

VARIAS DEFINICIONES DE LA ESTADÍSTICA

A. Zavrotsky

Facultad de Ingeniería
Universidad de Los Andes

Resumen.- El autor realiza un recuento histórico acerca de las definiciones de la Estadística. Comienza con H. Conring, quien en el siglo XVII, parece haber introducido la palabra alemana *Statistik* (estadística) basada en la raíz *Staat* (estado). Un diccionario inglés, del siglo XIX, menciona las aplicaciones de la estadística como su definición —colección de hechos relacionados a la condición de un pueblo— esta concepción a la cual el término debe su nacimiento. En este artículo se ofrece una definición, más acorde con todo lo que abarca el concepto: “ la estadística es la ciencia de los conjuntos empíricos” ;el autor genera una explicación minuciosa de los términos **conjunto y empírico** ilustradas con interesantes ejemplos.

Inesperadamente, en muchos casos resulta más fácil el adquirir un dominio relativamente sólido de una disciplina que dar una definición concisa de la misma. Esto hace pensar que para comprender cierta doctrina no es tan esencial el conocer la definición de la misma, y aún que tal definición no siempre es posible del todo. Tal es el caso aún de la ciencia exacta: la Matemática, pues entre los más conocidos matemáticos no hay unanimidad acerca de la definición de la misma. (Bell, 1938) cita varias definiciones y prefiere la de Bertrand Russell: "La Matemática suele definirse como materia en la cual nunca sabemos de qué hablamos, y no sabemos si lo que decimos es cierto". Naturalmente, Bertrand Russell lo dijo en broma, pues una oración en que todos los verbos son negativos, no puede definir nada: sólo

puede descartar lo que no es matemática, pero no puede afirmar qué es matemática: en los términos más técnicos, expresa una condición necesaria, pero no suficiente para que cierta tesis pueda incluirse en el dominio de la Matemática. en la vida común con tanta frecuencia oímos a unos oradores que no saben ni de qué hablan, ni si lo que dicen es verdad, que nos veríamos precisados a englobar todo galimatías en el dominio de las matemáticas. Como otros dos ejemplos, cítense: la definición de la probabilidad: "Probable es todo suceso cuyo contrario no es imposible, y que tiene de su lado más que la mitad de la verdad", insatisfactoria a pesar de llevar la firma venerable de Kant, y la definición de la Trigonometría: "La Trigonometría es el estudio de las relaciones entre las razones trigonométricas", que hace pensar en la célebre explicación de Molière: "El opio hace dormir porque tiene una virtud dormitiva" (Bruno,1930).

La estadística no representa en este sentido una excepción, y la abundancia de buenos textos de estadística contrasta con la escasez de definiciones aceptables de la misma. Las definiciones más antiguas se basan en la etimología, de la palabra **estadística** derivada del **estado**, pues el primero que la introdujo en la forma alemana *Statistik* parece haber sido H. Conring en el siglo XVII, basada (Johnson, 1962) en la raíz *Staat*. el término español **Estado** dice, sin embargo, mucho más que su homólogo alemán *Staat*, y aquí puede designar una entidad política, independiente o no, pues tanto puede hablar del **Estado Venezolano** en general, como del **estado Mérida**. Un diccionario inglés del siglo XIX define: **Estadística**, como primera entrada, "colección de hechos relativos a la condición de un pueblo, una nación o un país, en salud, longevidad, economía doméstica, etc.", y como segunda entrada, "la Ciencia que trata de estos sujetos" (Johnson,1962) . Aquí no se menciona sino una aplicación de la estadística: la original e históricamente la más importante,

a la cual el término debe su nacimiento, de la misma manera, como la **Geometría** en el original griego no significa otra cosa sino la medición de la tierra, y el **álgebra** en árabe es la reducción de los términos semejantes en ambos miembros de una ecuación (Eves, 1953). Pero desde el punto de vista moderno no se puede limitar a la Estadística a los susodichos temas demográficos y económicos, como la agrimensura no pasa de ser una de tantas aplicaciones de la geometría, y la reducción de los términos semejantes no es la única operación algebraica.

