



ESTADÍSTICA

“Hay tres clases de mentiras: la mentira, la maldita mentira y las estadísticas” (atribuida a Mark Twain).

*“Somos víctimas del abuso estadístico por nuestro analfabetismo numérico”
(Douglas Hofstadter)*

1. INTRODUCCIÓN

En la vida nada es seguro. Muchas de las decisiones que tomamos a lo largo de ella están basadas en informaciones incompletas, pero estamos acostumbrados a vivir con un cierto grado de incertidumbre... calculando siempre las posibilidades de obtener un buen resultado para lo que queremos o para lo que deseamos muchas veces. La Estadística como ciencia trata de cuantificar esa incertidumbre y es imposible separarla del cálculo de “Probabilidades”, se trata de cuantificar el azar.

Se cree que la *estadística descriptiva* (aquella que sólo trata de describir y analizar una serie de datos sobre un fenómeno determinado) tiene su origen unos dos mil años antes de Cristo en Egipto, China y Mesopotamia, donde ya se elaboraban censos para la recogida de tasas entre la población. Los egipcios usaban el “Nilometro”, aparato que estimaba el caudal del Nilo y a partir del cual definían un índice de fertilidad y la cantidad de impuestos que le correspondía. (La imposibilidad de fijar el clima aseguraba una cierta incertidumbre).

A diferencia de la anterior, la *estadística inferencial* trabaja con “muestras” y pretende generalizar los resultados obtenidos al resto de la población (inferir). Por ejemplo:

- Cambiamos de dentífrico porque según una encuesta 2 de cada 3 dentistas garantizan ausencia de caries.
- Modificamos nuestra ruta porque según el periódico los lunes hay atasco habitualmente y tendremos un alto porcentaje de probabilidad de llegar tarde.
- Se toman medidas de prevención ante una posible enfermedad porque estadísticamente “está demostrado” que corremos menos riesgo de contraerla.
- En campaña electoral se cambian estrategias basándose en resultados de continuas encuestas de opinión.



Pero... ¿nos podemos fiar de toda la información “estadística” que recibimos? Estamos inmersos en la llamada “era de la información” y gran parte de esa información puede ser fácilmente manipulable. Las estadísticas forman parte de esas posibles manipulaciones.

El desconocimiento de los números que posee gran parte de la población y la idea generalizada de que los números no mienten, hace que poca gente se dé cuenta de cuándo se le está tratando de engañar o los datos no son del todo fiables. En un principio, se suelen aceptar como incuestionables las conclusiones estadísticas sin hacer la más mínima crítica, y cuando leemos o escuchamos frases como “según las estadísticas...”, “estadísticamente está demostrado...”, “hay un porcentaje de...” inmediatamente se supone que es cierto. Sin embargo, según vamos madurando nos vamos damos cuenta de la cantidad de falsedades y engaños que nos han hecho llegar basándose en las supuestas “estadísticas”, con el objetivo de transmitir mensajes y conseguir objetivos más o menos lícitos. No es raro escuchar interpretaciones totalmente diferentes de unos mismos resultados estadísticos, dependiendo de quien esté informando. Llegamos a pensar que si con las estadísticas se puede probar cualquier cosa debe ser que no se puede probar nada. Pasamos de una fe ciega en los números a pensar que lo único que pueden hacer es engañarnos.

La verdad es que generalizar opiniones, propiedades, etc. al conjunto de la población, partiendo de las informaciones de una muestra, no es tarea fácil, por eso cuanto más grande y más representativa sea la muestra, más fiables serán los datos y ahí está la ardua tarea de los estadísticos ya que se tiende a trabajar con muestras pequeñas para, entre otras cosas, abaratar su coste.

El mal uso de las estadísticas tiene un importante potencial, pueden ser utilizadas para provocar sentimientos no siempre positivos entre la gente, distorsionando nuestra visión del mundo, e incluso a veces aprobar medidas gubernamentales que en principio rechazaríamos, en aras de datos estadísticos que avalan que son totalmente necesarias.

Pero tenemos que tener una idea clara: la estadística no engaña ni miente, engaña quien hace un mal uso de ella.

