matemáticos de proyecciones de población utilizando los niveles y tendencias de fertilidad, mortalidad y migración en el periodo intercensal; para estimar esos niveles y tendencias se deben utilizar registros ajenos al censo, como pueden ser determinados registros administrativos y diferentes encuestas demográficas.

Por otro lado, las encuestas demográficas pueden ser utilizadas para evaluar determinadas características de la población recogidos en el censo, bajo la hipótesis de que los errores ajenos al muestreo afectan a los datos de una encuesta son muy inferiores a los que afectan a un censo.

Por último, otro de los instrumentos más utilizados para evaluar la calidad de un censo son las llamadas Encuestas de Post-enumeración censal o Encuestas de Evaluación, que consisten en repetir el propio censo en una muestra o áreas que son recorridas exhaustivamente poco tiempo después de realizado el censo por agentes especialmente adiestrados. Estos agentes contrastan la enumeración obtenida en el censo con el existente en la realidad, prestando especial atención a las omisiones e inclusiones erróneas de unidades censales.

Como hemos visto, los censos pueden ser evaluados a partir de encuestas pero éstos a su vez proporcionan a las investigaciones por muestreo:

- Información indispensable para la construcción de las bases de muestreo.
- Información auxiliar para ser utilizada en:
 - o Procesos de estratificación.
 - o Procesos de estimación

3.2. Encuestas

La realización de una encuesta consta de varias etapas desde que se concibe hasta que se culmina. Estas etapas, según William Cochran, son las siguientes:

1. Indicar los objetivos de la encuesta.

Deben expresarse con la mayor precisión posible para obtener resultados que concuerden con los objetivos inicialmente planteados.

2. Definición de la población sobre la que se va a realizar la encuesta.

Debe tenerse un criterio que defina por comprensión a la población.

3. Determinación de las preguntas a efectuar.

El número de preguntas debe ser el menor posible y ha de estar en consonancia con el interés de los datos que se desean obtener. Deben ser de fácil comprensión y carentes de ambigüedad.

4. Expresión del grado de precisión que se desea.



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

Como en todo estudio estadístico, ha de decidirse a priori el grado de precisión deseado.

5. Métodos de medición.

Las formas de registrar las preguntas y respuestas pueden permitir codificar estas últimas y simplificar el tratamiento posterior de los resultados.

6. Elección de las unidades de muestreo.

Antes de seleccionar la muestra, la población debe ser dividida en unidades de muestreo que son clases de una clasificación de la población. Al conjunto de las unidades e informaciones de que disponemos para realizar la encuesta se la denomina marco de la encuesta.

7. Estratificación de la población.

La estratificación puede ser uniforme, cuando el tamaño de la muestra es el mismo para todos los estratos, proporcional, si el tamaño de la muestra en cada estrato es proporcional al número de elementos del estrato, y óptima, cuando se relaciona el tamaño de la muestra de cada estrato con la concentración de los elementos del estrato respecto a un valor medio de la población.

8. Selección de la muestra.

Para cada elección de método de selección de la muestra se pueden hacer estimaciones acerca del tamaño de la muestra, una vez prefijado el nivel de precisión a emplear. El tamaño de la muestra condicionará el tipo de muestreo que hemos de utilizar.

9. Encuesta piloto.

Es conveniente probar la encuesta a pequeña escala. De esta forma puede conseguirse una aproximación al costo real que va a suponer la encuesta.

10. Organización del trabajo de campo.

Han de limitarse las características espacio-temporales del trabajo del encuestador.

11. Tabulación y análisis de datos.

Una vez recogidos los datos, han de depurarse los cuestionarios obtenidos, así como codificar las respuestas si es posible. Igualmente habrán de hacerse valoraciones acerca de las respuestas desestimadas.

12. Información para futuras encuestas.

La información obtenida permitirá determinar el tamaño de la muestra para futuras encuestas de modo que se consigan resultados óptimos y una mayor rapidez en la obtención de resultados.

Una vez vistas las etapas de las que consta una encuesta, veamos ahora los elementos básicos de la misma. Entre ellos nos encontramos con:

- La información que se desea recibir y con qué precisión.
- A qué población va dirigida y qué muestra va a ser seleccionada.



