

METODOLOGIA BIOESTADISTICA PARA MEDICOS Y OFICIALES SANITARIOS

JOSE ALMENARA BARRIOS

Universidad de Cádiz

LUIS CARLOS SILVA AYÇAGUER

Instituto Superior de Ciencias Médicas de la Habana

RESUMEN

Analizamos el libro de D. Marcelino Pascua Martínez (1897-1977) titulado Metodología Bioestadística para Médicos y Oficiales Sanitarios. La obra fue publicada por la Editorial Paz Montalvo, en 1965. Su autor representa la culminación de la forma de pensar de los epidemiólogos españoles que desde el siglo XIX se venían preocupando por la utilización del método estadístico en la resolución de los problemas sanitarios, siendo el libro que nos ocupa el legado final del hombre clave en esa corriente de pensamiento. El texto está estructurado en siete capítulos, dos apéndices y cuatro tablas estadísticas. Cabe destacar del contenido el conjunto de técnicas estadísticas, demográficas, matemáticas y de diseño de estudios aplicados a la investigación sanitaria no habitual en los manuales de Bioestadística, no solo de su época, prácticamente inexistentes, sino de la actual.

ABSTRACT

We analyze the book by D. Marcelino Pascua Martínez (1897-1977), Biostatistics Methodology for Medical and Sanitary Officers. The work was published by Paz Montalvo in 1965. His author represents the culmination of the form of thinking of the Spanish epidemiologists that from the 19th century had been concerned about the use of the statistical method in the resolution of sanitary problems. The book approached is the final legacy of the key man in that trend of thought. The text is divided in to seven chapters, two appendixes and four statistical tables. The set of technical, demographic, mathematical statistics and of studies design applied to the sanitary investigation contained in this text is not usual in the Biostatistics handbooks, not only at that time, but also nowadays.

Presenta a lo largo del manual los contenidos de la Estadística, siempre con gran cantidad de ejemplos del ámbito sanitario. Destaca la primera parte dedicada a la Demografía Sanitaria, donde es pionero. El resto de temas por capítulos son: Medidas de Síntesis, correlación, teoría de probabilidad, teoría de tomas de muestra y estadística inferencial, la distribución χ^2 , el análisis de la variancia y una serie de apéndices donde trata el ajuste de curvas, los contrastes en los ensayos biológicos y las tablas estadísticas mas habituales.

The book presents the contents of Statistics, always with great quantity of examples of the sanitary area. The first part is devoted to the Demographic Sanitary, where it is pioneer. The rest of the topics by chapters are: Descriptive statistics, correlation, probability theory, samples and inference, the Chi-square distribution, the analysis of variance and a series of appendixes on curves adjustment, contrasts in the biological trials. The text ends with the presentation of the statistics tables but customary.

Palabras clave: Bioestadística, Epidemiología, Aplicaciones demográficas en sanidad, España, Latinoamérica, Siglo XX.

1. Introducción

El acercamiento al libro de Marcelino Pascua Martínez (1897-1977) que presentamos ha sido todo un ejercicio de rescate, descubrimiento, que hemos desarrollado con admiración, respeto y un tanto de rabia por el olvido imperdonable de su no-reedición moderna, de la que quizás tengamos culpa muchos de los que nos dedicamos a la Epidemiología y la Bioestadística. Obra realizada en la madurez del autor, no representa una tarea transversal en su vida sino un proyecto que creemos se gestó en su mente muchos años atrás y que cuando sus múltiples obligaciones se lo permitieron encaró, para dar lugar a un material clave de la Bioestadística hispana. No se puede aquilatar adecuadamente esta obra sin saber algo de la vida de D. Marcelino, porque sus reflexiones constituyen una manera de reaccionar frente a su existencia asumiendo el papel que le tocó desempeñar, comprometiéndose con sus ideas y con su labor científica. Por ello abordaremos algunos aspectos claves de su biografía.

Marcelino Pascua representa para nosotros el paradigma de los epidemiólogos españoles que desde finales del siglo XIX se venían interesando por la utilización del método estadístico para el estudio de los problemas de la salud y de la enfermedad en las poblaciones. Los datos biográficos mas relevantes han sido ya bien señalados por autores [BERNABEU I MESTRE,

1992; RODRIGUEZ y BERNABEU I MESTRE, 1996] quienes destacaron su formación inicial en la Universidad Complutense de Madrid, sus contactos con personas vinculadas a la Institución Libre de Enseñanza y especialmente su formación de postgraduado en Estadística y Epidemiología, tanto en EEUU en la Johns Hopkins University, como en Inglaterra en el University College y en el National Institute of Health de Londres, donde tuvo como profesores epidemiólogos y estadísticos de la talla de Greenwood, Pearl y Frost.