En numerosas ocasiones el objetivo es simplemente alimentar cierta corriente de opinión o llamar la atención sobre algo muy determinado y los datos que se dan no son significativos. Veamos algunos ejemplos:



Si estamos hablando de un problema social, una de las primeras preguntas que nos hacemos es a cuánta gente afecta. La cuantificación es inevitable y útil, pero se tiene que hacer y transmitir bien. Si leemos por ejemplo que 300 mujeres en Madrid son diagnosticadas mensualmente de osteoporosis, esa cifra aislada no nos dice nada si no sabemos ni las edades, ni el número total de mujeres censadas en la Comunidad de Madrid, ni estadísticas anteriores que nos permitan ver la evolución y comprobar si hay una mayor incidencia. A veces titulares de números absolutos lo único que pretenden es impresionar a la audiencia, pero dados de otra forma ya no serían titulares. Como se ve, necesitamos siempre más información sobre el contexto del que estamos hablando para poder valorar la noticia.

El ejemplo inverso lo vemos con el porcentaje. Si leemos que en determinada zona han aumentado los atracos en un 60% con respecto al año pasado nos impacta más que si nos dicen que han pasado de 6 a 10 atracos. Y en ambos casos la información no mentiría...

Sin embargo, existen dos cuestiones con las que hay que tener más cuidado. Una de ellas se refiere a cómo puede variar el resultado en el momento que modifiquemos levemente la definición de lo que estamos contando. Por ejemplo, el “umbral de la pobreza” dependerá de lo que se considere como importe mínimo de euros de que dispones por día, el resultado del “IPC” vendrá dado por los conceptos que contabilicen para calcularlo... en la mayoría de las ocasiones se adecua la definición al objetivo a lograr, es decir, seleccionamos aquellos valores con los que obtenga el resultado deseado.

La segunda es que los porcentajes de cualquier estadística siempre deberían darse junto con el tanto por ciento de seguridad que se tiene del estudio (“nivel de significación”) y con el error en más/menos del resultado (“intervalo de confianza”).

Por ejemplo, suponemos que en una muestra seleccionada de 1.000 personas sobre una población de 300.000, el salario medio de la muestra es de 600 euros y su desviación típica 30 €. Pues bien, sabemos que el sueldo de aproximadamente el 90% del total de la población estará entre +/- 3 veces la desviación típica¹, es decir, ganarán entre $600-90=510$ y $600+90=690$ euros. En caso de la desviación típica fuera 150, el 90% estaría entre 450 y 750, lo que no nos dice mucho.

¹ Desigualdad de Chebycheff



Veamos esta noticia: “los dentistas prefieren en un 70% el dentífrico marca Alfa”. En teoría lo que se ha dicho, estadísticamente hablando, es que el 70% de los dentistas consultados de una cierta muestra representativa prefieren esa marca. Ante esto la noticia, para que fuera más precisa, debería decir, por ejemplo: “con una seguridad del 90% (*nivel de significación*), entre el 65 y el 75% de los dentistas prefieren el dentífrico marca Alfa” (*intervalo de confianza +/- 5%, error total del 10%*).

Otro ejemplo: el índice de error puede hacer variar la bondad o no de una noticia. Imaginemos que nos dicen que el IPC ha pasado en porcentajes del 2,5 al 2,3. Está claro que es positivo porque ha bajado, pero si el intervalo de confianza es de más/menos 0,3%, este error puede hacer inexistente esa descenso (puede ir del 2 al 2,6%, por lo que ha podido subir).

Pero, ¿una persona corriente puede llegar a descifrar las claves ocultas de la estadística? Pues, con un poco de práctica podemos aprender a hacernos las preguntas adecuadas para entenderla:

- ¿cómo se ha definido lo que se está contando?
- ¿cómo se ha medido o contado?
- ¿se ha medido en toda la población afectada o sólo en una muestra?
- ¿es la muestra representativa?
- ¿sabemos qué índice de error tiene?
- ¿es lógico el resultado?
- ¿qué datos tengo del contexto?, ¿de sus antecedentes?
- ¿cuál es el cuestionario?, ¿qué se ha preguntado?, ¿cómo han sido las preguntas?

