

EVOLUCIÓN DEL CONCEPTO DE INFERENCIA ESTADÍSTICA. INFLUENCIA EN SU ENSEÑANZA

Christiane Ponteville, Cecilia Crespo Crespo

Instituto Superior del Profesorado "Dr. Joaquín V. González". (Argentina)

Facultad de Farmacia y Bioquímica. Universidad de Buenos Aires. (Argentina)

chponteville@gmail.com

Palabras clave: inferencia estadística, evolución histórica, probabilidades, pruebas de hipótesis

Key words: inferential statistics, historical development, probabilities, hypothesis test

RESUMEN

Podemos establecer dos tipos de tendencias en la enseñanza de conceptos de la inferencia estadística: interpretarlos como modelos matemáticos excluidos del contexto o como la aplicación de un algoritmo. Cualquiera de las dos involucra perder algún aspecto en la generación del concepto pues no tiene en cuenta su verdadera génesis en la historia de la ciencia. Proponemos una revisión de los diferentes aspectos que participaron en la evolución de los fundamentos de las probabilidades y la estadística para entender los marcos históricos y epistemológicos en los cuales se desarrollaron las pruebas de hipótesis como parte de la ciencia estadística.

ABSTRACT

We can establish two trends in inferential statistics teaching: as mathematical models excluded from the context or as an algorithm application. Either of them involves some concept generation lose due to not have into account its real genesis in science history. Our proposal is to review different aspects of probabilities and statistics fundamentals evolution to understand the historic and epistemological frames in which hypothesis test were developed as part of statistic science.

■ La estadística en nuestra sociedad y en la educación

La estadística se encuentra instalada dentro de nuestra sociedad como un instrumento de validación siendo utilizada y aplicada en una amplísima variedad de áreas que necesitaron de un desarrollo cualitativo significativo de ella. Así, la estadística marca su utilidad para el resto de las ciencias y se desarrolla como área de conocimiento independiente con un gran desarrollo teórico. Es transversal a una extensa variedad de disciplinas, desde el control de calidad hasta las ciencias sociales, desde las ciencias de la salud hasta la física. Podemos decir que el método estadístico es la matemática social por antonomasia (Boyer, 1994). Esta realidad se refleja en las instituciones educativas de nivel superior, donde, por ejemplo, existen departamentos académicos de matemática y estadística en forma paralela y la estadística se enseña en departamentos tan variados como los de medicina social, psicología, relaciones del trabajo, economía, entre otros.

En este marco, la estadística puede ser considerada no como una rama de la matemática, sino como una área de conocimiento en estrecha vinculación con ella que toma un status parecido al que tienen, por ejemplo, las nuevas ciencias relacionadas con la informática. Según sus enfoques, se ha superpuesto con, por ejemplo, la teoría de la decisión poniendo el énfasis en la posibilidad de hacer predicciones cada vez más acertadas y con las ciencias de la información en el procesamiento de datos. Es así como, la estadística, durante el siglo XX, ha sido considerada parte de la base del método científico y una estrategia metodológica fundamental.

De esta forma la sociedad del conocimiento demanda no solamente incluir conceptos estadísticos en la enseñanza de los diferentes niveles sino un proceso social de culturalización estadística debiéndose incorporar a los diferentes niveles de la educación conceptos y prácticas relacionados con las probabilidades y la estadística. Aunque ya no se discute la importancia de que exista formación estadística de los estudiantes en todos los niveles educativos, lo que genera permanente debate entre los docentes responsables son los resultados desalentadores en el aprendizaje de sus alumnos siendo estos niños, adolescentes, jóvenes o estudiantes de cualquier edad. Esta inquietud genuina de los docentes a largo de los años y de todos los niveles de la instrucción ha provocado que una comunidad cada vez más grande de investigadores trate de encontrar respuesta a la diferencia que existe entre lo que es enseñado y lo que es aprendido existiendo un aumento notable de publicaciones, propuestas de diseños curriculares e investigación relacionados con estos temas (Batanero, 2002).

