

Historia de la Estadística

Esta historia de la estadística es continuación a la historia presentada en la Revista Varianza N° 7. El presente testimonio, muestra a personajes del quehacer estadístico a partir del siglo XIX.

WALTER A. SHEWHART



Walter A. Shewhart (1891 - 1967) físico, ingeniero y estadístico estadounidense, es conocido como el padre del **Control Estadístico de la Calidad**, nació en New Canton, Illinois, U.S.A., acudió a la Universidad de Illinois, y se doctoró en física en la Universidad de California en Berkeley en 1917. Shewhart sirvió al Departamento de Guerra de U.S.A. de las Naciones Unidas y del gobierno de la India, de igual forma, trabajó en el Consejo Nacional de Investigación y en el Instituto Internacional de Estadística.

Fue miembro honorario de la Real Sociedad de Estadística en Inglaterra, y de la Asociación de Estadística de Calcuta. Fue director del Instituto de Estadísticas Matemáticas de la Asociación Americana para el Avance de la Ciencia y formó parte de la Asociación Americana de Estadística.

En 1924, Shewhart delimitó el problema de la calidad en términos de variación de causas asignables y ocasionales e introdujo la gráfica de control como una herramienta para distinguir entre las dos. Shewhart hizo hincapié en que para llevar un proceso de producción a un estado de control estadístico, donde sólo existan variaciones de causas ocasionales, y mantenerlo en control,

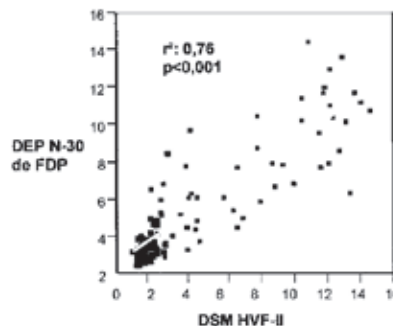
se requiere predecir los resultados futuros y administrarlo económicamente. Durante los 90's, el ingenio de Shewhart fue redescubierto por una tercera generación de directivos industriales como el enfoque "Seis Sigma".

JOHN WISHART



John Wishart (1898 - 1956) nació en Montrose Escocia, asistió a la Academia de Perth y luego, en 1916 ingresó en la Universidad de Edimburgo. Allí fue profesor de matemáticas de ET Whittaker.

La Primera Guerra Mundial hizo que la carrera universitaria de Wishart fuese interrumpida. Completó sus estudios universitarios en 1922, donde se graduó



en matemáticas y física. Había tomado un curso de formación del profesorado en la Casa de Moray, como parte de su carrera y, después de graduarse, se trasladó a Leeds tras aceptar un puesto como

profesor de matemáticas en la preparatoria West Leeds.

En 1924, después de una recomendación de Whittaker, le ofrecieron el puesto de asistente de Pearson en el University College de Londres. Wishart aprendió mucho de la estadística durante sus tres años con Pearson.

Wishart creó un laboratorio en Cambridge para sus alumnos de postgrado, donde Williams Cochran fue su alumno. Wishart tenía mucho talento para la estadística matemática y un instinto para realizar aplicaciones prácticas en diseño experimental.

Posteriormente, Wishart se convirtió en el Jefe del Laboratorio de Estadística de la Facultad de Matemáticas en Cambridge.

Algunas de las publicaciones más importantes de Wishart se realizaron en el período 1928-1932 antes de que se involucrara en la docencia en Cambridge.

En 1928 generalizó la distribución Chi-cuadrado que se denomina la Distribución de Wishart. También estudió las propiedades de la distribución del coeficiente de correlación múltiple que Fisher había considerado antes. También escribió numerosos artículos sobre las aplicaciones agrícolas de las estadísticas.

Wishart estuvo muy involucrado con el trabajo de la Royal Statistical Society. Él fue uno de los becarios que formaron Comité Organizador de la Sección de Investigación de la Agricultura en 1933. En 1945 se convirtió en presidente de la Royal Statistical Society de la sección de investigaciones.