Volviendo al ejemplo de los sueldos medios, suponemos que nos dicen:

“El salario medio de los trabajadores de una empresa es 1.100 €”

Nos deberíamos preguntar:

- ✓ ¿A qué se refieren con “salario medio”? ¿es el salario base?, ¿se han incluido otros complementos como antigüedad, horas extras, etc.?



- ✓ ¿Cómo se ha medido? ¿Preguntando directamente a cada persona? ¿Se ha utilizado información dada por la empresa?
- ✓ ¿La media se refiere a la totalidad de los salarios? ¿Han escogido solo a algunas personas?
- ✓ Si ha sido así, ¿la elección es representativa de todo el colectivo?, ¿se ha escogido basándose en información previa que nos ayude a obtener el resultado que deseamos?
- ✓ ¿Nos dan información del índice de error? ¿Cuál es su desviación típica?
- ✓ ¿Es lógico el resultado obtenido? o, ¿nos extraña esa “media”?
- ✓ ¿Conozco el contexto del estudio? ¿Qué clase de empresa es? ¿Son trabajadores cualificados? ¿Mano de obra sin cualificar?
- ✓ ¿En algún sitio puedo ver el tipo de preguntas que han hecho? ¿cómo han sido redactadas?

Pero lamentablemente, a veces no nos dan la información suficiente para responderlas y eso mismo nos tendría que dar una idea de la fiabilidad de los resultados. Un buen estudio estadístico debería ser transparente, dando a conocer cómo se ha realizado, la definición de los fenómenos, el procedimiento de medición y la elección de la muestra. Saber todo ello nos ayudaría a hacer un juicio de valor mucho más acertado e incluso a criticarlo.

Como escribe el matemático A.K. Dewdney: “Aquellos que abusan de las matemáticas también abusan de nosotros. Nos convertimos en presas de las triquiñuelas comerciales, las estafas financieras, la charlatanería médica y el terrorismo numérico de los grupos de presión, todo porque somos incapaces (o no estamos dispuestos) a pensar con claridad durante unos momentos”².

² <http://www.csd.uwo.ca/faculty/akd/akd.html>



2. ¿CÓMO SE CONFECCIONA UNA ESTADÍSTICA?

2.1 Obtención de datos

La materia prima de la Estadística son los datos. Partiendo de información obtenida por diferentes vías: encuestas, datos empíricos resultado de la repetición de un mismo suceso o experimento,... intenta extraer conclusiones. Su finalidad es, como ya se dijo, “cuantificar la incertidumbre”, establecer una serie de leyes, reglas fijas, normas a partir de las que generalizar posibles resultados en nuevas observaciones o que se prediga lo que posiblemente va a pasar.

Estos datos se obtienen mediante la realización de encuestas, es decir, preguntas realizadas a muchas personas sobre algo determinado, aunque en la industria (resistencia de materiales), demografía y biología (búsqueda de causas), se utiliza la recogida de datos obtenidos como consecuencia de la repetición de un experimento u observación aleatoria.

A la hora de diseñar una encuesta deberíamos tener en cuenta varios aspectos, como:

- La información que obtengamos tiene que tener unos resultados que se puedan expresar mediante números, es decir, que se pueda tabular.
- El tema en una encuesta debe ser claro y concreto; la información debe obtenerse mediante preguntas con un número limitado de respuestas.
- A la hora de realizar las preguntas el entrevistador deberá ser lo más imparcial posible para no influir en las respuestas.

Pero ¿a cuántas personas debo entrevistar o cuántas veces tengo que realizar el mismo experimento para obtener resultados representativos? Lo más cercano a la realidad sería contar con datos referentes a toda la población a la que afectara el carácter que quiero investigar, aunque eso resultaría poco práctico, muy costoso y a veces imposible³.

³ Por ejemplo, en determinadas grandes tiendas, nos muestran una puerta que mediante un mecanismo se abre y se cierra de forma constante, con objeto de demostrar la durabilidad y resistencia de las bisagras. Es evidente que este proceso no puede ser aplicado a la totalidad de la fabricación.



Para evitarlo se selecciona una muestra. Ésta, ante todo debe ser representativa y, por lo tanto, no sirve cualquier porción de la misma. Por ejemplo, en algunos casos habría que dividir una población en varias subpoblaciones que sean representativas, lo que se llama “muestreo estratificado”.