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

- El método elegido para llevarla a cabo.
- La forma de elaborarla para que el análisis posterior sea válido y fácil de realizar.
- Indicar si es ocasional, repetida o continua.

4. CARACTERES Y VARIABLES ESTADÍSTICAS

Una vez fijada la población debemos indicar cuáles son las características o cualidades que nos interesan estudiar en esa población, estableciendo la forma en la que deben medirse, las unidades de medida...

Estas características observables en una población se clasifican en cualitativas, que son aquellas que no se pueden cuantificar, tales como el color de pelo, el gusto musical, grupo sanguíneo,... Las características que no son cualitativas las llamamos cuantitativas, que son aquellas que sí se pueden cuantificar, como es la estatura, el número de hijos...

A su vez, las características cuantitativas se dividen en dos tipos, las discretas y las continuas. Las características cuantitativas discretas son aquellas que toman valores aislados, como es el número de televisores en una unidad familiar o el número de hijos de una pareja. Por el contrario, las variables continuas pueden tomar cualquier valor comprendido en un determinado rango o intervalo, aunque muchas veces la unidad de medida no nos permita tal hecho. Esto ocurre, por ejemplo, al estudiar la altura de una población, que aunque sabemos que es una variable continua, los aparatos de medida sólo nos permiten tomar éstas con una determinada aproximación.

Algunas veces también es preferible, en el caso de las variables discretas con un gran número de resultados, tratarlas como si fueran variables continuas y viceversa.

Notaremos por N al número de individuos de la muestra o de la población, y por X a la variable estadística. Si la variable es discreta, k representa el número de valores diferentes obtenidos en la muestra o población, y x_i el valor de dicha variable que ocupa el i -ésimo lugar en una ordenación de menor a mayor. Si la variable es continua, k representa el número de clases o intervalos I_i en que podemos dividir la población, donde $I_i = [e_i, e_{i+1})$ es el i -ésimo intervalo con $i = 1, 2, 3, \dots, k$; y $a_i = e_{i+1} - e_i$ es la amplitud de dicho intervalo. El punto medio del intervalo se llama marca de clase i -ésima y se nota

por c_i , es decir, $c_i = \frac{e_{i+1} - e_i}{2}$.

Cuando el número de valores distintos que toma la variable discreta sea grande, o cuando ésta sea continua, conviene agrupar los datos en clases. El número k de clases a considerar debe estar comprendido entre 5 y 20. Una regla usada es tomar k igual al entero más próximo a \sqrt{N} .

4.1. Variables Estadísticas Bidimensionales

De cada elemento de la población pueden analizarse simultáneamente dos o más caracteres obteniendo de cada individuo un dato referente a cada carácter. En este apartado estudiaremos dos



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

datos numéricos para cada uno de los elementos de la población, el primero de los cuales se representa por la variable estadística X y el segundo por la variable estadística Y . A este modelo matemático lo denominaremos variable estadística bidimensional. Por ejemplo, podemos estudiar para una población los ingresos netos mensuales (X), y los gastos totales al mes (Y), formando la variable (X, Y) .

Dado que las variables en estudio pueden ser discretas o continuas, las variables estadísticas bidimensionales pueden ser:

- Discreta-discreta. Por ejemplo, X = “número de goles en 40 partidos de fútbol”, Y = “número de espectadores en dichos partidos”.
- Discreta-continua. Por ejemplo, X = “número de calzado de los individuos de una población”, Y = “estatura de dichos individuos”.
- Continua-continua. Por ejemplo, X = “altura del padre del alumnado”, Y = “altura del alumnado”.

NOTA: Se extiende de forma natural a variable estadística n -dimensional: (X_1, X_2, \dots, X_n) .

5. TABLAS ESTADÍSTICAS

Las tablas estadísticas recogen las primeras informaciones que se obtienen al realizar la encuesta sobre los datos, relativa a uno o varios caracteres.

Consideremos que la población está formada por N individuos o elementos y se describe según un carácter C en k modalidades C_1, C_2, \dots, C_k .

Se denomina campo de la variable al conjunto de los valores que ésta toma. Por otro lado, llamamos recorrido o rango a la diferencia entre el mayor y el menor valor de la variable.

