

# La estadística en el Bachillerato

Por Ana GONZALEZ ARAGON (\*)

## 1. INTRODUCCION

En los programas de Matemáticas del BUP, aparecen temas dedicados a la Estadística de forma dispersa, inconexa e incompleta. En 1.º de BUP, se inicia el estudio de la Estadística Descriptiva (ordenación de datos, tablas de frecuencias, representaciones gráficas y principales medidas de centralización y dispersión). También se introduce el concepto de Probabilidad de Laplace o a priori. En 3.º de BUP se inicia al alumno en la Teoría de la Probabilidad (variable aleatoria, distribuciones de probabilidad más importantes, distribuciones bidimensionales).

Esta programación no permite obtener una visión global de la Estadística, no responde a ¿qué es la Estadística? y ¿para qué sirve?, preguntas que se plantea cualquier estudiante de esta materia, y además los problemas prácticos que se pueden abordar, son muy parciales (estudio de una muestra, cálculo de probabilidades de sucesos).

Esto tiene dos graves consecuencias. Una de carácter pedagógico, pues se trata de enseñar al alumno unos conceptos que no sabe para qué sirven, ni a que problemas reales responden, y, por tanto, el alumno no tiene una motivación suficiente para aprender esas nuevas ideas. Otra consecuencia es de tipo profesional, ya que el alumno no adquiere una preparación suficiente para enfrentarse con los programas de Estadística, que tendrá en gran parte de los estudios universitarios que pueda realizar (Medicina, Biología, Económicas, Psicología, etc.).

Para resolver estos problemas, es fundamental:

- a) Buscar la interrelación entre todos los conceptos que adquieren en los distintos cursos, e incluso ampliar, aunque de forma poco profunda, con tal de lograr una visión global de los métodos estadísticos.
- b) Trabajar con problemas reales y concretos que faciliten la comprensión de todos esos conceptos.
- c) Adaptar didácticamente este método al Bachillerato.

## 2.A) Visión global de la estadística

La Estadística se ocupa del estudio de los fenómenos aleatorios, es decir, de aquellos fenómenos cuyo resultado es impredecible. Por ejemplo: las alturas de los trabajadores de una fábrica, el n.º de varones nacidos en un determinado año, el n.º de hematies que hay en la sangre, el n.º de piezas defectuosas que se producen en una fábrica, el n.º que sale al tirar un dado, etc.

Nos encontramos con un conjunto de «entes» (personas, tornillos, etc.) sobre los que queremos estudiar

una o varias características. Si medimos o contamos esa característica, obtenemos un conjunto de datos que se llama POBLACION. En el caso de la altura de los trabajadores de un taller, la población estaría formada por el conjunto de las alturas de todos esos trabajadores; en el caso del n.º de varones nacidos en un año, el carácter que estudiamos, el sexo, lo expresamos por V para varón y H para hembra, obteniendo un conjunto de la forma {V, H, V, V, H, H, H,...} que constituiría la población.

En general, nos interesará algo concreto de ese fenómeno aleatorio para tomar una determinada decisión. Por ejemplo, se van a confeccionar «monos» de trabajo y se considera interesante hacer tres tallas diferentes. Entonces, podríamos estudiar la altura media de esos trabajadores, la dispersión o «diferencia» que existe entre esa altura media y el conjunto de las alturas, etc. Podría interesarnos el porcentaje de varones que nacen en un año, o qué posibilidad existe de que al comprar 10.000 tornillos a una fábrica, menos del 50 por 100 sean defectuosos.

Surge ahora un problema. Y es que, en la mayoría de los casos, obtener la población es prácticamente imposible, bien porque ésta sea muy numerosa (y por consiguiente suponga un gasto de tiempo y dinero que pueda eliminar el interés del estudio que ibamos a realizar), o bien porque el estudio de la población implique la destrucción de los objetos que la constituyen. Por ejemplo, imaginemos que nos interesa averiguar la duración media de las bombillas que hemos adquirido en una fábrica; es evidente, que si obtuviéramos la población nos quedaríamos sin bombillas.