2.2 Organización de Datos – Estadística Descriptiva

Para poder trabajar de forma más cómoda con los datos obtenidos los ponemos en tablas, los organizamos (“tabulamos”) y les asignamos una serie de parámetros estadísticos (media, mediana...). De esto se encarga la Estadística descriptiva.

2.2.1 Tablas y Gráficos

Podemos empezar resumiéndolos en una tabla de frecuencias. Sobre una línea colocamos todos los valores que puede obtener la variable de estudio (peso de un grupo de personas, edad a la que empezaron a trabajar, valoración que hacen de un determinado político o personaje público, etc.) y los dividimos en intervalos de igual tamaño. Su FRECUENCIA es el número de anotaciones que hay en cada intervalo. La FRECUENCIA RELATIVA es la razón entre la anotación de cada intervalo y el total de datos recogidos.

El porcentaje de veces que aparece un determinado valor se obtiene multiplicando su frecuencia relativa por 100.

TABLA DE FRECUENCIAS				
peso en kg	40-50	50-60	60-70	70-80
nº personas	5	7	11	2
Frecuencia Número de personas que hay en cada intervalo de pesos				
Frecuencia Relativa nº total de casos : 25				
40-50 kg	5/25=0,20	20%		
50-60 kg	7/25=0,28	28%		
60-70 kg	11/25=0,44	44%		
70-80 kg	2/25=0,08	8%		



Para hacer los intervalos debemos tener en cuenta:

- Que éstos sean preferiblemente iguales y tengan su punto medio en números apropiados, es decir, cómodos a la hora de trabajar con ellos.
- Si hay pocos datos, utilizaremos pocos intervalos.
- Si los datos son muy numerosos es conveniente usar muchos intervalos.

La tabla de frecuencias nos muestra los datos que hay alrededor de cada valor.

A partir de ellas puedo construir histogramas. Una barra cubre un intervalo y tiene su punto medio en el centro de éste, de ahí la importancia al diseñar su amplitud. La altura de la barra representa la cantidad de puntos u observaciones de cada intervalo.

Los datos de una tabla de frecuencias se pueden presentar mediante diferentes representaciones gráficas o diagramas. En “Tablas y Gráficas” aparecen descritas, por lo que aquí no nos detendremos a hacerlo. Sólo mencionaremos los tipos:

- Pictogramas.
- Diagramas de sectores.
- Diagramas de barras o histogramas.
- Polígonos de frecuencias.

Cada uno de ellos tiene unas características y permite destacar diferentes aspectos.

2.2.2 Parámetros estadísticos

A partir de aquí vamos a pasar de los gráficos a las fórmulas. Otro de los fines de la Estadística Descriptiva es resumir grandes cantidades de datos en unos pocos números que nos proporcionen una idea, lo más clara y aproximada posible del comportamiento general de todos los individuos o elementos que constituyen la población objeto de estudio.

En todo conjunto de datos encontramos dos propiedades principales que constituyen dos tipos de parámetros estadísticos:

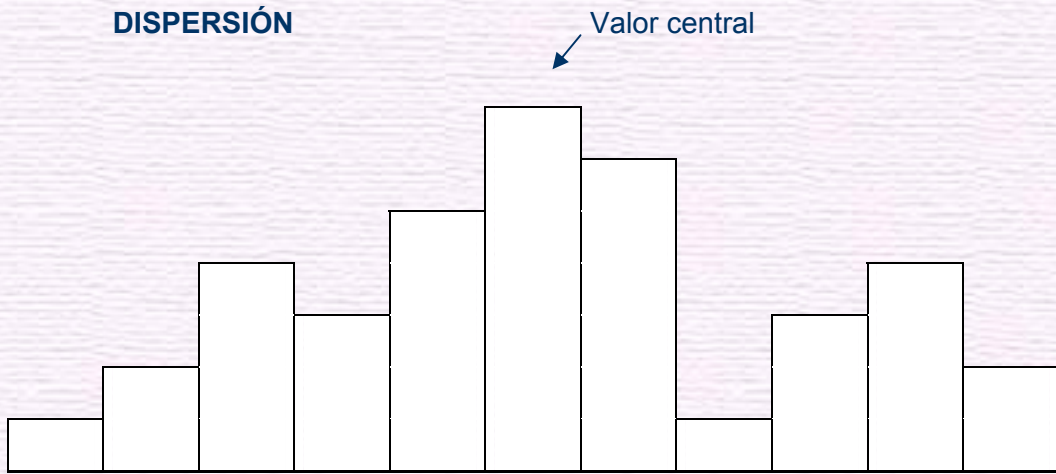
- El *valor central o típico*, cuyo objetivo es agrupar los datos correspondientes a toda la población en un solo valor numérico que represente al conjunto total.



