

Comprender las bases de la estadística (Understanding the basics of statistics)

Badii, M.H. & A. Guillen

UANL, San Nicolás, N.L., mhbadii@yahoo.com.mx

Keywords: Analysis-synthesis, experimental design, inference, sampling, statistics

Abstract. Simple fundamental basics of statistics are briefly described. The historical origin and the necessity to understand statistics are noted. An especial emphasis is placed upon some aspects such as collection, analysis, synthesis, statistical phases and decision making in dealing with statistics.

Palabras clave: Análisis-síntesis, diseño de experimentos, estadística, inferencia, muestreo

Resumen. En este trabajo se describen de forma breve y simple las bases fundamentales de la estadística. Se menciona el origen de esta rama de la ciencia y la necesidad de la comprensión de la misma. Se pone un particular énfasis sobre la colección, análisis, síntesis, los métodos estadísticos, las fases de la estadística y el aspecto inferencial de la misma.

Introducción

La estadística es la ciencia que se trata de *verificar la validez probabilística* de cualquier fenómeno, concepto, proceso, evento o objeto en el espacio y el tiempo. Por tanto, si aceptamos que *el lenguaje universal es la matemática*, debemos también por la misma lógica acordar que *el idioma de la ciencia es la estadística*. La palabra *statistik* proviene de la palabra italiana *statista* que fue utilizada por primera vez por Gottfried Achenwall (1719-1772), un profesor de Marlborough y de Göttingen, y el Dr. E. A. W. Zimmerman introdujo el término *estadística* a Inglaterra. Su uso fue popularizado por sir John Sinclair en su obra *Statistical Account of Scotland* (1791-1799). Sin embargo, mucho antes del siglo XVIII, la gente utilizaba y registraba datos.

La estadística gubernamental oficial es tan vieja como la historia registrada. El viejo testamento contiene varios informes sobre levantamiento de

censos. Los Gobiernos de los antiguos Babilonia, Persia, Egipto y Roma reunieron registros detallados sobre la población y recursos. En la edad media, los gobiernos empezaron a registrar la propiedad de la tierra. En el año 762 de nuestra era, Carlomagono pidió la descripción detallada de las propiedades de la Iglesia. A principio del siglo IX terminó la enumeración estadística de los servicios que había en los feudos. Por el año 1086, Guillermo el Conquistador ordenó que se escribiera el *Domesday Book*, un registro de la propiedad, extensión y valor de las tierras de Inglaterra. Este trabajo fue el primer resumen estadístico de Inglaterra. Durante el siglo XVI los gobiernos inglés y francés empezaron a registrar el número de muertos por el brote de peste, el número de bautismos, defunciones y matrimonios.

La historia del desarrollo de la teoría estadística y su práctica es larga. Sólo hemos empezado a nombrar las personas que hicieron contribuciones significativas al campo. Más adelante encontraremos a otros cuyos nombres están relacionados con leyes y métodos específicos. Mucha gente ha contribuido al estudio de la estadística con refinamientos a innovaciones que, en conjunto, constituyen la base teórica de lo que se va a estudiar en esta obra (Badii & Castillo, 2007, 2009).

El propósito de este trabajo es la identificación de la naturaleza de la estadística, sus objetivos y cómo desempeña un papel importante en las ciencias, en la industria y, finalmente, en nuestra vida cotidiana. En esta investigación describiremos los objetivos de la ciencia estadística. En especial, se identificarán los tipos de problemas que la metodología estadística puede resolver y después explicaremos cómo se puede utilizar éste medio valioso para contestar algunas preguntas de la práctica.

La necesidad de la comprensión

Un banco ha aprendido de la experiencia que existen cuatro factores que influyen en gran medida en la determinación de si un cliente pagará a tiempo el préstamo que se le hizo o si se va convertir en moroso. Tales factores son:

- a) El número de año que tenga viviendo en la dirección actual.
- b) Una antigüedad en su trabajo.
- c) El hecho de si el cliente es dueño o no de la casa que habita.

- d) El hecho de que el cliente tenga una cuenta de cheques o de ahorros en el mismo banco.