Otro diccionario, también inglés, del año 1919, adopta un punto de vista algo más general, pero todavía, demasiado restringido:

"**Estadística**: como usado antiguamente, era aquel ramo de la ciencia política que trataba de la colección, clasificación y discusión de los hechos (especialmente de índole numérica) que se refieren a la condición de un estado o de una comunidad. En los tiempos recientes, es la parte de los estudios que tiene por objeto la colección y arreglo de los hechos numéricos relativos a los asuntos humanos o a los fenómenos naturales" (Johnson, 1962).

Ni aún esta extensión agota la variedad de las aplicaciones modernas de la Estadística, que actualmente presta sus servicios, por ejemplo, hasta a la **Reina de las ciencias exactas**: Teoría de los Números, en los problemas tales como la distribución de los números primos, o de los dígitos decimales en el desarrollo de p , de e , ó de otras constantes. Sólo se puede concluir que una definición de la estadística no necesita referirse del todo al objeto al cual ésta se aplique.

Otros textos contemporáneos pecan por tratar de limitar la estadística a la Teoría del Muestreo, como por ejemplo:

"La Estadística Inferencial se refiere al proceso de usar los datos para sacar decisiones del caso más general del cual estos datos forman parte" (Mood,1950).

No hay razones para creer que toda estadística sea selectiva, pues por ejemplo un censo de población trata de abarcar la totalidad de los objetos del estudio y si el resultado está sujeto a los errores de observación, lo mismo sucede también con la medición de un objeto físico. El mismo texto define también, a la estadística (Mood,1950) como "Tecnología del método científico"; expresión ésta demasiado vaga y poco clara.

Desde nuestro punto de vista, el objeto de la Estadística puede definirse en dos palabras: La Estadística es la ciencia de los **conjuntos empíricos**. Cada una de estas dos palabras requiere una explicación minuciosa. **Conjunto**, porque las conclusiones estadísticas nunca pueden basarse en un caso aislado. Si cierto cráneo hallado en Neanderthal mide tantos centímetros de largo y tantos de ancho, esto es dato valioso, pero no en el sentido estadístico.

Hay que disponer de cierto número; de un conjunto de tales cráneos, comparables en algún sentido, para poder someterlos a un cálculo estadístico: promedio, dispersión, asimetría, plurimodalidad, homogeneidad de una sola característica, correlación entre dos características distintas, etc. (Von Mises, 1951) cita un caso curioso del abuso del método estadístico por falta de cumplir con este requisito: asociaciones provocadas por la palabra latina ALIQUID.

Un conocido sociólogo cuyo nombre callamos por compasión, sometió a pruebas caracterológica a cuatro niños y pretendió descubrir cuatro tipo de reacciones. Naturalmente, no pudo encontrar más de cuatro!. Ya que cada tipo estaba representado por un solo individuo, aquí no había conjunto, luego no había estadística. Un crítico objetó

que en este estudio figuraban en realidad no cuatro sino cinco niños, pues el mismo profesor se había portado como un niño.

Y se dice **empíricos**, por tratarse de los conjuntos, sujetos a la **observación** y no al **cálculo**. Sería tal vez preferible ilustrar esta distinción por medio de varios ejemplos paralelos:

<ul style="list-style-type: none">• ¿Cuántos eclipses del sol sucedieron durante el siglo XIX? Por cuanto existe una fórmula matemática exacta que permite calcular las fechas de todos los eclipses pasados, presentes y futuros, éste no es problema estadístico. Se podría formular la misma pregunta acerca de cualquier siglo futuro, y resolverla con la misma facilidad.	¿ Cuántos terremotos sucedieron durante el mismo período? Aquí ya no se conoce un método tan sencillo, y no hay más remedio que revisar los registros históricos, éste es problema estadístico . Si la ciencia llaga algún día a predecir los terremotos con la misma precisión como lo hace para los eclipses, el problema pasará del dominio de la Estadística al del Análisis Matemático (Wauch,1952).
--	--

(Sigue...)

(Continuación...)