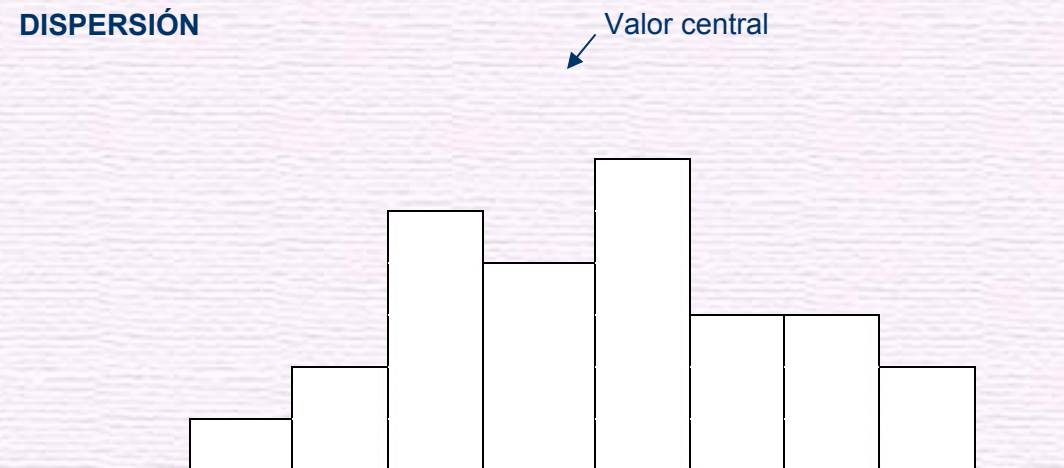
- La *dispersión* de este valor central que nos da idea de la distancia que presentan los datos con respecto al valor que hemos tomado como central.

Lo anterior lo podemos observar en los siguientes histogramas:

**MUCHA
DISPERSIÓN**



**POCA
DISPERSIÓN**





A) Parámetros centrales: su objetivo es agrupar o concentrar la información recogida con respecto a toda la muestra en unos pocos números que representen al total. Son:

- Media
- Mediana
- Moda

La **Media** de un conjunto de valores es el resultado de realizar la suma de todos ellos y después dividir por el número de datos que tenemos. En la vida cotidiana se utiliza esta herramienta con frecuencia. Para saber cuál es el gasto medio de un servicio (calefacción, teléfono, alimentación...) recurrimos a la media de lo que hemos gastado en diferentes meses; los escolares calculan si van a aprobar una asignatura calculando la media de las notas obtenidas en diferentes ejercicios...

Este parámetro estadístico presenta ventajas e inconvenientes. Entre los primeros están que en su cálculo se recogen todos los datos con los que se está trabajando y que los cálculos para obtenerla son sencillos. En cuanto a los inconvenientes, destaca el hecho del efecto que sobre el valor obtenido ejercen los datos extremos. Por eso a veces la media no es representativa, no informa de nada. No hay más que pensar en los valores que, según las estadísticas, alcanzan los sueldos medios en nuestro país y compararlos con el nuestro, o calcular la cantidad de euros que gasta alguien en lotería, desde el punto de vista de un no jugador.

Para solucionar estos fallos se recurre a otros parámetros centrales, que a veces son más representativos: la moda y la mediana.

La **Moda** de un conjunto de datos es aquel valor con mayor frecuencia. No es tan representativo como la media, pues en su cálculo no intervienen todos los datos recogidos. No tiene demasiado interés salvo cuando esta frecuencia se destaca claramente de la del resto de los valores. Cuando hay una frecuencia, por ejemplo una calificación de 6 en una evaluación, que se repite muchas más veces que el resto, la moda sí que da información.