Desafortunadamente, el banco no conoce el efecto individual que tiene cada uno de tales factores sobre el resultado del préstamo. Sin embargo, posee archivos de computadora con información sobre los clientes y tiene conocimiento, también del resultado de cada préstamo. Una persona física solicita un préstamo vive hace 4 años en su dirección actual, es dueña de la casa, sólo tiene una antigüedad de 3 meses en su trabajo actual y no es cliente del banco. Mediante el uso de estadística, el banco puede calcular la probabilidad de que esta persona pague su préstamo si éste se le otorga. En este ejemplo la estadística es la probabilidad de que la persona pague a tiempo el préstamo. En el planteamiento de diferentes tipos de problemas, cada persona utiliza la palabra de manera correcta. Todos ellos recurren a la estadística para auxiliarse en la *toma de decisiones*.

En cierta ocasión, de acuerdo con Levin & Rubin (1996), Benjamin Disraeli hizo la siguiente aseveración: "Existen tres tipos de mentiras, las mentiras ordinarias, las grandes mentiras y las mentiras estadísticas". Este juicio, tremendamente severo, respecto a la estadística, que fue hecha ya hace varios años, se ha vuelto una descripción bastante acertada de muchos de los fracasos estadísticos que encontramos en la vida diaria.

Significado de los datos

La estadística como campo de estudio, es el arte y la ciencia de dar sentido a los datos numéricos. Las décadas de los ochenta y noventa fueron testigo de la creciente toma de conciencia de que el pensamiento estadístico es una de las claves para el desarrollo de la ciencia. Cuando un grupo de investigadores tiene que decidir cómo elaborar un nuevo producto alimenticio, puede guiarse por sus propios gustos e intuición u obtener datos tomados de una encuesta acerca de las preferencias de los consumidores. Cuando los gerentes de personal desean consultar a un asesor en inversiones, lo pueden elegir siguiendo la moda o tomando en consideración los datos relativos a la trayectoria de los candidatos. Obtener y utilizar datos en forma inteligente resuelve muchos problemas. El pensamiento estadístico es indispensable para todos los investigadores, tanto al tratar con las operaciones cotidianas como al buscar oportunidades para mejorar. El pensamiento estadístico abarca los siguientes conceptos:

- a) Cómo recabar datos provechosamente.

- b) Cómo dar a los datos en bruto una forma comprensible.
- c) Cómo utilizar los conceptos de la teoría de la probabilidad para entenderlos.
- d) Cómo inferir y hacer predicciones con base en la siempre limitada información disponible.
- e) Cómo utilizar las computadoras modernas en el proceso.

Colectar y ordenar datos

Hay que conocer y comprender desde el comienzo los siguientes conceptos:

- a) **Data:** Hechos numéricos sobre cualquier cosa.
- b) **Información:** Poner los datos en el contexto de interés.
- c) **Conocimiento:** Proporcionar significado de la información.
- d) **Sabiduría:** Búsqueda de la propiedad, esencia, relación causa-efecto de los conocimientos.

Los datos son colecciones de cualquier cantidad de observaciones relacionadas. Una colección de datos se conoce como *conjunto de datos*, y una sola observación es *un punto de datos*. Para ser útiles, los datos se deben recopilar y poner a nuestra disposición. La simple decisión de medir y recopilar los datos relevantes es el comienzo indispensable para usarlos en la solución de problemas. En la recopilación de datos el primer paso consiste *en determinar lo que se quiere medir*. A menudo, las variables más fáciles de medir no son las más relevantes para resolver un problema. Por ejemplo, a un productor de plantas de orquídea le gustaría saber cuál es la demanda mensual que tiene la planta. Sería relativamente fácil registrar cada mes las órdenes de compra de dicho producto, pero no es lo mismo que la demanda de los consumidores. Se debe recalcar que puede suceder que algunos clientes quieran comprar la planta pero no la encuentren en existencia cuando van a los viveros. Sería mejor obtener información no sólo sobre las órdenes de compra, sino también sobre las existencias en los viveros. Precisar un poco el problema ayuda a determinar los datos que se deben de reunir y cuáles son las variables importantes, no las más fáciles de medir.

La segunda decisión tiene que ver con la forma en que se obtiene la información. Con frecuencia los datos pueden recopilarse en el trabajo; si se establece algún plan, la información acerca de la producción de plantas (por ejemplo) se puede registrar y hacer accesible como parte de la operación cotidiana. Otras veces, los datos se deben buscar deliberadamente.