Así pues, debemos elegir una MUESTRA representativa de la población, es decir, seleccionar un subconjunto del conjunto de personas u objetos sobre los que se centra el estudio. Obtenida esta muestra trabajaremos con ella como si se tratase de la población.

Aparecen ahora dos nuevos problemas:

1) ¿Cómo elegir esa muestra de modo que sea representativa de la población? ¿Qué tamaño o número de elementos deben formar parte de la muestra?

Es evidente que si trato de hallar la altura media de los españoles y elijo una muestra constituida sólo por niños, no será representativa, y los resultados que obtenga con ella, difícilmente podrán ser válidos para la población.

De la misma forma, si la población consta de 36.000.000 de personas, será mejor en principio, tomar una muestra de tamaño 1.000.000, que una de tamaño 10.000. Sin embargo, nos interesará tomar la

(\*) Catedrática de Matemáticas del IB de Onteniente (Valencia)

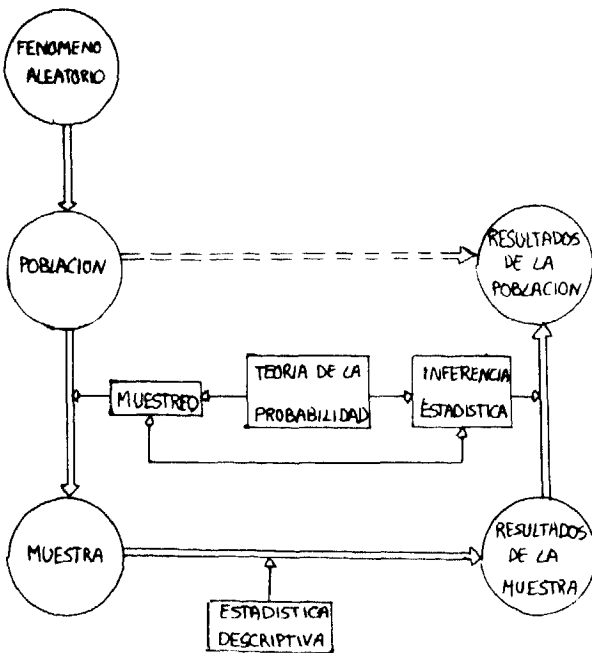
muestra cuyo tamaño sea lo más pequeño posible para evitar gastos inútiles de tiempo y dinero.

Estos problemas los resuelve la **TEORÍA DE MUESTRAS** o **MUESTREO**, que es una parte de la Estadística, basada en gran medida en la **TEORÍA DE LA PROBABILIDAD**.

2) Estudiada la muestra, utilizando la **ESTADÍSTICA DESCRIPTIVA**, obtenemos unos resultados de esa muestra. ¿Serán válidos para la población? Si en el ejemplo anterior, obtuviéramos que la altura media de los elementos que integran la muestra es de 1,70 m., ¿podemos afirmar que la altura media de los españoles es 1,70 m.?

A este proceso de inferir resultados para la población a partir de la muestra, se le llama **INFERENCIA ESTADÍSTICA**. Esta parte de la Estadística está basada, también en la Teoría de la probabilidad.

Resumiendo este proceso, quedaría:



## 2.B) Adquisición del método estadístico a través de problemas concretos

Una forma de adquirir esta idea global de la Estadística, es a través de la resolución de problemas reales y concretos en los que el alumno esté interesado.

Las fases a seguir serían las siguientes:

1. **Elección del problema.**—Se trata de seleccionar un fenómeno o problema aleatorio y algunas características que interese averiguar del mismo. Por ejemplo: Averiguar la edad media de la población activa de la localidad en que nos encontremos.

2. **Muestreo.**—2.1. Elección del tipo de muestreo: Indudablemente el alumno deberá seleccionar una muestra de la población para su estudio.

Hay diversos métodos para obtener la información requerida en una investigación por muestreo. Naturalmente, interesa que el método que elijamos nos proporcione una muestra representativa de la población y, por tanto, conviene hacer muestreos probabilísticos, es decir, muestreos en los que cada elemento

de la población tenga una probabilidad no nula de ser extraído para la muestra. En estos muestreos, es posible con la ayuda de la Teoría de la Probabilidad, estimar la distribución en la población, derivada de la distribución en la muestra. Esto exige que el muestreo sea aleatorio y de ahí que se llame muestreo aleatorio.