La **Mediana** es el valor central de la distribución, cuando hemos podido ordenar los valores que la componen. Es el valor que deja la mitad (50%) a la izquierda y la otra mitad a la derecha (percentil 50). Si el número de éstos es par la mediana será la media aritmética de aquellos dos que ocupen el lugar central. Se utiliza cuando trabajamos con datos cualitativos



susceptibles de ser ordenados; si en el conjunto de frecuencias hay valores muy alejados de la media, lo que afecta mucho a ésta o si se está trabajando con datos estadísticos agrupados en clases y las que se sitúan en los extremos son abiertas (sueldos mayores de 3.000 € al mes, extensión mayor a...).

De acuerdo con los siguientes datos, donde quedan expresadas las estaturas aproximadas de 4.350 estudiantes, veremos los valores que toman los tres parámetros anteriores:

estatura (metros)	1,52	1,56	1,60	1,64	1,68	1,72	1,76	1,80	1,84	1,88
nº estudiantes	62	186	530	812	953	860	507	285	126	29

Altura (m)	nº estudiantes	Altura x nº est.
1,56	186,00	290,16
1,60	530,00	848,00
1,64	812,00	1.331,68
1,68	953,00	1.601,04
1,72	860,00	1.479,20
1,76	507,00	892,32
1,80	285,00	513,00
1,84	126,00	231,84
1,88	29,00	54,52
	4.288,00	7.241,76

$$\text{Media} = \frac{\text{Altura x nº estudiantes}}{\text{nº total estudiantes}} = 1,69 \text{ m}$$

$$\text{Mediana} = \text{Valor intermedio} = 1,72 \text{ m}$$

$$\text{Moda} = \text{Valor con mayor frecuencia} = 1,68 \text{ m}$$



B) Parámetros de dispersión: pretenden mostrar la distancia, los que los datos reales se alejan de los valores que hemos calculado como centrales o representativos. Hablaremos de:

- Rango
- Desviación
- Varianza
- Desviación Típica

Como su nombre indica y ya hemos visto, los parámetros centrales intentan agrupar los resultados de una serie de datos en unos números que generalicen las características de todo el grupo. Pero con ellos perdemos la información de cómo se encuentran distribuidos estos valores en torno al central, es decir, a cuánta distancia de él se encuentran. Esta información la obtenemos mediante los parámetros de dispersión. Hay varias formas de medirla.

El **Rango**, o recorrido de un conjunto de datos estadísticos, es la diferencia que existe entre los valores extremos. Si el rango es grande, encontraremos valores que estén muy alejados unos de otros y por tanto los valores centrales no serán muy representativos. Esto no sucede cuando el rango es pequeño. Los valores no están muy alejados entre sí ni de los centrales por lo que puede suponerse que sí representa a toda la muestra

La **desviación** de un valor cualquiera con respecto a la media es la diferencia entre este valor y la media. Esta operación puede tener como resultado un número positivo, negativo o ser cero.

Vale como ejemplo esta tabla donde indica cuántas horas de televisión ven a la semana 5 personas diferentes:

PERSONA	1	2	3	4	5
Nº HORAS	5	7	3	33	12

$$\text{Valor de la Media} = \frac{5 + 7 + 3 + 33 + 12}{5} = 12 \text{ horas}$$



Rango: diferencia entre los valores extremos >> 33-3=30

Valor-media	Desviación
5 - 12 =	-7
7 - 12 =	-5
3 - 12 =	-9
33 - 12 =	21
12 - 12 =	0

Con el estudio de las desviaciones de cada una de las desviaciones por separado ya puedo tener una idea acerca de la desviación de los datos con respecto a la media (si están muy alejados o muy cercanos a ésta).

Ya que la desviación tiene valores negativos y positivos, al sumarlos para calcular su media se anularían unos con otros, perdiendo la información sobre las desviaciones antes mencionadas. Para que esto no ocurra y obtengamos un valor positivo, recurrimos a los cuadrados de las desviaciones, lo que da lugar a la **varianza** (σ^2). Ésta presenta el inconveniente de que sus resultados no aparecen en las mismas unidades de los datos con los que estoy trabajando, sino con ellos elevados al cuadrado (en el ejemplo, la varianza se daría en horas al cuadrado).

$$\sigma^2 = \frac{(5 - 12)^2 + (7 - 12)^2 + (3 - 12)^2 + (33 - 12)^2 + (12 - 12)^2}{5} = 60 \text{ horas}^2$$

Como al elevar al cuadrado, tanto las unidades como el resultado están distorsionados, extraeremos la raíz cuadrada y utilizaremos el valor positivo de ésta para obtener unidades del mismo orden de los datos originales (que no estaban al cuadrado). Esto se llama desviación típica.