Un paso importante en la obtención de datos es el muestreo. Un grupo de estudiantes dedicados a la investigación sobre la dinámica de dengue que examina la incidencia de la enfermedad en diferentes comunidades de una ciudad no puede, en realidad, entrevistar a todos los habitantes de la ciudad, y por tanto, inevitablemente, los datos se obtienen de una muestra limitada de éstos.

Otra forma de recopilar datos es la experimentación. Los experimentos diseñados son un factor clave en la tendencia actual para mejorar la calidad de los productos y de los procesos, y no se limitan al laboratorio. En otras palabras, prácticamente, se puede experimentar con todos los productos o procesos que sea parte de una investigación científica y tecnológica.

Síntesis de los datos

Los datos en bruto, las largas listas de números, no son de mucha utilidad por sí solos. Se les debe *sintetizar o resumir de manera que se les pueda entender y utilizar*. Lo más importante y lo primero que hay que hacer para resumir los datos es presentarlos en forma de tabular o de manera gráfica. Algunas gráficas y diagramas relativamente simples pueden hacer obvios algunos aspectos sumamente importantes de los datos.

El siguiente paso al analizar los datos consiste, por lo general, en *determinar estimaciones o encontrar un valor en cierto sentido típico o promedio*. La mayoría de las personas, tengan o no una preparación en estadística, tienen una idea razonable de lo que significa la palabra "*promedio*". No obstante, incluso una noción tan simple requiere de alguna consideración. Por ejemplo, la media de las alturas de una especie animal difiere de la mediana de las mismas alturas. Medio y mediana son conceptos diferentes del "*promedio*". Es más, puede que el promedio no sea el estadístico más importante de una variable. En estadística la idea de *variabilidad o dispersión* es por lo menos tan importante como la de *promedio*. De hecho, entender la variabilidad puede ser el beneficio más importante que uno recibe del estudio de la estadística.

Un último paso al resumir los datos consiste en *determinar cualquiera forma o patrones que éstos presenten*. Queremos buscar las "*asimetrías*" en ellos y los valores "*atípicos*"; valores poco usuales, "*rarezas*" que difieren considerablemente de la mayoría de los datos. La *asimetría* puede distorsionar los

estadísticos que resumen los datos. Por ejemplo, un diagrama que presente la riqueza faunística en comunidades de zonas áridas, no será simétrico; la mayor parte de las comunidades tendrá un valor muy pequeño, otros tendrán valores moderadamente grandes y puede existir algunos que cuentan con valores grandes. Si no tomamos esto en cuenta y sólo consideramos el promedio de la riqueza, nuestra apreciación será errónea. Los valores *atípicos* también distorsionan los promedios. Es más, los valores atípicos pueden indicar casos peculiares.

Las síntesis de datos pueden ayudar a los investigadores de tomar decisiones a hacer suposiciones bien pensadas acerca de las *causas* y, por tanto, de los *efectos* probables de ciertas características en situaciones determinadas. Antes de depositar nuestra confianza en cualquier conjunto de datos interpretados, ya vengan éstos de una computadora o no, debemos probarlo mediante las siguientes preguntas:

- a) ¿De dónde vienen los datos? ¿La fuente es parcial?, es decir, ¿Es posible que haya un interés en proporcionar datos que conduzcan a una cierta conclusión más que otras?
- b) ¿Los datos comprueban o contradicen otras evidencias que se poseen?
- c) ¿Hace falta alguna evidencia cuya ausencia podría ocasionar que se llegue a una conclusión diferente?
- d) ¿Cuántas observaciones se tienen? ¿Representan a todos los grupos que se desea estudiar?
- e) ¿La conclusión es lógica? ¿Se ha llegado a conclusiones que nuestros datos no confirman?
- f) ¿Vale la pena usar los datos o debemos esperarlos y recabar más información antes de actuar?

Búsqueda de un patrón significativo en los datos

Existen muchas formas de organizar los datos. Podemos sólo coleccionarlos y mantenerlos en orden; o si las observaciones están hechas con números, entonces podemos hacer una lista de los puntos de datos de menor a mayor según su valor numérico. Pero si los datos son trabajadores especializados de una construcción; o

los distintos tipos de automóviles que ensamblan todos los fabricantes; o los diferentes colores de suéteres fabricados por una empresa dada, debemos organizarlos de manera distinta. Necesitaremos presentar los puntos de datos en orden alfabético o mediante algún principio de organización. Una forma común de organizar los datos consiste en dividirlos en categorías o clases parecidas y luego contar el número de observaciones que quedan dentro de cada categoría. Este método produce una *distribución de frecuencias*.