Su *curriculum* profesional es ciertamente denso, ocupó puestos docentes relevantes en la Escuela Nacional de Sanidad donde enseña Estadística, Epidemiología y Demografía y en los períodos que van de 1929 al 31 y de 1933 al 36 lo encontramos al frente de los Servicios de Estadística Sanitaria de la Dirección General de Sanidad. Alcanza a ser diputado socialista durante la República, ocupando en el periodo de 1931 a 1933 la Dirección General de Sanidad, desde donde da un gran impulso a la Salud Pública española con la implantación de numerosos programas de Salud y la creación del Consejo Superior de Investigaciones Sanitarias. Tras la guerra civil, durante la cual desarrolla una agitada actividad como embajador en Moscú y París, se ve abocado al exilio que le lleva de nuevo a la docencia, esta vez ejercida en Baltimore en la Johns Hopkins University, en concreto en su emblemática Escuela de Higiene y Salud Pública. Es aquí donde desarrolla una gran parte de su labor investigadora que facilitará su salto en 1948 a la Jefatura del Departamento de Estadísticas Sanitarias de la Organización Mundial de la Salud, cargo que ocupa hasta el año 1957, fecha de su jubilación.

En esta última etapa es cuando D. Marcelino, ya jubilado, se centra en uno de sus últimos y más acariciados proyectos, el manual de *Metodología Bioestadística para médicos y oficiales sanitarios*, que publicó la editorial Paz Montalvo en 1965 y que es objeto de estas páginas. Consta el mencionado manual de siete capítulos, dos apéndices, cuatro tablas estadísticas, las notas bibliográficas, un índice de autores y otro índice alfabético de materias, que se presentan en 627 páginas. En una primera aproximación, llama la atención lo minucioso del contenido que abarca un conjunto de técnicas estadísticas, demográficas, matemáticas y de diseño de estudios no habitual en los manuales al uso de Bioestadística, no ya de su época, prácticamente inexistentes, sino de la actual. Otro dato inicial que cabe resaltar son las fuentes que utiliza el autor, siempre asociados a los trabajos originales. Una pieza importante de la bibliografía que maneja es la clásica revista *Biometrika*, fundada en 1901 bajo el subtítulo *A journal for the statistical study of biological problems*, por Sir Francis Galton, W.F.R. Weldon, Karl Pearson y C.B. Davenport, y que continúa publicándose hasta nuestros días y que fue durante muchos años la revista más importante no solo de Bioestadística sino de Estadística en general, y posiblemente una de las de mayor impacto en el desarrollo de las Ciencias

empíricas. Lo que sí parece claro es que D. Marcelino la conocía de manera exhaustiva, y que la estudió desde el primer número hasta el último publicado en su época. Es expresivo de lo que decimos la cita que hace en la página de su libro del trabajo de W.R. Macdonell, [1902, vol. I, p. 375] titulado *On the influence of previous vaccination in cases of smallpox* contenido en el primer volumen de *Biometrika* (1901-1902), por ser el primer artículo de Epidemiología cuantitativa que publica la mencionada revista y que por supuesto conoce Pascua, trabajo que extrañamente no aparece citado en autores posteriores a Pascua en España, cuando como decimos es el primer trabajo de Epidemiología propiamente dicha que aparece en *Biometrika*, y donde se presenta unas de las primeras aplicaciones de las tablas estadísticas de contingencia 2 x 2 en Medicina, para evaluar la efectividad de la vacunación de la viruela en seis ciudades inglesas, tan en boga en la Epidemiología posterior.

Expuestos estos aspectos iniciales, describiremos y analizaremos los contenidos del libro. Del Prólogo cabe llamar la atención sobre dos consideraciones que, a nuestro entender, revelan la madurez estadística de este epidemiólogo y de su visión de futuro en cuanto a lo que debe ser la formación del médico. El primer aspecto se puede valorar en el siguiente párrafo:

"[...] adjudicase un cierto espacio a las nociones indispensables del Cálculo de Probabilidades, ya que él constituye base estadística general y fundamento esencial de la Teoría de Toma de Muestras, médula del moderno análisis cuantitativo".

La segunda idea queda plasmada en este otro:

"[...] no se precisa ser consumado arúspice, dado el rumbo que se observa en la investigación científica y su metodología estadística, para percibir que en plazo seguramente no largo la preparación en disciplinas matemáticas deberá ser más vasta que la demandada al presente".

A continuación de estas aseveraciones se desgranar los distintos capítulos que constituyen la obra. Los presentaremos, e iremos comentado sus aspectos más relevantes. Antes, una sola advertencia: quien siga leyendo estas notas, encontrará citas de don Marcelino escritas en el estilo propio *de su tiempo*. Pero no del período estricto en que fueron redactadas (finales de los 50) sino de su tiempo generacional, acaso ligeramente anticuado incluso en el momento de la publicación: es el lenguaje culto, a veces grandilocuente y ampuloso, siempre refinado, de sus padres intelectuales (Cajal, Ortega, ...), el estilo de quienes aún podían darse el lujo de escribir reposadamente, sin apremios ni tensiones editoriales.

Capítulo I

Los preliminares de este capítulo I representan el punto de partida de la obra y constituyen una declaración de intenciones sobre lo que es la Bioestadística para Pascua y lo que se puede esperar de su aplicación en el campo Bio-sanitario. Empieza por dar la definición de Estadística apoyándose en autores como Fisher, Yule y Gini, para, a continuación, dar su propia definición de Bioestadística de una manera pragmática:

"[...] entiéndese por Bioestadística su aplicación (la de la Estadística) al campo general de los problemas biológicos. Dentro de éste, nosotros nos ocuparemos en esta obra en particular de aquellos atinentes a la Medicina clínica y a la Estadística sanitaria. Pero nótese bien: tal tarea no depende pura y exclusivamente de técnicas de tipo matemático, sino asimismo de la conjunción de ellas con las propias del investigador en cuanto al planeamiento adecuado de la experiencia en estudio, y también de la aplicación de los principios de la Lógica en que puedan y deban apoyarse las deducciones e inferencias a sacar de los resultados numéricos".