Wishart murió en un accidente en Acapulco, México cuando se encontraba de visita en calidad de representante de la Organización de las Naciones Unidas para la Agricultura y organizando la creación de un centro de investigación para aplicar técnicas estadísticas en la investigación agrícola.

GERTRUDE MARY COX



Gertrude Mary Cox (1900 - 1978) nació en Dayton Iowa USA. En 1918 se graduó de la Escuela Secundaria de Perry en Iowa. Tras convertirse en diaconisa de la Iglesia Metodista en 1925 continuó sus estudios en la Universidad de Iowa (Iowa State College), donde estudió matemática y estadística graduándose en 1929 y posteriormente concluyó su maestría en estadística el año 1931.

En el periodo 1931 a 1933, efectuó estudios de postgrado en estadística en la UCB (University of California, Berkeley), luego, retornó a la Universidad Estatal de Iowa, a trabajar como asistente en un laboratorio de Estadística, donde empleó metodologías alternativas en Diseño de Experimentos. En 1939 fue nombrada profesora asistente en el Departamento de Estadística de la Universidad. En 1940 fue nombrada profesora de estadística en la Universidad de Carolina del Norte en Raleigh.

En 1945 se convirtió en directora de la División de Investigación en Estadística de la Universidad Estatal de Carolina del Norte (North Carolina State College) que fue dirigida por Cochran. En 1947 se convirtió en miembro fundador de la Sociedad de Biometría.

En 1950 en colaboración con William Cochran publicó el libro *Diseño Experimental*, que se convirtió en un clásico sobre el tema. En 1960 realizó su último trabajo como Directora de Estadísticas de Research Triangle Institute en Durham, Carolina del Norte. Permaneció en este cargo hasta su jubilación en 1964.

En 1949 se convirtió en la primera mujer elegida por el Instituto Internacional de Estadística (ISI - Instituto Internacional de Estadística) por su eficiente labor a favor de la profesión. En 1956 fue elegida presidente de la American Statistical Association (ASA - American Statistical Association) y en 1975 fue tomada como miembro de la Academia Nacional de Ciencias.

FRANK YATES



Frank Yates (1902–1994) nació en Manchester, Inglaterra. Fue educado en Wadham House, una escuela privada. Obtuvo una beca para Clifton College en 1916. Cuatro años más tarde se le concedió una beca para estudiar en la Universidad de San Juan en Cambridge. Se graduó en la “Clase Matrícula de Honor” (con honores de primera clase) en 1924.

Después de dos años como docente en matemáticas, decidió abandonar la enseñanza y se unió a la compañía Gold Coast Survey como asesor matemático. Debido a su estado de salud, decidió trabajar en Inglaterra y después de hacer una solicitud a RA Fisher, fue nombrado asistente estadístico en la Estación Experimental de Rothamsted en 1931. En 1933 se convirtió en jefe de Estadísticas de la mencionada institución, ocupó este cargo hasta su jubilación en 1968.

Yates trabajó bastante con el Diseño Experimental, colaborando a Fisher en varios proyectos de investigación. Desarrolló los métodos computacionales para el análisis de experimentos a través de mínimos cuadrados y creó los bloques incompletos modernos como los lattices.

Durante la Segunda Guerra Mundial, estudió el suministro de alimentos y la aplicación de fertilizantes para aumentar los rendimientos de la tierra. Aplicó sus técnicas estadísticas de diseño de experimentos en problemas tales como el control de plagas. Después de 1945 se continuó aplicando sus técnicas estadísticas a los problemas de la nutrición humana.

Frank fue nombrado parte de la Comisión Estadística de Muestreo de las Naciones Unidas y publicó los métodos de muestreo para encuestas y censos en 1949.

Yates era un entusiasta usuario de la computadora y comentaba: *“Para ser un buen estadístico teórico también se debe utilizar el equipo y por lo tanto debe tener los mejores recursos de computación”*.

Fue una de las personas que influyeron en la creación de la British Computer Society (Sociedad Británica de Computación) y fue su presidente en 1960 y 1961. Después de su retiro se convirtió en investigador senior en el Imperial College de Londres.

En 1966 La Royal Society de Londres le dio la Medalla Real por sus contribuciones de largo alcance profundo en los métodos de la biología experimental.