El objetivo de organizar los datos es permitirnos ver rápidamente algunas de las características de los datos que hemos recogido. Buscamos cosas como el *alcance* (los valores mayor y menor), *patrones* evidentes, el *promedio*, es decir, alrededor de qué valores tienden a agruparse los datos, qué valores aparecen con más frecuencia (la *moda*), etc. Cuanta más información de este tipo podamos obtener de nuestra muestra, mejor será el entendimiento de la población de la cual proviene, y mejor será nuestra toma de decisiones.

Fases de la estadística

La estadística es un área de la ciencia que se ocupa del diseño de experimentos o de métodos de muestreo, del análisis de datos y de obtener inferencias acerca de una población de mediciones a partir de la información contenida en una muestra, en resumen, esta ciencia se encarga de las nociones y validaciones probabilísticas. La estadística se ocupa del desarrollo y aplicación de procedimientos para el diseño, el análisis y la realización de inferencias que darán la mejor información a un costo mínimo. También se puede definir la estadística como una herramienta aplicada a la descripción de conjunto de datos para resolver problemas mediante el análisis de datos, extraer información de un gran cúmulo de datos, la estadística además estudia la estructura de la disciplina de los "métodos estadísticos y de sus características históricas".

La estadística es una rama de las matemáticas aplicada para los profesionales. La estadística aplicada con frecuencia recurre a la intuición, la aritmética, y al álgebra elemental.

Los científicos aplican diferentes técnicas estadísticas a virtualmente todas las ramas de la ciencia. Estas técnicas son tan diversas que los estadísticos, por lo general, las dividen en dos grandes categorías: *estadística descriptiva* y *estadística inferencial*.

La *estadística descriptiva* dedicada a llevar registros ordenados de datos del estado, sirve como una herramienta o instrumento para tabular, representar, describir o resumir una serie de datos o elementos que pueden ser cuantitativos o cualitativos (el procesamiento de un conglomerado de datos mediante el muestreo

representativo). Por ejemplo, un profesor desea calcular la calificación promedio de un grupo de la materia de la administración. Como la estadística describe el desempeño del grupo pero no hace ninguna generalización acerca de los diferentes grupos, podemos decir que el profesor está utilizando estadística *descriptiva*.

Supongamos ahora que el profesor decide de utilizar el promedio de calificación obtenido por uno de sus grupos para estimar la calificación promedio de las diez unidades del mismo curso de la administración. El proceso de estimación de tal promedio sería un problema concerniente a la *estadística inferencial*. Los estadísticos se refieren también a esta rama como *inferencia estadística*. Obviamente, cualquier conclusión a la que llegue el profesor sobre diez unidades del curso estará basada en una generalización que va más allá de los datos del grupo original de la materia, y éste puede no ser completamente válida, de modo que el profesor debe establecer cual será la probabilidad de que la inferencia sea cierta.

De manera similar, la inferencia estadística implica generalizaciones y afirmaciones con respecto a la *probabilidad* de su validez. Con base a este ejemplo se puede definir que la *estadística inferencial* como un cuerpo uniforme de técnicas esta diseñada para resolver otro tipo de problemas tentativa de inferir propiedades de grandes números de datos mediante el análisis de muestreo de muy difícil solución por la mente humana no educada matemáticamente, y sirve para dar información de la que se puede sacar conclusiones acerca de un grupo grande de elementos mediante el análisis de muestreo.

La *estadística inferencial* esta basada fundamentalmente, sobre la *nociones de las probabilidades y diseños experimentales* que ambos se relacionan esencialmente en la comprensión de la teoría de probabilidades.

Los métodos y las técnicas de la inferencia estadística se pueden utilizar también en una rama de la estadística que se conoce como *teoría de decisiones*. El conocimiento de teoría de decisiones es muy útil para los científicos, ya que se le usa para tomar decisiones en condiciones de incertidumbre, cuando por ejemplo, un fabricante de aparatos de sonido no puede especificar precisamente la demanda de sus productos o cuando en una escuela se deben asignar grupos y crear horarios sin tener el conocimiento preciso del número de estudiantes que entrarán al primer grado.