Importante, vigente y acertado párrafo, que pone las cosas en su sitio y obliga al investigador biomédico a pensar. Investigar no es almacenar datos, para después tratarlos estadísticamente y obtener una conclusión, algorítmica que será falaz en la mayor parte de las ocasiones, por no decir en todas. Hoy tal adocenamiento se ha generalizado con el uso del ordenador, ya que se acopian datos y fácilmente se someten a los más sofisticados tratamientos estadísticos, que parecen avalar nuestras conclusiones. Por eso reparar en el alcance del párrafo anterior es de obligado cumplimiento para todo investigador que necesite del método estadístico para verificar hipótesis de naturaleza empírica.

Brillante también en este capítulo I son los comentarios que hace en el apartado Colección y registro de datos originales; algunos apuntan sobre el importante aspecto de la calidad de los datos originales y que tantos textos de Estadística ni siquiera mencionan; tal es el caso del que dice:

"[...] de la calidad que esencialmente posean las observaciones y medidas, así como del cuidado y método puesto en la colección y registro de ellas, dependerá en su mayor parte la propiedad y la pertinencia de las deducciones y conclusiones de la indagatoria".

Otros nos alertan sobre la posibilidad de incurrir en sesgos, como cuando se dice:

"Otro punto de cardinal importancia, por lo que respecta a la obtención y registro de datos estadísticos, que en particular debe ser subrayado por lo que toca a las investigaciones usualmente llamadas de tipo científico, es el de *evitar*

parcialidad y selección en el carácter de los resultados, sacando muestras que no representen la realidad del conjunto del fenómeno en estudio y viciando con ello, en buena lógica, el juicio y las conclusiones de índole general que se deriven".

Es exigente y pide un último esfuerzo, no solo en el análisis de los datos sino en el origen de los mismos,

"[...] y para ello, el estadístico y el analista cuantitativo deberán *revisar cuidadosamente las circunstancias operantes en la producción de los datos*".

En la sección 1ª de este capítulo comienza la Estadística Descriptiva, con la exposición de los temas relacionados con la tabulación y la representación gráfica de variables. En el apartado de gráficos presenta un número importante de ellos, no habitual en los libros de Estadística general o aplicada a un área del conocimiento en particular: polígonos de frecuencia, histogramas, ojivas, diagramas de barras, de barras compuestos, de exceso y defecto, en pirámide, circulares, polares, de dispersión, cartogramas, mapas epidemiológicos, estereogramas y gráficos para series temporales o históricas. No obstante tal profusión, nos aconseja prudencia a la hora de la representación gráfica, alentando la búsqueda de simplicidad y claridad.

Termina la sección con un apartado sobre maquinaria estadística, tampoco habitual en los tratados de su tiempo, que hoy a tan solo algo más de 40 años nos parece pura arqueología hablándonos de las máquinas taladradoras o perforadoras de tarjetas, las clasificadoras o separadoras y las tabuladoras eléctricas que permitían el recuento de varias características recogidas en las tarjetas, todo en la misma operación, (sin separarlas en grupos) y que daban los totales parciales y subtotalet, registrando los resultados en una hoja de salida. Qué lejos nos queda esta tecnología que nos apunta Pascua con el advenimiento de los ordenadores y los potentes programas estadísticos de los 10 últimos años.

Es en la Sección 2ª donde el libro empieza a alcanzar singular brillantez. Dicha Sección, dedicada a los números índices, exhibe gran precisión, conecta con aplicaciones en el campo de la Epidemiología, la Demografía y la Higiene Pública fácilmente, sobre todo con una cantidad y calidad de ejemplos valiosos, fruto de su extensa obra investigadora en estas áreas, tanto en EEUU como en la O.M.S., la cual es citada de manera frecuente en el texto. Los ejemplos son presentados en tablas y en diferentes cálculos para ejemplificar los conceptos desarrollados teóricamente con sobrada sencillez. Comienza diferenciando claramente los conceptos de razón, proporción y tasa. Deja claro que una tasa es una *medida de la intensidad con que un fenómeno ocurre referida de modo relativo a una base de 100, de 1.000 o de 100.000 individuos,*

o de unidades de experimentación u observación, según sean los casos o las conveniencias a nuestro fines expositivos, en un cierto período fijado de tiempo, y para un determinado territorio o localidad. Si bien la definición es aceptable, hoy podríamos entablar una discusión si es del todo correcta, ya que no recoge la idea actual de tasa, referida al cambio de una magnitud en una unidad de tiempo, de forma que la tasa de prevalencia que define Pascua para nosotros hoy, en sentido estricto, sería una proporción. Bien, al margen de discusiones en esta línea, que sobrepasa la pretendida reseña del libro, ya comentamos que esta sección es brillante y abarca en más de 100 páginas todos los índices o *coeficientes* habituales de los trabajos de Epidemiología y Demografía. Para el lector interesado y dada su importancia exponemos a continuación cada uno de los epígrafes de este apartado.