<ul style="list-style-type: none"> • En el primer millón de cifras de desarrollo decimal de $1/7$, ¿Cuántas veces ocurre el guarismo 1? ¿Cuántas veces 5? ¿Cuántas 3?. Para resolver este problema, no hace falta escribir efectivamente un millón de dígitos decimales y contar los guarismos “1”, uno por uno, por tratarse de una decimal periódica: $0, (142857) (142857)... (142857) 1428$ con un millón de cifras exactas. Aquí evidentemente se forman 166.666 períodos completos y uno incompleto de 4 cifras solamente, porque se tiene $1000000=6 \times 166.666 + 4$. Luego el guarismo 1 se encuentra 166.667 veces, y el guarismo 5 166.666 veces, y el guarismo 3 ninguna. Esto no es problema esta-dístico, por cuanto se conoce la ley exacta de la distribución de los guarismos en una decimal periódica. 	<ul style="list-style-type: none"> • ¿Cuántas veces se encuentra el guarismo 1 en el primer millón de cifras del desarrollo decimal de $\sqrt{2}$ ó de π? Estas decimales no son periódicas, luego no hay otro camino sino calcular efectivamente un millón de dígitos decimales, y contar uno por uno los guarismos “1” cuantas veces se encuentren: éste es problema estadístico. NOTA: Para la expresión decimal, carece de importancia que $\sqrt{2}$ sea número algebraico y π trascendente. El caso sería distinto si en vez de decimales usáramos fracciones continuas, que en el primer caso es periódica, y en el segundo no.
<ul style="list-style-type: none"> • ¿Cuántos cuadrados perfectos están comprendidos entre 2.000.000 y 3.000.000? Por cuanto $1414^2 < 2.000.000 < 1415^2$; y $1732^2 < 3.000.000 < 1733^2$, el número requerido es $1.732-1.414=318$, éste no es problema estadístico. 	<ul style="list-style-type: none"> • ¿Cuántos números primos están comprendidos entre 2.000.000 y 3.000.000? Aquí no hay fórmula exacta alguna y la única solución posible es la calcular todos los números primos en el intervalo de 2.000.000 hasta 3.000.000 y contarlos uno por uno. Es un problema estadístico.
<ul style="list-style-type: none"> • ¿Cuántos funcionarios emplea cierto Ministerio? La pregunta no es de índole esta-dística, sino legal, y para contestarla hay que consultar la Ley de Presupuesto Nacional. 	<ul style="list-style-type: none"> • ¿Cuántos habitantes tiene cierto pueblo? Es la tarea tradicional de la Estadística.

Zavrostsky, A: Revista Economía No. 2, 1988. 143-146.

Y como ejemplo más sensacional de un censo no estadístico, cítese el número de los electrones en el universo, iguala 15.747.724.136.275.002.577.605.653.961.181.555.468.044.717.914. 527.116.709.366.231.425.076.185.631.296 según Eddington (Wauch,1952), quien seguramente no los contó uno por uno.

En nuestra opinión, la expresión “Ciencia de los conjuntos empíricos” posee suficiente generalidad para abarcar la esencia de los objetos que caen dentro del dominio de la Estadística.

Bibliografía

- Bell, E. T. (1938): **The Queen of Sciences**, G.E. Stechert &C°, New York, p16.
- Bruño, G. M. (1930): **Elementos de Trigonometría**, Librería Bouret, París, p. 5
- Jonhnsom, N, L, & Tetley H. (1962) : **Statistic an Intermediate Text Book**, Cambridge University Press, Vol 1, p.8
- Eves Howard: **An Introduction to the History of Mathematics**, Holt, Rinehart & Windston, New York, p. 197.
- Mood, A. M. & Graybill, F. A. (1950): **An Introduction to the Theory of statistic**, Mac Graw Hill, New York, p.1.
- Von Mises, R. (1951): **Wahrscheinlichkeit**, Statistik und Wahrheit, Springer Verlag Wien, p.192.
- Centeno Grau, M.(1968): “**Estudios Sismológicos**”, Boletín de la Academia de Ciencias Físicas, Matemáticas y Naturales, N° 79, Caracas, p.225.
- Wauch, A. E. (1952): **Statiscal Tables and Problems**, McGraw-Hill, New York, Toronto & London, p.117.