JOHN WILDER TUKEY



John Wilder Tukey (1915-2000) estadístico nacido en New Bedford, Massachusetts. Tukey obtuvo el título de Bachiller en Artes en 1936 y una maestría

en Ciencias en 1937, ambas en química, en la Universidad de Brown, luego obtuvo su doctorado en Matemáticas en la Universidad de Princeton.

Durante la Segunda Guerra Mundial, Tukey trabajó en la Oficina de la Investigación de Control de Fuego de Artillería y colaboró con Samuel Wilks y William Cochran. Después de la guerra, regresó a Princeton dividiendo su tiempo entre la universidad y los laboratorios AT&T Bell.

Tukey ha sido uno de los grandes talentos estadísticos del siglo XX, con importantes contribuciones a la Topología, Visualización de Información y en especial a la estadística, incluyendo la filosofía.

Tukey introdujo las modernas técnicas para estimar el espectro de las Series Temporales. En 1965, en un artículo conjunto con J. W. Cooley, publicado en la revista *Mathematics of Computation*, introdujo el algoritmo de la Transformada Rápida de Fourier (FFT), fundamental para crear el procesamiento digital de datos.

En 1970, contribuyó significativamente en lo que ahora conocemos como la Estimación Jackknife, también denominada la Quenouille-Tukey Jackknife.

Fundó el Análisis Exploratorio de Datos o EDA (Exploratory Data Analysis), una nueva aproximación a la estadística que usa fuertemente un conjunto de técnicas basadas en el uso de gráficos. Introdujo los diagramas de caja (Box Plot) en su libro de 1977, *Análisis Exploratorio de Datos*. Tukey está detrás de muchas nuevas y sencillas formas de tener en cuenta las magnitudes estadísticas. Entre ellas destacan los gráficos “Box and Whisker Plot” (diagrama de caja y bigotes), el “Stem-and-Leaf Diagram” (diagrama de tallo y hojas), los “radigramas” (rootograms) y los diagramas de ajuste.

GEORGE EDWARD PELHAM BOX



George Edward Pelham Box (1919) químico, matemático, estadístico inglés, en el área de la estadística realizó importantes contribuciones en las áreas de Control de Calidad, Análisis de Series Temporales, Diseño de Experimentos y la Inferencia Bayesiana.

Box nació en Gravesend, Kent, Inglaterra y se formó como químico. Durante la Segunda Guerra Mundial, trabajó en experimentos bioquímicos sobre el efecto de gases tóxicos en los animales pequeños para el ejército británico. Necesitaba asesoramiento estadístico para analizar los resultados de sus experimentos, pero no pudo encontrar un estadístico que podría darle orientación, por lo que él mismo estudio la estadística de los textos disponibles. Después de la guerra, se matriculó en la Universidad College de Londres y obtuvo una licenciatura en matemáticas y estadística. Recibió un doctorado de la Universidad de Londres en 1953.

Box ha escrito artículos de investigación y libros publicados, estos incluyen *Análisis de Series de Tiempo: Previsión y Control* (1979, con Gwilym Jenkins) e *Inferencia Bayesiana en el Análisis Estadístico* (1973, con George C. Tiao). Hoy, su nombre está asociado con resultados importantes en las estadísticas como el Modelo de Box-Jenkins, transformaciones de Box-Cox, diseños de experimentos y muchos otros.

George Box se casó con Joan Fisher, la segunda hija de Ronald Fisher.

El estadístico George Box acuñó, en 1953, el término robustez, para designar los métodos estadísticos que procuran asegurar resultados aceptables, cuando no se cumplen los supuestos estándares en que se basan los métodos estadísticos regulares. Ya desde fines del siglo pasado hubo científicos que se preocuparon del tema. Se dieron cuenta de los peligros de hacer inferencias, cuando los datos aparecen contaminados con valores extraños, y llegaron a proponer modelos y estimadores robustos como alternativas para estos casos. Sin embargo, no es sino a partir de la década de 1960, que este tópico es reconocido como un tema de investigación en estadística.