Los índices que presenta son:

A. De Natalidad

- Bruto
- Específico
- Tasa de reproducción bruta
- Tasa de reproducción neta
- Alumbramientos múltiples

B. De Mortalidad

- Tasas brutas
- Datos recientes
- Tasas específicas
- Análisis de la mortalidad por tasas específicas
- Tasas específicas por sexo y edades
- Por estado o situación económico-social
- Por causas
- Por sexo, edad y causas
- Clasificación de las causas de mortalidad
- Certificación médica de causas de defunción
- Mortalidad por causas desconocidas y mal definidas
- Datos para 1957.

C. De Mortalidad proporcional

D. De Mortalidad infantil

- Datos recientes
- Corrección del cálculo de sus tasas
- Mortalidad neonatal. Datos recientes
- Mortalidad infantil por atributos
- Mortalidad perinatal y neonatal en relación con el peso al nacer

- Mortalidad infantil por el estado económico de los padres
- Distribución mensual de la mortalidad infantil

E. De Mortinatalidad

- Precaución por definición de nacido-muerto
- Por orden del nacimiento
- Mortalidad perinatal

F. Mortalidad en la niñez

G. Mortalidad materna

- Especificidad
- Por edades de las madres y causas de defunción
- Datos recientes

H. De morbilidad

- Tasas de incidencia y de prevalencia
- Tasas específicas
- Fuentes generales de datos

I. De ataque secundario

- Usos en Sanidad

J. De letalidad por casos

- Fuentes principales de datos
- Varios aspectos epidemiológicos importantes

K. De incremento natural

- Datos recientes

L. De Nupcialidad

M. Tasas standardizadas

- Método directo
- Método indirecto
- Para dos atributos conjuntos
- Sobre el standard de población
- Otros índices globales comparativos

N. Tasa de mortalidad media equivalente

Ñ. Supervivencia y expectativa de vida

- Datos para España
- Datos para otros países
- Ganancias en expectativa de vida por reducción en causas de muerte

O. *Método de la observación continuada o progresiva por tasas de supervivencia*

- Porcentaje de enfermos de cáncer que sobrevivieron cinco años, por sexo y sitio de lesión.
- Bibliografía del método
- Encuestas *prospectivas* y *retrospectivas*.

P. *Tasas de mortalidad potencial*

Q. *Otras descripciones analíticas de la mortalidad*

R. *Indices de Salud e Higiene*

Como puede observar el lector, se aborda de manera minuciosa un amplísimo espectro de temas demográficos, pertenecientes a lo que posteriormente hemos venido a llamar sanitaria. Insistimos en que cada uno de los conceptos presentados lleva aparejado un ejemplo que aclara la situación y las ambigüedades que hubieran podido quedar, y que en la mayor parte de las ocasiones los ejemplos se construyeron con datos reales, fruto de su trabajo investigador.

Son de una gran riqueza e inteligencia los comentarios y las discusiones que plantea Pascua en cada uno de las situaciones del capítulo. Presentamos a continuación algunos que nos llamaron particularmente la atención. Por ejemplo, en el excelente apartado de estadísticas de mortalidad, hablando de las tasas específicas por sexo y edades en una comparativa que hace entre dos países, al señalar las diferencias observadas entre las mismas, escribe:

"[...] debiendo hacerse notar que esta variabilidad es uno de los más fundamentales hechos de la Demografía y de la Estadística sanitaria, y por tanto, asunto del más grande interés para nosotros".

De nuevo aparece la idea de dar sentido a lo que hacemos en Estadística, no recoger datos simplemente, sino ir mas allá, y buscar diferencias con intención explicativa, dando un sentido a nuestras hipótesis. Diremos hoy que necesitamos para llevar a cabo una investigación epidemiológica, medir, comparar e interpretar resultados. Pasaremos así de un primer nivel empírico destinado a describir a un segundo nivel teórico que intenta esquematizar la realidad, penetrando en la esencia del proceso [GUTIERREZ, 1991, p. 46].

Sin duda todo el apartado de clasificación de las causas de enfermedad y su análisis estadístico no pudo hacerlo nadie mejor que Pascua; nadie tenía la experiencia en su época para abordar de manera tan magistral esta parcela. Ya nos había legado con anterioridad su obra magna sobre la Mortalidad en

España, con un trabajo titulado *Mortalidad específica en España I. Cálculo de poblaciones(1934)*, un segundo trabajo publicado en 1935 con el título *Mortalidad específica en España II. Mortalidad por sexos, grupos de edades y causas en el periodo 1911-1930* [CLAVERO, 1992]. Y un tercer trabajo realizado en Baltimore, (*Mortalidad específica en España III. Mortalidad por sexos y causas de defunción de la lista larga internacional en el periodo 1901-1930. IV. Tablas de vida*) que ve la luz en 1942. Pero no olvidemos además su intenso paso por la O.M.S., donde ocupó la primera jefatura de la Sección de Estadísticas Sanitarias y donde termina su carrera profesional como decíamos antes en 1955, con el cargo de director consultor de la citada sección.