Dentro del muestreo probabilístico, existen distintas modalidades de especial interés en las aplicaciones prácticas.

a) **Muestreo aleatorio simple:**

Se caracteriza por la equiprobabilidad de todos los elementos poblacionales, para ser seleccionados. El sistema de elección podría ser mediante las tablas de números aleatorios. Las muestras que se obtienen suelen ser representativas aunque presenta el inconveniente de tener que hacer un listado de la población.

b) **Muestreo aleatorio sistemático:**

Las unidades de la población deben estar numeradas correlativamente. Si la proporción de elementos para la muestra es de 1 por 10, eligiendo aleatoriamente un número comprendido entre 1 y 10, los demás se obtienen de forma sistemática, sumando diez unidades al número anterior.

Es un tipo de muestreo cómodo de realizar, pero como podría constatarse con diversos ejemplos, no siempre da muestras representativas.

c) **Muestreo aleatorio estratificado:**

Las poblaciones muy heterogéneas, se dividen en subpoblaciones o estratos lo más homogéneos posibles, y la muestra se reparte entre los diferentes estratos. Este reparto o afijación puede ser uniforme, proporcional u óptima, según sea, igual entre los estratos, proporcional al número de elementos de cada uno de ellos o teniendo en cuenta, además de los criterios anteriores, la variabilidad o falta de homogeneidad de los estratos.

Se aumenta la precisión en las estimaciones, pero tiene el peligro de que la formación de los estratos no se haga correctamente.

d) **Muestreo por conglomerados o áreas:**

Se divide la población en grupos o conglomerados lo más heterogéneos posible y se seleccionan algunos de ellos, que pasan íntegramente a formar parte de la muestra.

Tiene el inconveniente del muestreo anterior, la formación de los conglomerados, pero presenta la ventaja de disminuir el gasto en desplazamientos y evitar el listado de la población.

e) **Muestreo polietápico:**

Consiste en la aplicación reiterada del muestreo por conglomerados.

Se trata ahora de seleccionar el tipo de muestreo óptimo, para obtener la muestra en nuestro problema particular. Es decir, se trata de ver qué muestreo proporciona una muestra representativa a la vez, que se minimicen los gastos que conlleve su obtención.

2.2. Elección del tamaño de la muestra: En el ejemplo que hemos puesto al principio, se trataba de averiguar una media poblacional, es decir, se intentaba hallar la media de edades de la población activa. El tamaño de la muestra depende de la característica que queremos estudiar (media poblacional) y con qué precisión queremos resolver el problema. Concretamente, un buen estimador de la media poblacional  $\mu$  es la media muestral  $\bar{x}$ . La probabilidad de que pertenezca la media poblacional  $\mu$  al intervalo  $[\bar{x} - e, \bar{x} + e]$  es  $1 - \alpha$  siendo  $1 - \alpha$  una probabilidad que se asigna según el grado de certeza que queramos tener respecto a la conclusión final. El valor de  $e$  depende del

tamaño de la muestra, cuanto mayor sea ésta, menor será el valor de  $e$  y, por tanto, mayor será la precisión del resultado que demos para la población. En la práctica lo que se hace es asignar el valor de  $e$  y deducir qué tamaño de la muestra debemos utilizar, dentro del tipo de muestreo que hayamos seleccionado. Por ejemplo, si se tratara del muestreo aleatorio simple o sistemático:

$$n = \frac{N K^2 S^2}{N e^2 + K^2 S^2} \quad (1)$$

siendo:

$N$  = tamaño de la población.

$e$  = error máximo que estamos dispuestos a admitir al estimar la media poblacional.

$$S^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{N - 1}$$

donde  $x_1, x_2, \dots, x_k$  los diferentes valores poblacionales y  $n_1, n_2, \dots, n_k$  la frecuencias de aparición de los mismos.