Este interés, unido al rápido desarrollo de la estadística como ciencia y como herramienta fundamental en otras áreas, impulsado significativamente por el desarrollo y la difusión de los métodos computacionales con el crecimiento de su potencia y velocidad de procesamiento de datos así como por las posibilidades de envío de la información, propone un gran desafío. Todo ello ha facilitado el uso de la estadística a un número cada vez mayor de alumnos, provocando, en consecuencia, una gran demanda de formación básica en esta materia.

Uno de los aspectos a ser tenidos en cuenta en la enseñanza de la estadística es que los fenómenos aleatorios aparecen fuertemente en la sociedad actual. De esta forma, aunque, tradicionalmente, la mayor parte de las aplicaciones se basaban en los juegos de azar, ya que estos son corrientes y producen, en general, espacios muestrales fácilmente visibles y en gran número finitos, si esperamos

que los alumnos valoren el rol de la probabilidad y estadística, es fundamental acercarse a problemas del mundo biológico, físico, social y político: las características genéticas, la previsión climática, el resultado de las elecciones, el crecimiento de la población, la extinción de las especies, el efecto de los hábitos alimenticios o las drogas sobre la salud, la extensión de enfermedades, los resultados deportivos, el índice de precios o el censo de la población son claras evidencias del mundo que los rodea. De esta forma, aparece fuertemente el concepto de modelo dando una oportunidad muy importante para entender y aprender a medir dentro de las probabilidades. Para entender la idea de modelo como herramienta en la sociedad actual, en este trabajo se busca realizar una descripción general de algunos de los hechos de la evolución de las probabilidades y la estadística a través de la historia con el fin de entender los marcos históricos y epistemológicos en los cuales se desarrollaron las pruebas de hipótesis como parte de la ciencia estadística.

■ El comienzo de las ideas estadísticas

Desde los inicios de la civilización podemos identificar estadísticas sencillas pues se han encontrado símbolos y representaciones en rocas de cuevas que muestran que se estaban contando personas, ganado o alguna otra pertenencia.

De esta forma el origen de la estadística en la civilización puede explicarse como la recopilación de datos sobre poblaciones con el fin de informar algo sobre la realidad de un grupo humano. Por ejemplo, existen tablillas con datos sobre producciones agrícolas y los objetos de canje en la civilización babilónica y registros de datos para la construcción de pirámides en la civilización egipcia. En China podemos encontrar registros numéricos relacionados con censos con anterioridad al año 2000 a.C. con datos acerca de la actividad agrícola y comercial.

Se encuentran registros de censos periódicamente realizados hacia el año 594 a.C. en la civilización griega con el objetivo de recaudar impuestos, establecer divisiones de tierras y la existencia de recursos y hombres. Los emperadores romanos, establecieron en forma estructural el uso de los recursos de la estadística en su poderosa organización política, realizando censos de la población y de las riquezas contenidas en las tierras conquistadas con el objetivo de la recaudación de impuestos. En América precolombina los incas hicieron uso de elementos estadísticos utilizando el quipu como instrumento de archivo de información.

Siguiendo con la visión romana hasta los siglos XVII y XVIII, la estadística se fue identificando con el arte de gobernar los estados ya que mostraba una fuente de información muy importante para la toma de decisiones. De esta manera se pudo estudiar las tasas de mortalidad y esperanza de vida mediante registros estadísticos de Londres desde finales del año 1500 y resolver problemas de beneficios perpetuos en temas de seguros a finales del año 1600. Aquí, se instala la lectura del ajuste de curvas a través de datos apareciendo las primeras ideas de continuidad e inferencia estadística. La creación de sociedades de estadística permitió realizar comparaciones entre países determinando factores de crecimiento económico.

■ Aparición de las probabilidades. Los juegos de azar

La probabilidad matemática tiene sus orígenes en los juegos de cartas y de dados. Estos juegos de azar, populares desde tiempos antiguos, conducen a los primeros estudios sobre el conteo del número de posibles resultados de lanzar un dado varias veces y el problema del reparto de apuestas que consistía en la distribución de las ganancias entre jugadores antes de la finalización del juego, este problema abordado por varios autores.