Capítulo II

Este capítulo segundo lleva el título de Medida de la Variación y en él D. Marcelino aborda las medidas de tendencia central, de dispersión, de forma, y un interesante apartado final sobre series temporales que abarca métodos para la determinación de la tendencia secular, métodos para estimar manifestaciones estacionales de procesos morbosos y de la mortalidad y análisis de las variaciones cíclicas de estos fenómenos, cerrando el capítulo con la definición de epidemia desde una perspectiva estadística.

Nada parece indicar que este capítulo se separe de los manuales corrientes de Estadística general o de Epidemiología, sin embargo de nuevo la sorpresa, nuevamente la minuciosidad en los detalles, en las citas bibliográficas, la ejemplificación continua, el apunte preciso para la aclaración del concepto explicado. Los comentarios a pie de página sobre la utilización correcta en castellano de algunos términos de la estadística o el recuerdo a los clásicos (Karl Pearson, R.A. Fisher, etc..) cuando presenta la notación por vez primera de las fórmulas y como nació la misma. O algo que se sucede en todo el libro: dar para cada termino nuevo, para cada concepto estadístico, su traducción al inglés, francés y alemán, recurso que facilita al estudioso su búsqueda en la bibliografía internacional. Son detalles que dan personalidad a un apartado repetido en cualquier manual de Estadística. En cuanto al contenido, nos lo podemos imaginar. Abarca todas las medidas de síntesis habituales en Estadística univariante: la media aritmética, tanto para datos agrupados como para datos no agrupados, la mediana, la norma, y medias menos familiares como la geométrica y la armónica. Dentro de las medidas de dispersión se ocupa de la amplitud, la desviación media, desviación estándar en datos agrupados y no agrupados; presenta la prueba de Charlier, que hoy prácticamente no aparece en ningún texto de Estadística general, ya que se utiliza para comprobar la exactitud de las operaciones aritméticas previamente al cálculo de las medidas de dispersión (obviamente hoy día dichas operaciones

las hace el ordenador y no dudamos de su infalibilidad), y la prueba de Sheppard, para corregir *el efecto causado por la agrupación de la frecuencia en el punto medio de las clases*. Continúa explicando la variancia, el coeficiente de variación y las medidas de forma como oblicuidad y kurtosis.

Capítulo III

En él presenta Pascua las técnicas de correlación, en tres direcciones:

- 1º Como cálculo del grado de asociación entre variables en una base relativa, por el coeficiente de correlación.
- 2º Estudiando la forma o relación funcional entre las dos variables de modo que ella nos permita estimar una en virtud de la otra, mediante la línea o ecuación de regresión.
- 3º Juzgando de la confianza que ofrecen las estimaciones o cálculos antedichos, por el llamado a veces *error estándar de la estimativa o del cálculo*.

Una vez establecidos los enfoques nos presenta la temática siguiente: calculo del coeficiente de correlación para datos no agrupados y agrupados, prueba de Charlier, regresión y cálculo de las líneas de regresión, correlación curvilínea, prueba de rectilinearidad de la regresión, ejemplos de correlación en Medicina e Higiene, coeficiente de correlación biserial, correlación por rangos, correlación entre series históricas, correlaciones espurias y correlación múltiple y parcial.

Hoy se escribe y se enfatiza sobre la diferencia entre causalidad y asociación estadística entre dos variables y son numerosos los autores que tienen que poner en evidencia esta diferencia ya que previamente a lo largo de los años en muchos trabajos de Epidemiología, Psicometría o Econometría se presentaban hechos de manera que la virtual coincidencia entre asociación causal y correlación era asumida. El mero sentido común nos advierte de esta falacia, y ya Pascua nos decía en esta obra (en su página 230):

"Mas nótese bien esta advertencia desde el comienzo: correlación no significa ni forzosamente supone causalidad [...]".

Advierte además en el pie:

"Pero la proposición inversa, esto es la de que dos variables que no se correlacionan son necesariamente independientes, no es, sin embargo, válida".

Dos últimos aspectos cabe resaltar de este capítulo, que creemos lo engrandece. Por un lado, dejar un apartado para las llamadas *Correlaciones espurias*, o *aquellas correlaciones que pueden surgir por la acción de otros elementos o factores que aquellos que deberían entrar en la determinación de los verdaderos coeficientes de correlación*. Para argumentar el concepto y desarrollarlo se ampara en el trabajo del primer hombre que utilizó el termino en Estadística, en 1897, que no podía ser otro que el eminente Karl Pearson, en su obra "On the Form of Spurious Correlation which may arise when indices are used in the Measurement of Organs" *Pro. Roy. Soc.*, vol. LX, 1897. Y que Pascua cita.