Para $i = 1, 2, \dots, k$, llamaremos frecuencia absoluta asociada a la modalidad C_i del carácter C , al número de elementos n_i de la población o de la muestra estudiada, que posee dicha modalidad. Se verifica

$$\sum_{i=1}^k n_i = N.$$

Para $i = 1, 2, \dots, k$, llamaremos frecuencia absoluta acumulada asociada a la modalidad C_i del carácter C , a la suma N_i de las frecuencias absolutas de las modalidades C_j , con $j \leq i$. $\sum_{j=1}^i N_j = N_i$. Se comprueba

fácilmente que $N_1 = n_1$ $N_i \leq N_{i+1}$ $N_k = N$.

Para $i = 1, 2, \dots, k$, llamaremos frecuencia relativa asociada a la modalidad C_i del carácter C , al cociente f_i entre la frecuencia absoluta correspondiente a la modalidad y el número de elementos de la población

$$N f_i = \frac{n_i}{N}. \text{ De manera obvia se comprueba: } 0 \leq f_i \leq 1 \quad \sum_{i=1}^k f_i = 1.$$



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

Para $i = 1, 2, \dots, k$, llamaremos frecuencia relativa acumulada asociada a la modalidad C_i del carácter C , al cociente F_i entre la frecuencia absoluta acumulada correspondiente a la modalidad y el número de elementos de la población o muestra. $F_i = \frac{N_i}{N}$. De manera obvia se comprueba:

$$0 \leq F_i \leq 1 \quad F_i \leq F_{i+1} \quad F_k = 1.$$

Las frecuencias relativas sirven para comparar resultados obtenidos en encuestas realizadas sobre muestras o poblaciones de distintos tamaños.

En algunas ocasiones es preferible expresar las proporciones en porcentajes: $P_i = 100 \cdot f_i = 100 \cdot \frac{n_i}{N}$.

Llamaremos tabla estadística de frecuencias a la disposición actual de presentar la información relativa al estudio de un carácter. Suelen presentarse en una columna las modalidades y frente a cada una de ellas su frecuencia absoluta o relativa. A veces, también se incluyen las acumuladas.

C	n_i	f_i	N_i	F_i	P_i
C ₁	n ₁	f ₁	N ₁	F ₁	P ₁
C ₂	n ₂	f ₂	N ₂	F ₂	P ₂
...
C _i	n _i	f _i	N _i	F _i	P _i
...
C _k	n _k	f _k	N _k	F _k	P _k
	N	1			

Ejemplo 1. Un dentista observa el número de caries de 100 niños de un colegio. La información recogida aparece resumida en la siguiente tabla:

Nº caries	n_i	f_i	N_i	F_i	P_i
0	25	0,25	25	0,25	25 %
1	20	0,20	45	0,45	45 %
2	35	0,35	80	0,80	80 %
3	15	0,15	95	0,95	95 %
4	5	0,05	100	1	100 %



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

Ejemplo 2. Estudio sobre el color de los ojos de 20 niños y niñas de una clase.

Color de ojos	n_i	f_i	N_i	F_i	P_i
Azules	5	0,25	5	0,25	25 %
Marrones	12	0,6	17	0,85	60 %
Verdes	3	0,15	20	1	15 %

Ejemplo 3. La siguiente tabla muestra los pesos de 20 alumnos y alumnas de 3º E.S.O.

Peso (kg)	c_i	n_i	f_i	N_i	F_i	P_i
[45,55)	50	9	0,45	9	0,45	45 %
[55,65)	60	8	0,4	17	0,85	40 %
[65,75)	70	1	0,05	18	0,90	5 %
[75,85)	80	1	0,05	19	0,95	5 %
[85,95)	90	1	0,05	20	1	5 %

En el caso de datos agrupados, vemos como en la tabla de frecuencias se añade una nueva columna que contiene a las marcas de clase de cada intervalo.

Habitualmente a las tablas estadísticas se le pueden añadir columnas para facilitar los cálculos de los diferentes parámetros estadísticos.