Box se desempeñó como presidente de la Asociación Americana de Estadística en 1978 y del Instituto de Estadística Matemática en 1979. Recibió la medalla Shewhart de la Sociedad Americana de Control de Calidad en 1968, el Premio en memoria a Wilks de la Asociación Americana de Estadística en 1972, el Conferenciante RA Fisher en 1974, y la Medalla de Guy de Oro de la Royal Statistical Society en 1993. Box fue elegido miembro de la Academia Americana de Artes y Ciencias en 1974 y miembro de la Royal Society en 1979.

HAROLD HOTELLING



Harold Hotelling (1895 – 1973) norteamericano, nació en Fulda Minnesota, fue un estadístico matemático y un economista teórico influyente.

Demostrando que no todo depende de Fisher, trabajó con la prueba multivariada (T de Hotelling) que generaliza la prueba de Student a varias dimensiones.

Su nombre es conocido por todos los estadísticos debido a la Distribución de T-Hotelling y su uso en pruebas de hipótesis y de confianza.

También presentó el Análisis de Correlación Canónica, y es el epónimo de la Ley de Hotelling, el Lema de Hotelling y la Regla de Hotelling en la economía.

Hotelling ocupa un lugar crucial en la genealogía de la teoría económica moderna, varias áreas de investigación activa se vieron influidos por sus artículos de economía.

Además de su enseñanza, escritura y actividades profesionales, Hotelling trabajó en los consejos editoriales de la revista *Annals of Mathematical Statistics* y la *American Journal of Economics and Sociology*. Fue elegido presidente del Instituto de Estadística Matemática y la Sociedad Econométrica. También fue elegido miembro de la Asociación Americana de Estadística en 1937.

JERZY NEYMAN



Jerzy Neyman (1894-1981) matemático y estadístico nacido en Moldavia Polonia.

En 1934, Jerzy Neyman publicó en la Royal Statistical Society, de Londres, lo que puede ser considerado como el primer trabajo científico sobre el Muestreo de Poblaciones Finitas. Estableció, sin lugar a dudas, que la selección aleatoria es la base de una teoría científica que permite predecir la validez de las estimaciones muestrales.

También dejó establecido toda una filosofía sobre la eficiencia de la estrategia muestral. Neyman y Egon Pearson, hijo de Karl Pearson, presentaron en 1936 una teoría sobre la forma de probar hipótesis estadísticas, en base a datos. Esta presentación promovió mucho interés, estimuló una considerable cantidad de investigación, y muchos de los resultados hasta hoy aún se usan.

Ellos resolvieron dificultades fundamentales para la comprensión de las pruebas de hipótesis, introduciendo las nociones de hipótesis alternativa, y los dos tipos de error, el de rechazar una hipótesis que es verdadera, y el de no rechazar una hipótesis que es falsa. De ello, surge un resultado fundamental, el Lema de Neyman - Pearson, y se crea una larga controversia con R.A. Fisher, que visualizaba la prueba de hipótesis como un procedimiento mediante el cual el investigador podía formarse una opinión sobre alguna característica de la población, o parámetro. Neyman y Pearson vieron la prueba de hipótesis como un medio para que el investigador tomara una decisión sobre un parámetro de la población. El Lema de Neyman - Pearson es considerado como la base de la moderna teoría de la decisión.

Neyman introdujo, en 1934, la teoría de los intervalos de confianza. Es una forma de estimar un parámetro, contrapuesta a la estimación puntual, que determina un intervalo que contiene al parámetro, y un coeficiente de confianza, que representa la probabilidad que el intervalo efectivamente contenga al parámetro. Los intervalos de confianza y las pruebas de hipótesis son dos elementos de la inferencia estadística fundamentales.

En las décadas de los años 30 y 40 se centra en el desarrollo de la técnica del Muestreo Estratificado en la población. Jerzy Neyman fue un arduo defensor del método de muestreo estratificado, y su trabajo en este tema, abrió nuevas áreas de investigación. También desarrolló el método de muestreo por conglomerados. Durante el periodo de 1930 y 1940 se establecen las condiciones bajo las cuales estos métodos, el muestreo estratificado y el muestreo por conglomerados, resultan más eficientes.