El otro aspecto sobresaliente se refiere a la inclusión de un apartado de correlación múltiple y parcial, metodología hoy algo más en boga por lo ya señalado a lo largo de este escrito (la incorporación del ordenador en la mesa del estadístico), pero de alta especialización en su época. Y como siempre, nos sorprende su erudición; a la hora de cerrar el apartado remite al lector deseoso de ampliar a los trabajos de Hotelling sobre análisis de componentes principales de 1935 (*The analysis of complex of statistical variables into principal components. J. Educat. Psychology, 1935*) y de Thurstone de 1943 (*Multiple factor analysis, University of Chicago Press*). Ambas obras son pioneras del análisis multivariante, que había nacido años antes, en concreto de nuevo con Pearson, quien en 1901 publica sus primeros trabajos sobre análisis de componentes principales. Cuando los cita Pascua en su obra eran desconocidos en la Ciencia aplicada española y mucho más en la Biomedicina; incluso actualmente son de restringido uso. De nuevo, acierto pleno de D. Marcelino.

Capítulo IV

Titulado Probabilidad y Curva Normal, el Capítulo comienza con algo que, necesariamente, conociendo la trayectoria de la obra, no nos debe de sorprender: un escrupuloso intento de definir un concepto tan crucial en la Epistemología, en la Filosofía de la Ciencia y en la Estadística como es el de Probabilidad. Para ello Pascua no se olvida de nadie a la hora de analizar las obras de los estudiosos, y fundamentalmente vuelve a los clásicos, se ayuda de: Laplace, Von Mises, Edgeworth, Arne Fisher, Von Kries, Boole, Bernouilli, Jeffreys, Poincare, Borel, Darmais, etc..e incluso de nuestro don Miguel de Unamuno, del que rescata la siguiente frase *el cálculo de probabilidades no es más que la racionalización del azar, de lo irracional*.

Con estos cimientos construye el capítulo, donde nos encontramos las definiciones de probabilidad dada por los diferentes autores para las diferentes épocas.

Se repasa la Teoría Combinatoria y se postulan los teoremas fundamentales del cálculo de probabilidades, siempre con ejemplos aclaratorios, teniendo en cuenta al lector medio a quien va dirigido su manual. Continúa con la presentación de las distintas distribuciones teóricas como la Binomial, la de Bernouilli y la Normal, este caso, se detiene en su ecuación, propiedades, áreas, gráfica de probabilidad normal, prueba de la normalidad de una distribución y distribuciones no-normales, para entrar a explicar luego la distribución de Poisson, la binomial negativa, y terminar citando otras distribuciones teóricas.

Capítulo V

Entramos de lleno en la llamada Estadística Inferencial, capítulo que estructura Pascua en dos partes, una primera dedicada al Muestreo y que titula *Teoría de la Toma de Muestras*, y una segunda y extensa titulada *Situaciones de comparación*, donde aborda las pruebas de hipótesis.

En la primera parte se centra el autor a dar las definiciones de universo, muestra y toma de muestra, refiriéndose a ellos de la siguiente manera,

"A una parte, que forzosamente tiene que ser finita o limitada en su tamaño, extraída del referido universo, denominásele *muestra* de él, y al mecanismo o proceso mediante el cual ella es obtenida conócese, en técnica estadística, como *toma de muestras*".

Explica los diferentes tipos de toma de muestras (hoy decimos muestreo, palabra que a Pascua no le gusta y argumenta a pie de página en su contra) dando importancia a la toma de muestra aleatoria apoyándose en la cita de autores tan importantes como Yule y Kendall, cuando dicen:

"La experiencia ha mostrado, de hecho, que el ser humano es un instrumento extraordinariamente pobre para realizar toma de muestras al azar. Y en donde haya una posibilidad cualquiera de elección o juicio, muy fácilmente surgirá parcialidad".

Tras las definiciones previas, esta primera parte continúa desarrollando los diferentes apartados. Distribución al azar de las medias aritméticas (lo que hoy llamamos distribución muestral), cuya conclusión final queda enunciada de manera magistral en esta proposición:

"La media \bar{x} rendida por la muestra es la evaluación de máxima verosimilitud de la media del universo μ ".

Se explica también en este apartado el error estándar de la media y su calculo, la diferencia entre dos medias aritméticas y su *significación*, se

introducen los importantes conceptos de *hipótesis nula y alternativa, errores de la primera clase y errores de la segunda clase* (error tipo I y error tipo II) y expone el concepto de niveles de significación, advirtiendo:

"El juicio de significación estadística se forma en términos de probabilidad de ocurrencia, no de un modo absoluto, categórico ni apodíctico".

Sin comentarios, aviso para los que de manera cegadora se agarran a la significación estadística para lanzar *verdades* científicas.

A continuación va presentando las distintas situaciones de comparación y las diferentes pruebas; prueba t, prueba de los signos, método rápido de Wilcoxon, etc. Hay un comentario en la página 341 que nos parece magistral, orientado a combatir un adocenamiento absolutamente vigente hoy, y fomentar el espíritu crítico en investigación médica y en Estadística aplicada, que reproducimos:

"Reiteramos una vez más que, como impone la lógica, en el tratamiento de esta clase de problemas las deducciones obtenidas con las pruebas estadísticas afectan estrictamente a la complejión estadística de la cuestión tal como ella aparece planteada, pero no solventan de por sí los factores de causalidad que los temas llevan envueltos, pues el enjuiciamiento y dilucidación de éstos trasciende, en general, a la mera esfera de pruebas estadísticas. Ellas pueden ayudar poderosamente al investigador en aquella tarea y finalidad suya, pero sería erróneo el adjudicarlas, en general, una plena virtud cognoscitiva, sentenciadora y finalista".