Diseño de experimentos y métodos estadísticos

El segundo elemento de un problema estadístico es la decisión de *cómo se seleccionará la muestra*. Este elemento, llamado *diseño del experimento o procedimiento de muestreo*, es significativo porque los datos cuestan dinero y tiempo. Un buen diseño de muestreo puede reducir, a veces, los costos de la

recopilación de datos hasta en una décima parte, o hasta un centésimo del costo de otro tipo de diseño muestral (Badii et al., 2007).

Diseños del experimento es la tercera rama de los *métodos estadísticos* desarrollada para determinar y confirmar relaciones causales entre las variables en estudio, mediante la utilización de métodos apropiados de análisis de datos para extraer la información deseada a partir de ellos, no importa cuanta información contengan acerca de la cuestión práctica. La causa es lo que interesa a los investigadores y a los científicos como conceptos administrativos, finanzas, economías que son importantes en el diseño estadístico para el estudio de dichas relaciones.

El cuarto elemento de un problema estadístico es el *uso de los datos muestrales para formular una inferencia acerca de la población*. Para dicha inferencia pueden utilizarse muchos procedimientos para hacer una estimación o tomar una decisión respecto a alguna característica de una población o para predecir el valor de algún elemento de la misma.

El elemento final de un problema estadístico identifica lo que es probablemente la contribución más importante de la estadística a la confiabilidad de las inferencias para su realización. En resumen, un problema estadístico consta de lo siguiente:

1. Una *definición clara del objetivo del experimento* y de la *población del interés*.
2. El *diseño del experimento* o del *procedimiento de muestreo*.
3. La *colección y el análisis de los datos*.
4. El procedimiento para formular *inferencias acerca de la población* con base en la información muestral.
5. La obtención de una medida de la *bondad o confiabilidad de la inferencia*.

Cabe mencionar que los pasos en la resolución de un problema estadístico son secuenciales; es decir, se tiene que identificar primero la población de interés y planear cómo se reunirán los datos antes de poder recopilarlos y analizarlos. Todas estas operaciones tienen que proceder a la última meta, la realización de inferencias acerca de la población basada en la información contenida en la muestra.

Estadística aplicada e inferencial

La estadística es un área de la ciencia que se ocupa de la extracción de información a partir de datos numéricos y su uso en las inferencias acerca de una población de la cual se obtienen los datos. En algunos aspectos el estadístico cuantifica la información y estudia varios diseños y procedimientos de muestreo, buscando el procedimiento que conduzca una cantidad especificada de información en un caso dado a un costo mínimo. Por lo tanto, una gran contribución de la estadística reside en el diseño de los experimentos y encuestas, reduciendo así el costo y su tamaño. La segunda gran contribución es la realización de las propias inferencias. El estadístico estudia varios procedimientos inferenciales, buscando el mejor proceso estimador o para la toma de decisiones en un caso dado. Más importante aún, el estadístico proporciona información acerca de la bondad del procedimiento inferencial. Cuando se predice, convendría saber algo respecto al error de la predicción. Si tomamos una decisión, se desea conocer la probabilidad de que tal decisión sea correcta. Los sistemas humanos internos de predicción y toma de decisiones no proporcionan respuestas inmediatas a estas preguntas tan importantes. Se puede evaluarlas solamente mediante observación sobre un periodo largo de tiempo. Por lo contrario, los procedimientos estadísticos sí dan respuestas a estas preguntas. De esta manera, la estadística permite obtener inferencias a partir de los datos muestrales y evaluar la exactitud de dichas inferencias; esta información será muy útil al tomar decisiones en campo de los negocios o la administración.

Inferir de los datos

En términos generales, es difícil obtener una infinidad de datos sobre un proceso de producción. Al usar los datos, un gerente ha de inferir algunas características de la población subyacente o del proceso industrial basándose en la limitada información con que cuenta. Nunca podremos estar seguros de que la inferencia es absolutamente correcta. Por consiguiente, quienes utilizan los datos deberán considerar que casi siempre hay un cierto grado de incertidumbre.

La teoría estadística se vale de los resultados de la probabilidad para determinar la incertidumbre o grado de error probable de una inferencia. Por ejemplo, se puede llevar a cabo un proceso de producción de 1000 tarjetas para la computadora, bajo ciertas condiciones, y encontrar la proporción de tarjetas que necesitan ser reelaboradas, digamos el 18%. ¿Cuál es el grado de error probable de esta estimación? ¿Es el 1%? ¿Es el 10%? Si se duplica la muestra a 2000 tarjetas, ¿se reduce el grado de error probable a la mitad? Para contestar a estas preguntas necesitamos recurrir a la teoría de la inferencia estadística.