WILLIAM GEMMELL COCHRAN



William Gemmell Cochran (1909 – 1980) fue matemático y un destacado estadístico escocés.

Estudió matemáticas en las universidades de Glasgow y Cambridge y posteriormente trabajó en la “Estación Experimental de Rothamsted” entre los años 1934 a 1939, fecha en que emigró a EE.UU. Allí colaboró en la creación de varios departamentos de estadística y trabajó en la Universidad de Harvard desde 1957 hasta la fecha de su jubilación en 1976.

En el área de la estadística, realizó contribuciones al Diseño de Experimentos y a la Teoría del Muestreo.

Escribió muchos artículos y libros, algunos de los cuales se han convertido en referencias fundamentales y de los que podemos mencionar:

- Cálculo e Interpretación del Análisis de Varianza y Covarianza, que publicó en 1934 - Experimental Designs (con Gertrude Mary Cox), 1950.

- Statistical Methods Applied to Experiments in Agriculture and Biology, de George W. Snedecor, pero al que Cochran contribuyó a partir de la quinta edición.

FRANK WILCOXON



Frank Wilcoxon (1892-1965) químico y estadístico nacido en Estados Unidos, entre sus contribuciones en la estadística se tienen las Pruebas de Rango, las Comparaciones Múltiples, Rangos Secuenciales y Experimentos Factoriales.

A lo largo de su investigación, encontró que los métodos estadísticos eran numéricamente simples, fáciles de comprender y de aplicar.

Wilcoxon introdujo sus dos pruebas de rango, la prueba de la suma de rangos para dos muestras y la prueba del rango con signo para muestras apareadas, en un artículo el año 1945. Con este artículo y el de su contemporáneo Mann Whitney, comenzó el amplio desarrollo de la Estadística No Paramétrica. No hay duda que este artículo fue la contribución más importante de Wilcoxon ya que se convirtió en la mayor inspiración para el desarrollo de los métodos no paramétricos, además la metodología introducida ha tenido un amplio impacto en la estadística aplicada, particularmente por sus aplicaciones a las ciencias sociales, convirtiéndose en la herramienta estadística más popular.

En 1963 estudió las propiedades de la prueba de la suma de rangos donde se permite la dependencia dentro de las muestras. Wilcoxon estaba interesado

en la generalización de su procedimiento básico de la prueba del rango a nuevas situaciones. Frank fue el iniciador de la investigación en métodos secuenciales no-paramétricos. Debido a que las técnicas no paramétricas fueron tan exitosas en situaciones con muestras de tamaño fijo, pensó que sus buenas propiedades se podrían ampliar en forma natural a las situaciones secuenciales. Esta idea condujo a un gran número de procedimientos de rangos secuenciales desarrollados por Wilcoxon.

ANDRÉI NIKOLÁYEVICH KOLMOGÓROV



Andréi Nikoláyevich Kolmogórov (1903 - 1987) matemático ruso que hizo progresos importantes en los campos de la Teoría de la Probabilidad y de la Topología. En particular, desarrolló una base axiomática que supone el pilar básico de la Teoría de la Probabilidad a partir de la Teoría de Conjuntos. Al principio de su carrera trabajó en lógica constructivista y en las Series de Fourier.

También trabajó en la mecánica clásica. Asimismo, fue el fundador de la Teoría de la Complejidad Algorítmica. Obtuvo su doctorado en la Universidad Estatal de Moscú.

En 1922 Kolmogórov publicó sus primeros resultados en teoría de conjuntos y un año más tarde, construyó una Serie de Fourier, obteniendo un notable reconocimiento internacional.

Alrededor de este tiempo, decidió dedicar su vida a la matemática y publicó ocho trabajos sobre la

Teoría de la Integración, Análisis de Fourier y sobre la Teoría de Probabilidad.

Kolmogórov y el matemático Sydney Chapman desarrollaron el conjunto de ecuaciones fundamentales en el campo, las ecuaciones de Chapman-Kolmogórov.