Si este dato no se conoce de algún otro estudio (será lo más probable en el caso concreto de los alumnos), se elige una muestra piloto del tamaño  $n'$  que queramos y se calcula  $S'^2$  que sustituiremos en (1) en lugar de  $S^2$ .

$$S'^2 = \frac{\sum_{i=1}^h (x_i - \bar{x})^2 n_i}{n' - 1}$$

donde  $x_1, x_2, \dots, x_h$  y  $n_1, n_2, \dots, n_h$  son los diferentes valores y sus frecuencias en la muestra piloto, y

$$\bar{x} = \frac{\sum_{i=1}^h x_i n_i}{n'}$$

$K$  es el valor correspondiente (según la distribución normal o según la Desigualdad de Chebyscheff) al coeficiente de confianza  $p_k = 1 - \alpha$  que asignamos.

**3. Estudio de la muestra. (Estadística Descriptiva).**—Se hace un estudio estadístico descriptivo de los elementos de la muestra: tablas de frecuencias, representaciones gráficas, cálculo de medidas de centralización y dispersión, etc. Es decir, se estudia en la muestra, todo lo que nos interesa conocer de la población. En nuestro problema concreto interesaría conocer fundamentalmente la media muestral.

**4. Inferencia.**—Obtenidos los datos muestrales, se da una estimación poblacional de la característica que intentábamos conocer. En nuestro caso daremos el valor de la media poblacional, con una probabilidad de que el error que cometemos es menor que una determinada cantidad, o lo que es lo mismo, daremos un intervalo, al que pertenecerá la media poblacional con una determinada probabilidad.

## 2.C) Adaptación didáctica de este método al Bachillerato

**1.º de BUP.**—En este curso, el alumno deberá adquirir una idea clara de fenómeno aleatorio (objeto de la Estadística). Se le ayudará a elegir un determinado problema aleatorio y a plantearse cuestiones de interés acerca del mismo.

Se dará el concepto de población, y será fácil para el alumno, ver la necesidad de trabajar con una muestra.

Se le dará a conocer los distintos tipos de muestreo, haciendo ver, claramente, las ventajas e inconvenientes que presentan cada uno de ellos, usando cuantos ejemplos sean necesarios para lograrlo.

Una vez el alumno decida qué tipo de muestreo es más conveniente para su problema concreto, se planteará el problema del tamaño de la muestra.

En este curso, se deberá hacer de forma intuitiva teniendo en cuenta el tamaño que tiene la población y la facilidad o dificultad que encierre la obtención de los datos.

Una vez obtenida la muestra se hará un estudio de la misma. Para ello se iniciará al alumno en Estadística Descriptiva, aprendiendo las técnicas de:

- Ordenación de datos.
- Tablas de frecuencias.
- Representaciones gráficas.
- Cálculo de las principales medidas de centralización.
- Cálculo de las principales medidas de dispersión.

Inferir los resultados de la muestra, a la población, será un proceso casi natural para el alumno (este tipo de inferencias se hacen continuamente en la vida real). Sin embargo se le hará ver la poca fiabilidad de sus resultados, debido a la «arbitrariedad» cometida al elegir el tamaño de la muestra y la muestra misma. Es decir, que no se tiene una certeza absoluta de que los resultados obtenidos sean válidos para la población y que, por tanto, es necesario saber, qué posibilidad tenemos de que nuestra inferencia sea correcta. De alguna manera, se le introduce ya el concepto de probabilidad y su necesidad para dar una solución aceptable al problema.

**3.º de BUP.**—El proceso a seguir, es el mismo que el iniciado en 1.º de BUP teniendo en cuenta los conocimientos que ya posee de entonces e intentando ya, dar soluciones a las cuestiones que quedaron planteadas.

Se le inicia en la Teoría de la Probabilidad, a través de los conceptos:

- Definición de Probabilidad.
- Variable aleatoria.
- Distribuciones de probabilidad.
- Distribución Binomial y Normal.

A continuación se le pueden dar unas nociones de Teoría de Muestras. Por último, se le introduce en las técnicas fundamentales de la Inferencia Estadística; estimación puntual, estimación por intervalos y contraste de hipótesis estadísticas.