El papel de la computadora

El dar sentido a los datos se ha hecho mucho más fácil con el advenimiento de las computadoras modernas. Algunos métodos requieren tal volumen de cálculos que sería absolutamente imposible realizarlos sin ellas. Lo que tal vez más importante, las computadoras pueden hacer gráficas de los datos con sencillez y claridad. Todo método estadístico entraña hipótesis; con un programa adecuado para la computadora es posible comprobar lo razonable que son tales hipótesis. Las computadoras reducen el tedio en el análisis de los datos y permiten que éste sea más eficiente.

El hecho de que el análisis de datos se haga con una computadora no garantiza que éste sea un buen análisis. Los datos pueden ser muy sesgados o pueden haberse violado las hipótesis. La elección de métodos para analizar los datos puede no haber sido la adecuada. De acuerdo con Hildebrand & Lyman Ott (1998) es el ser humano quien decide qué datos utilizar, escoge los métodos de análisis, comprueba la hipótesis e interpreta los resultados razonablemente. Mientras la inteligencia artificial no haga avances considerables, la inteligencia humana seguirá siendo indispensable.

Hay literalmente cientos de paquetes de software estadístico para analizar los datos. Algunos, como SAS, SPSS, MINITAB, BMDP, etc. fueron adaptados a microcomputadoras. Para utilizar un paquete es necesario saber los pasos que permiten obtener los resultados particulares necesarios, dando entrada a los datos, asignarles nombres de modo que los resultados sean legibles, seleccionar las variables y observaciones que desea utilizar en el análisis y llamar a la parte adecuada del programa que lleva acabo las tareas específicas que se quiere realizar.

Conclusiones

Todo lo que ocurre en la vida a partir de que nace cualquier especie en el universo es (con la excepción de la muerte y el pago de impuesto tanto a la madre naturaleza y/o al padre del tiempo) probabilístico. Por tanto, la apreciación, el estudio y la comprensión de al menos las bases fundamentales de la ciencia de la estadística, lo cual se encarga de el determinar la validez probabilística de todas las cosas, es una necesidad imperativa para cualquier ciudadano universal, inteligente y responsable. Es con este objetivo que se dio a la tarea de escribir de forma simple y mencionar de manera breve un resumen de los conceptos esenciales e introductorios de la ciencia de la estadística.

Referencias

- Badii, M.H. & J. Castillo. (eds.). 2007. Técnicas Cuantitativas en la Investigación. UANL, Monterrey.
- Badii, M.H. & J. Castillo. 2009. Muestreo Estadístico: Conceptos y Aplicaciones. UANL, Monterrey.
- Badii, M.H., J. Castillo, M. Rodríguez, A. Wong & P. Villalpando. 2007. Diseños experimentales e investigación científica. *InnOvaciones de NegOcios*, 4(2): 283-330.
- Bancroft, T.A. 1968. Topics in intermediate statistical methods. Vol. 1. Iowa State University. Ames. Iowa, 129 pp.
- Box, G.E.P., W.G. Hunter, y J.S. Hunter, 1999. Estadística para investigadores. Editorial Reverté, S.A. 675 pp.
- Daniel, C. and F.S. Wood. 1980. Fitting equations to data. 2nd ed. John Wiley, New York. 458 pp.
- Dixon, W.J. and F.J. Massey. 1960. Introduction to statistical analysis. 3rd ed. McGraw-Hill. New York 638 pp.
- Hildebrand, D.K., y R. Lyman Ott, 1998. Estadística Aplicada a la Administración y a la Economía. Tercera Edición Addison Wesley Longman, 943 pp.
- Li, C.C. 1964. Introduction to experimental statistics. McGraw-Hill, New York. 460 pp.
- Ostle, B. 1963. Statistics in research 2a ed. Iowa State Press, Ames, Iowa.
- Tanur, J.M. 1978. Statistics: A Guide to the Unknown. 2a Ed. San Francisco Holden – Day.
- Wine, R.L. 1964. Statistics for Scientists and Engineers, Printice Hall, Inc. Englewood Cliffs, N.J.