ABRAHAM WALD



Abraham Wald (1902-1950) matemático austríaco, nacionalizado estadounidense. Exiliado a EE.UU. en 1938, se especializó en estadística y aportó a esta ciencia con un elevado rigor matemático. Fue el fundador del Análisis Secuencial.

Desde 1927 curso estudios en la Universidad de Viena, doctorándose en 1931. Formó parte del Institute for Business Cycle Research. Debido a la invasión nazi, se va a los Estados Unidos en 1938, donde es parte del equipo de la Cowles Commission Research y alumno de Harold Hotelling en Columbia University (New York).

Después fue profesor, caracterizándose por explicar complejos conceptos de manera muy sencilla. También desarrolló proyectos militares para el Statistics Research Group y desarrolló la teoría del Análisis Secuencial, aplicada a la II Guerra Mundial para mejorar los estudios de control

de la calidad. Destacó en programación Lineal, geometría y estadística, propuso varias aplicaciones matemáticas a la economía de cierta relevancia. Se interesó en la econometría, proponiendo modelos para suavizar la estacionalidad de las series temporales, y creando la “Prueba de Wald”, que data de 1939.

CALYAMPUDI RADHAKRISHNA RAO



Calyampudi Radhakrishna Rao nació en Hadagali, Estado de Karnataka India, el 10 de septiembre de 1920. Actualmente es profesor emérito de la Universidad Estatal de Pensilvania.

Trabajó en el Instituto Estadístico de la India y en el Museo Antropológico de Cambridge antes de conseguir el doctorado en el King's College de la Universidad de Cambridge. El tutor de tesis doctoral, fue R.A. Fisher.

Sus mayores descubrimientos fueron la Cota de Cramér-Rao, y el Teorema de Rao-Blackwell, ambos relacionados con la calidad de los estimadores. También trabajó en Análisis Multivariante y Geometría Diferencial.

Rao obtuvo el premio de la American Statistical Association, el Wilks Memorial Award en 1988. También fue premiado con la Medalla Mahalanobis. En junio de 2002, Rao fue premiado con la Medalla Nacional de Ciencias de Estados Unidos.

GEORGE CASELLA



Distinguido profesor en el Departamento de Estadísticas y miembro del Instituto de Genética para la Universidad de Florida. Investigador activo en muchos aspectos de la estadística, contribuyendo en la Teoría Estadística en las áreas de Teoría de la Decisión y Estadística confidencial para la Estadística medioambiental. Su más recientemente esfuerzo, está concentrado en la Estadística Genoma y Metodología en Ciencias Políticas. También mantiene activa investigación en la Teoría y Aplicación de Monte Carlo y otros métodos intensivos computacionales.

Es miembro de la Asociación Americana de Estadística. El instituto de Estadística Matemática, y un electo miembro del Instituto Internacional de Estadística (IIE) donde se habla de él como un muy buen investigador y recientemente fue electo miembro extranjero de la Academia de Ciencia Española Royal.

Casella se desenvuelve como editor de teoría y métodos de la revista de la Asociación de Estadística Americana, actualmente es editor conjunto de la revista de la Sociedad Estadístico Royal.

Casella es autor de 8 libros:

- Variance Components.
- Theory of Point Estimation.
- Statistical Inference.
- Monte Carlo Statistical Methods.
- Statistical Analysis of Quantitative Traits.
- Statistical Design.
- Introduction to Monte Carlo Methods with R.

ROGER BERGER



Recibido en matemáticas de la Universidad de Kansas, tiene una maestría y doctorado en Estadística de la Universidad de Purdue, finalizó sus estudios de PhD en 1977. Brindó sus servicios como asistente y profesor para la Universidad del Estado de Florida. En 1982, paso a la facultad del departamento de Estadística en la Universidad del Estado de Carolina del Norte como docente destacado.

Su interés profesional incluye la teoría estadística y la educación. Recientes investigaciones incluyen métodos exactos para datos categóricos y biomédicos incluyendo los test de equivalencia. Fue uno de los primeros en desarrollar la teoría del test de unión e intersección. Es co-autor con George Casella del libro, Inferencia Estadística.

“La esencia de la vida es la improbabilidad estadística a escala colosal.

Richard Dawkins