Naturalmente, todo este bagaje teórico deberá estar enfocado a aquellos conceptos o resultados concretos que el alumno va a precisar para resolver el problema concreto que el alumno va a precisar para resolver el problema concreto que tenga planteado, puesto que indudablemente, es imposible que el alumno de 3.º de BUP adquiera una preparación completa acerca del Muestreo y la Inferencia Estadística.

## 3. CONCLUSIONES

Con esta forma de enfocar la enseñanza de la Estadística, creo que se logran dos objetivos importantes. Uno, es estrictamente pedagógico, interesado el



biblioteca del estudiante

- Libros de bolsillo
- Con comentarios de texto

### NOVEDADES:

Marta Portal:

«ANÁLISIS SEMIOLOGICO DE PEDRO PARAMO»

Manuel Machado:

«LA GUERRA LITERARIA»

Tirso de Molina:

«POESIA LIRICA. DELEITAR APROVECHANDO»

### EN LA MISMA COLECCION:

Pedro Salinas:

«TEATRO»

Jorge Guillén:

«ESTUDIOS»

Gerardo Diego:

«ANGELES DE COMPOSTELA. VUELTA DEL PEREGRINO»

Juan Ramón Jiménez:

«CRITICA PARALELA»

Antonio Machado:

«POESIA»

Solicite catálogo



**NARCEA, S. A. DE EDICIONES**  
Dr. Federico Rubio, 89.  
MADRID-20. Tel. 254 61 02

alumno en resolver problemas reales y concretos, necesita unos conocimientos que él mismo va pidiendo. Su aprendizaje obedece a unas necesidades que tiene claramente planteadas, tiene un objetivo útil y concreto. En resumen, posee una fuerte motivación para ir aprendiendo ideas nuevas. Además, su aprendizaje le proporciona una utilidad inmediata y muy concreta, que él mismo ve plasmada en la resolución de su problema.

El otro objetivo importante, es la adquisición de una idea bastante general de la Estadística. Que el alumno sepa qué tipo de problemas pueden resolverse desde un punto de vista estadístico, y en líneas generales, cuál es el procedimiento que debe usarse. Además, ese mínimo de conocimientos estadísticos que ha aprendido, podrá usarlos en su vida cotidiana, pudiéndose enfrentar a numerosos problemas de una forma más científica, y lo que es también muy importante, estará en condiciones de ampliarlos y profundizar de forma exhaustiva, si sus posteriores estudios o su trabajo lo requiere.

#### 4. OBSERVACIONES

Los objetivos fundamentales de este trabajo han sido, plasmar de alguna forma, la interrelación entre los distintos temas estadísticos, y su posible enseñanza a través de problemas reales en los cuales se puede «recorrer» todo el proceso estadístico. Es por ello que no me he detenido en desarrollar ninguna de las partes teóricas del trabajo, ni la resolución detallada de ningún problema concreto. Esto, sería en todo caso, objeto de otros posibles trabajos.

#### REFERENCIAS BIBLIOGRAFICAS

AZORIN POCH, F.: *Curso de Muestreo y aplicaciones*. Aguilar, S. A. 1969.

CALOT, G.: *Curso de Estadística Descriptiva*. Paraninfo, 1970.

CRAMER, H.: *Métodos Matemáticos de Estadística*. Aguilar, S. A., 1953.

GUTIERREZ CABRIA, S.: *Introducción al Cálculo de Probabilidades y a la Estadística*. Facsimil, Valencia, 1974.

GUTIERREZ CABRIA, S.: *Estadística Aplicada*. Facsimil, Valencia, 1976.

MOOD, M. y GRAYBILL, F. A.: *Introducción a la teoría de la Estadística*. Aguilar, S. A., 1972 Ed. Castillo, 1967.

---

## REVISTA DE BACHILLERATO

---

#### REDACCION

- Paseo del Prado, 28, 7.<sup>a</sup> planta. Madrid-14

#### VENTA Y SUSCRIPCIONES

Servicio de Publicaciones M.E.C.

- Ciudad Universitaria, s/n. Tel. 449 67 22
  - Alcalá, 34
  - Paseo del Prado, 28
-