La **DESVIACIÓN TÍPICA** (o estándar) que mide la dispersión de los datos alrededor de la media y es la raíz cuadrada positiva de la varianza en cada uno de los datos, lo que se hace para evitar el inconveniente dicho antes de la pérdida de unidades

$\sigma = 7,7 \text{ horas}$



Así, si tenemos una media de 12 horas y una desviación típica de 7,7 horas, se puede concluir que hay mucha desviación/dispersión con respecto a la media.

La relación de este parámetro junto con la media nos da una información más fiable de la dispersión de los valores obtenidos con respecto a esta última.

“Entre los hombres se dan más genios y más monstruos, más gigantes y enanos, más superdotados y más deficientes, es decir, más casos extremos que entre las mujeres, cuyos rasgos tienden a agruparse algo más en torno a unas tendencias centrales...”, (Extraído de un informe “Sobre la forma de conducir” del Prof. Pinillos, Catedrático Psicología Universidad Complutense, Madrid).

Suponiendo que el anterior párrafo fuera correcto y sin entrar a valorarlo, podríamos decir que la distribución correspondiente a los hombres tiene mayor desviación típica que la de las mujeres, puesto que en los primeros el número de casos extremos es mayor.

2.3 Predicción sobre los datos - Estadística Inferencial

Por último, aunque esto se nos sale de nivel, citaremos el objetivo de la **inferencia estadística**, ciencia que extrae conclusiones a partir de datos concretos, basándose en el cálculo de probabilidades⁴. Interpreta los resultados de una encuesta para inferir propiedades del conjunto global de sujetos estudiados, sobre la base de una muestra que debe cumplir una serie de características en cuanto a su amplitud y al modo en que cada elemento ha sido seleccionado que asegure su representatividad con respecto a la población total.

Únicamente nos gustaría mostrar un elemento del que deberíamos disponer cuando se nos quiera transmitir información estadística, es la FICHA TÉCNICA con la que se ha realizado la encuesta que ha permitido obtener esos datos que se nos hacen llegar.

⁴ El juego ha sido un elemento generador de múltiples reflexiones e investigaciones a lo largo de la historia de la Humanidad, al intentar asegurar lo más posible el éxito de los resultados. La probabilidad, el estudio formal de las leyes del azar, se ha desarrollado con objeto de predecir los resultados de estas actividades, que en principio no parecen sujetos a ningún tipo de reglas. Las leyes que rigen los fenómenos del azar son aplicables a los hechos recogidos como sucesos de una observación estadística.



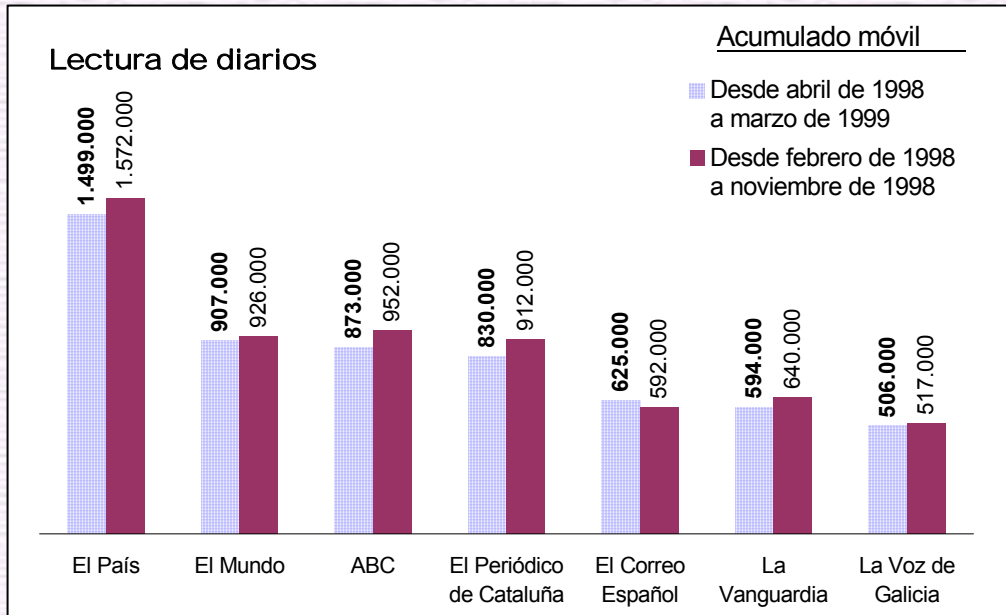
A continuación vemos un ejemplo de Ficha Técnica de una encuesta realizada por la Dirección General de Tráfico sobre “Opiniones sobre la Seguridad Vial de los Ciudadanos”, en diciembre 2005.