El capítulo continúa analizando la diferencia entre proporciones y presenta un interesante apartado titulado la experiencia clínico-quirúrgica, donde se desarrollan los elementos necesarios de diseño para la evaluación y experimentación en clínica y terapéutica. Analiza también el problema del calculo del tamaño de la muestra necesario para que los resultados puedan ofrecer confianza estadística.

También desarrolla en este capítulo de manera excelente el concepto de límites o intervalos de confianza, nos presenta los hoy ya obsoletos diagramas de Clopper y Pearson, para calcular intervalos de confianza de una proporción dada en una muestra concreta, publicados en *Biometrika*, vol. XXVI del año 1934. Hoy, que se está dando tanta importancia a los intervalos de confianza como sucedáneos de las pruebas de hipótesis en las publicaciones científicas, nos sorprende que ya en el año 65, Pascua exaltara su utilidad con comentarios como el que aparece en la página 414:

"Añadamos que la ventaja que para muchos presenta el establecimiento de límites de confianza sobre la práctica de pruebas de *significación* es la de que el

intervalo viene dado en las mismas unidades que las medidas originales de las observaciones, y más inteligible, por tanto".

El capítulo termina con dos apartados poco usuales en los libros de Bioestadística, por un lado, el titulado *Errores en los recuentos sanguíneos*, donde encara desde un punto de vista estadístico dicho problema. Y por otro, una breve aproximación al *Análisis Secuencial*.

Capítulo VI

Dedica Pascua, este capítulo sexto por completo a la distribución χ^2 (Ji-cuadrado), justificando su extensión por la *gran importancia de esta técnica y sus numerosas aplicaciones a problemas biológicos*.

Recuerda que la prueba de χ^2 fue presentada por primera vez por Karl Pearson, al objeto de *probar y juzgar de la divergencia que existe entre las observaciones actuales y los valores teóricos que las corresponderían*. Considera Pascua que poder evaluar esta divergencia, es *uno de los más importantes pasos en el progreso científico*.

Continúa el capítulo presentando el concepto de grados de libertad y las aplicaciones más comunes de la prueba en el ámbito biosanitario. Empieza por presentar el caso concreto del análisis en tablas 2 x 2, y mediante un acertado ejemplo sobre vacunaciones, explica el cálculo de las frecuencias teóricas asociadas a una tabla de contingencia.

Luego se presenta la corrección de Yates para aquellos casos en los que algunas frecuencias esperables son pequeñas. El texto aborda también la prueba exacta de Fisher y termina profundizando con el tratamiento de tablas de 2 líneas por k columnas y de m líneas por k columnas, siguiendo su terminología.

Un comentario acertado a toda sistemática de aplicación de la prueba χ^2 y en el que deberíamos insistir, dado el olvido aparente que hacen muchos investigadores en la actualidad, reza de la siguiente manera (página 432):

"Pero nótese, desde el comienzo, que la χ^2 no es una medida del grado de asociación".

Efectivamente, el grado de asociación entre variables no es equiparable con la significación estadística medida por un test de hipótesis; son otras

medidas las que determinaran dicha asociación, dentro de un contexto metodológico correcto.

Capítulo VII

Termina el libro con este capítulo dedicado al Análisis de la Variancia, inspirado en los trabajos de R.A. Fisher, cuya función Pascua empieza, con su proverbial lenguaje, por aclarar que se realiza: [...] *para efectuar confrontaciones, [...], no sólo entre dos series, sino simultáneamente entre cualquier número de ellas.* Y profundiza, *...pretendiéndose estimar sobre la base de documentación de sus cifras la heterogeneidad de la variabilidad o discrepancia que existe entre ellas, y su significación o no, objetivo, sin duda, uno de los más habituales y de más interés en el progreso de las ciencias médico-biológicas.*

Más adelante, se expone la subdivisión de la variabilidad total y la tabla de la distribución F. Presenta el análisis de la variancia para una *clasificación sencilla de las variables y para doble clasificación de las variables.* Todo los conceptos son ejemplificados como es norma habitual, y ya hemos comentado, a lo largo del texto.

Termina el capítulo adquiriendo alta relevancia con la presentación, muy poco frecuente en otros textos, de un apartado que aborda la experimentación factorial y la planificación de experiencias, citando y apoyándose en los tratadistas clásicos en la materia, Snedecor, Yates, Cochran y Cox.

Finaliza dándonos dos ultimas pruebas, una para evaluar la homogeneidad de variancias, en concreto la prueba de Bartlett, y otra para evaluar la rectilinearidad de la regresión.

El libro se completa con dos extensos apéndices, el primero con dos apartados, uno sobre representación algebraica y ajuste de curvas; y otro sobre métodos de interpolación. Allí se repasan los métodos elementales de ajuste, el método de mínimos cuadrados, las curvas exponenciales y logarítmicas, los métodos de interpolación aritmética y geométrica y sus aplicaciones en Demografía, el concepto de curva logística (tan de moda en la actualidad), etc. En el segundo apéndice, con el título de contraste biológico, ofrece un repaso fundamentalmente a los métodos cuantitativos aplicados a la experimentación biológica y terapéutica, y desarrolla conceptos como el de Dosis Letal 50 y su cálculo por los métodos de Trevan, Reed-Muench, Kärber y Gaddum. Aporta datos importantes sobre el diseño de ensayos clínicos y su tratamiento

estadístico, y termina con una aportación bibliográfica al tema de los ensayos terapéuticos.