Si la variable estadística fuera bidimensional (estudiamos dos caracteres), los datos se disponen en una tabla de doble entrada. Si disponemos de 2 variables X e Y, una de ellas toma los valores x_1, x_2, \dots, x_n y la otra, y_1, y_2, \dots, y_m , los datos se disponen de la siguiente forma:

$x \backslash y$	y_1	y_2	...	y_j	...	y_m
x_1	n_{11}	n_{12}	...	n_{1j}	...	n_{1m}



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

x_2	n_{21}	n_{22}	...	n_{2j}	...	n_{2m}
.
.
.
x_i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{im}
.
.
.
x_n	n_{n1}	n_{n2}	...	n_{nj}	...	n_{nm}

donde n_{ij} representa la frecuencia absoluta del par (x_i, y_j) .

En el caso de que las variables sean cualitativas la tabla recibe el nombre de tabla de contingencias.

Debemos advertir que la forma de proceder en la elaboración de una tabla actual se puede acelerar notablemente con la utilización de una hoja de cálculo.

6. CONCLUSIÓN

En sus orígenes históricos, la Estadística estaba íntimamente ligada a las cuestiones del Estado, y de ahí proviene el nombre de Estadística. Hoy en día está presente prácticamente en todos los campos de estudio, además de las cuestiones de Estado.

La Estadística resulta fundamental en ciencias como Física (control de errores de medida, leyes de distribución de los gases, principio de incertidumbre...), Química (velocidad de reacciones químicas, composiciones...), Medicina (métodos de diagnóstico, comparación de tratamiento...), Veterinaria (control de epidemias), Biología (estudio y reproducción de especies), Meteorología (previsión del tiempo), etc. En las Ciencias Sociales, desde los primeros estudios sobre tablas de mortalidad aplicados por las compañías de seguros hasta los complejos modelos económicos, la Estadística ha sido y es una herramienta fundamental para las ciencias relacionadas con el estudio de los seres



ISSN 1988-6047 DEP. LEGAL: GR 2922/2007 Nº 14 – ENERO DE 2009

humanos y sus relaciones. Son habituales las estadísticas sobre estudios económicos (modelos de funcionamiento del mercado), demográficos (pirámides de población, tasas de natalidad...), psicológicos (test psicotécnicos), sociológicos (encuestas del CIS)... Otros campos de actuación son las ingenierías (resistencia de materiales, estudio de plantaciones agrícolas, diseño de nuevos productos y tiempos de vida...) o en el comercio (control de calidad, evolución de la producción...).

Es difícil encontrar algún sector ajeno a la Estadística, especialmente con el uso del ordenador y el acceso a los métodos estadísticos para personas no específicamente cualificadas. Así, y por citar otros ejemplos, se utilizan modelos estadísticos para establecer la aparición de números primos en intervalos (Matemáticas), para establecer la autoría de obras (Lingüística), para la identificación de acusados (Derecho), para el tráfico en redes (Informática), para establecer el precio de los anuncios (Publicidad), etc.

El lenguaje estadístico es además muy utilizado en la información de todos los medios de comunicación. Continuamente observamos, en estos medios, tablas estadísticas acompañadas de sus respectivos parámetros estadísticos.

Estos parámetros nos informan sobre la distribución de los datos; proponen un representante de todos los datos, indican si éstos están separados entre sí o, por el contrario, tienden a agruparse alrededor de un valor. Además nos informan sobre la forma global de la gráfica de la distribución.

Por último, puede destacarse las múltiples ventajas que el uso de los potentes ordenadores actualmente disponibles han traído a la recopilación y tratamiento de datos estadísticos. Una simple hoja de cálculo supone un gran ahorro a la hora de realizar las tablas.

7. BIBLIOGRAFÍA

Cochran, W. G. (1986). *Técnicas de muestreo*. México: Continental

Cramer, H. (1977). *Elementos de la teoría de probabilidades*. Madrid: Aguilar

López Cachero, M. (1996). *Fundamentos y Métodos de la Estadística*. Madrid: Ed Pirámide

Quesada, V.; Isidoro, A. y López, L.A. (1989). *Curso y ejercicios de estadística*. Madrid: Alhambra

Ríos, S. (1983). *Análisis estadístico aplicado*. Madrid: Paraninfo

Ríos, S. (1994). *Iniciación a la estadística*. Madrid: Paraninfo

Autoría

- Nombre y Apellidos: SILVIA BORREGO DEL PINO
- Centro, localidad, provincia: I.E.S. ÁNGEL DE SAAVEDRA. CÓRDOBA. CÓRDOBA
- E-mail: DEPIS79@HOTMAIL.COM