- Proyecto y Dirección: Dirección General de Tráfico (Ministerio del Interior)
- Técnica de investigación: Entrevista telefónica asistida por ordenador.
- Ámbito geográfico: España
- Universo: Población de 16 y más años
- Número de entrevistas (muestra): 1.800 entrevistas
- Error de la muestra: +/- 2,5% para un nivel de confianza del 95%
- Tipo de muestreo: estratificado por Comunidad Autónoma y dimensión del municipio. Selección aleatoria de los hogares y selección de la persona a entrevistar mediante cuotas cruzadas de sexo y edad de acuerdo con la distribución real de la población.
- Trabajo de campo: del 19 al 22 de diciembre 2005

Como conclusión podríamos comentar que la inferencia estadística siempre tiene un punto de subjetividad que viene dada por el objetivo que tiene (y quiere hacernos llegar) aquel que hace el análisis, por lo que siempre hay que prestar atención de cuál es la fuente de la conclusión a unos determinados datos, porque ya se sabe...“Interprete los sondeos como le dé la gana, sea feliz” (Alsina)

Incluso con los datos fijos, las interpretaciones que se pueden dar son muy dispares, basta recordar las noches electorales o el siguiente ejemplo donde cada periódico interpreta el número de ejemplares vendidos:

ESTUDIO GENERAL DE MEDIOS DE PRENSA Y RADIO



Fuente: diario "EL País", 13/04/99

Diferentes titulares según quien los escribe:

- "EL PAÍS sigue líder pese a bajar la lectura de los diarios". (*Diario "El País"*).
- "EL MUNDO se consolida como el 2º diario de mayor difusión en España". (*Diario "El Mundo"*).
- "EL PERIÓDICO DE CATALUÑA se reafirma como líder en Cataluña". (*Diario "El Periódico de Cataluña"*).
- "Retroceso generalizado de los periódicos diarios". (*Diario "ABC"*).
- "Pérdida de lectores de la prensa de información general". (*Diario "La Vanguardia"*).





3. BIBLIOGRAFÍA

SANCHÍS, C.; SALILLAS, J.; RIERA, T. ; FONTANET, G.; (1986): *Hacer Estadística* Madrid: Alhambra.

GONICK, L.; SMITH, W.;. (1993): *La Estadística en cómic*. Barcelona: Zendera Zariquiey.

CENTRO NUEVO MODELO DE DESARROLLO; (1994): *NORTE-SUR: La fábrica de la pobreza*. Madrid: Editorial Popular.

LAFUENTE,A.; GALLEGO, F.; SECO, U.; (1999): *Matemáticas. Educación Secundaria de Adultos*. Sevilla:MAD.

ALVIR. P. Y COLABORADORES; (2006): *La Comunidad de Madrid en cifras 2006*. Madrid: Comunidad de Madrid.

www.malaprensa.com, *Errores y chapuzas de la prensa española: números equivocados, gráficos incorrectos, fallos lógicos, conceptos erróneos...*

www.atlas.org.ar (12/02/06)

LUQUE SERRANO, B.; “Mentiras, pecados y abusos estadísticos”, tomado de la página web www.matap.dmae.upm.es

LACOURLY, N; (Julio 2000): *Una pequeña historia de la Estadística*.