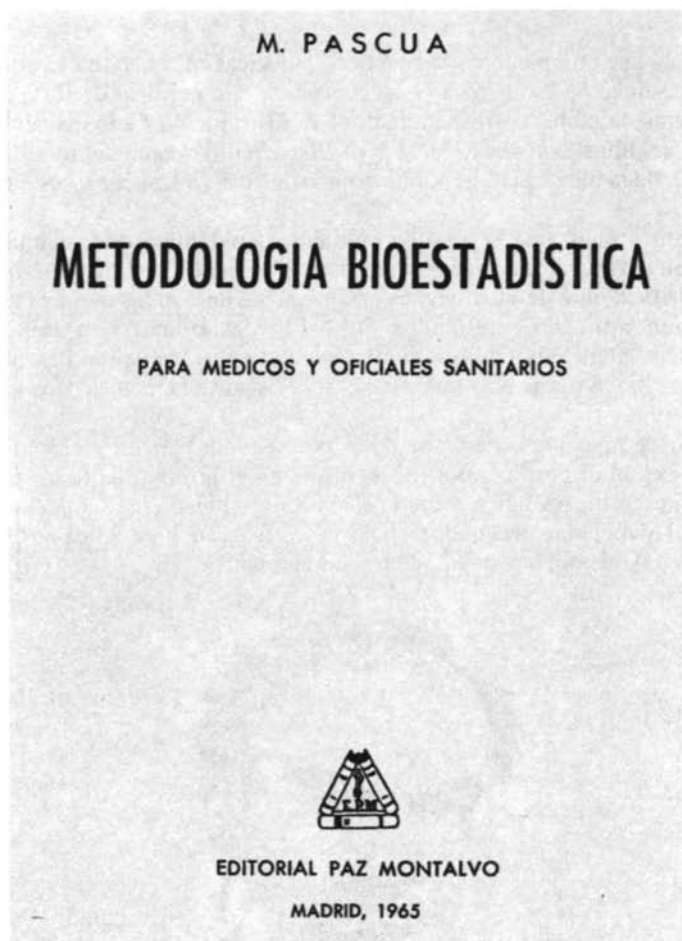
La obra se completa con las tablas estadísticas más habituales empleadas en la investigación biomédica, y así reproduce con permiso de R.A. Fisher y sus editores, la tablas correspondientes a la distribución t y la distribución χ^2 tomadas del libro *Statistical Methods for Research Workers*, del referido autor. Y la tabla de la distribución F , con permiso de G.W. Snedecor y sus editores.

Estamos pues ante la obra de un médico epidemiólogo adelantado a su época, que hoy se nos hace un clásico al leerlo. Su libro Metodología Bioestadística, obra de madurez, es posiblemente uno de los mejores textos de Estadística aplicada publicados en habla castellana, con importantes aportaciones a la Demografía sanitaria, que hace que Marcelino Pascua se nos presente como la figura más importante de la Estadística médica española.

La obra aquí presentada, se editó por segunda vez en 1973 con pocos cambios según el propio autor nos confiesa en el prólogo de la misma, pero con un interesante apéndice sobre la *electrónica en medicina y sanidad*, donde aborda el papel del ordenador en la investigación biomédica y el futuro alentador del mismo, que hoy podemos comprobar.



Figura 1. D. Marcelino Pascua Martínez (1897-1977).
Fotografía tomada de la portada del libro de Ponencias del V Encuentro
Marcelino Pascual. Escuela Andaluza de Salud Pública, Granada 1995



BIBLIOGRAFIA

BERNABEU I MESTRE, J. (1992) "Marcelino Pascua desde la perspectiva histórica". En: *I Encuentro Marcelino Pascua. Estadísticas demográfico-sanitarias*. Madrid, Instituto de Salud Carlos III, 11-15.

CLAVERO GONZALEZ, G. (1992) "La obra de Marcelino Pascua en su entorno". En: *I Encuentro Marcelino Pascua. Estadísticas demográfico-sanitarias*. Madrid, Instituto de Salud Carlos III, 17-23.

GUTIERREZ CABRIA, S. (1994) *Filosofía de la Estadística*. Servicios de Publicaciones Universidad de Valencia.

MACDONELL, W.R. (1902) "On the Influence of Previous Vaccination in cases of Smallpox". *Biometrika*, Vol. 1, 375-383.

PASCUA MARTINEZ, Marcelino (1965) *Metodología bioestadística para médicos y oficiales sanitarios*. Madrid, Editorial Paz Montalvo, 627 pp., DL. M-13187-65.

RODRIGUEZ OCAÑA, E. BERNABEU I MESTRE, J. (1995) "El legítimo criterio aritmético. Los métodos cuantitativos en la Salud Pública española, 1800-1936". En: *V Encuentro Marcelino Pascua. La Epidemiología y la Estadística*. Granada, Escuela Andaluza de Salud Pública, 9-33.