La primera publicación importante relacionada con los juegos de azar fue debida a Girolamo Cardano y se llamó “El libro de los juegos de azar”. Esta obra, un manual para jugadores, contiene una pequeña parte dedicada al estudio del azar en la cual introduce la idea de asignar una probabilidad p con valor entre 0 y 1 y esbozó una idea rudimentaria de la Ley de los Grandes Números. Además, Galileo Galilei establece la teoría de la medida de errores diferenciando los errores sistemáticos y los errores aleatorios y que la mayoría de las mediciones se agrupan alrededor de un valor. Con estas ideas, estableció las bases del nacimiento de la estimación como la conocemos actualmente.

Resumiremos cuatro contribuciones, desde el siglo XVIII y principios del siglo XIX, que sobresalen por su importancia y muestran la evolución de las ideas vinculadas con una nueva visión científica de las probabilidades: Bernoulli vio la importancia social de las probabilidades y explicitó algunas probabilidades inversas y enunció el resultado de que aumentando suficientemente el número de observaciones se puede conseguir cualquier grado de precisión prefijado; Abraham de Moivre ideó métodos para aproximar las funciones de números grandes y sus observaciones de las sumas de los términos de un desarrollo en serie binómica, se suele decir que fue el autor de la curva de la distribución normal; Thomas Bayes trabaja sobre el teorema de la probabilidad inversa que explica probabilidades de causas desconocidas deduciéndolas de acontecimientos observados siendo el primero que empleó la probabilidad matemática inductivamente es decir razonando de lo particular a lo general, o sea del individuo a la masa; Pierre Simon Laplace muestra que la teoría de probabilidades no es más que sentido común reducido al cálculo con argumentos que generan polémica entre sus contemporáneos. Laplace utiliza el análisis matemático en forma sistemática de manera amplia y consecuente a lo que, por su naturaleza, era un capítulo de las matemáticas combinatorias. El concepto de medición y la posibilidad de obtener diferentes resultados hicieron que Gauss considerara que cuando se dan un número cualquiera de medidas “igualmente dignas de confianza” de una magnitud incógnita el valor más probable es su media aritmética. De esta forma, frases como “igualmente dignos de confianza”, “igualmente probable”, generaron muchos debates sobre el significado de la probabilidad (Bell, 1995).

■ Génesis y desarrollo del método estadístico

La palabra estadística se atribuye a un profesor de Gotinga, Gotfried Achenwall y su irrupción en el mundo científico través de concepciones diversas y el desarrollo de su método pueden ser analizados consignando algunos de sus episodios sobresalientes.

El primer análisis estadístico de un censo nacional fue el que hizo Lambert Adolphe Quetelet del primer censo de Bélgica refiriendo la influencia que tenían sobre la mortalidad diferentes variables como la edad, el sexo, la estación, la profesión y la situación económica. Habla de “L’homme moyen” (hombre medio). Se necesitaron setenta años para que las pruebas de inteligencia realizadas en el ejército de los

Estados Unidos de América en 1917 se aceptaran sin oposición como instrumento útil para la educación y la criminología. Entre los partidarios de estos métodos se encontraba Florence Nightingale, que de acuerdo con Pearson decía que “para comprender lo que Dios piensa hay que estudiar las estadísticas, ya que son estas las medidas de sus propósitos”. Otro de los científicos que cultivó estos métodos relacionados con los seres vivos fue Gregor Mendel que, en sus trabajos sobre cruces de garbanzos, fue por primera vez la herencia genética con la matemática.

Una síntesis de los dos anteriores se dio en Francis Galton. Sus intereses eran muy diversos y sus gráficas abarcaban desde los garbanzos y las polillas hasta los perros y los seres humanos. Era un matemático biólogo que crea el moderno método matemático introduciendo la línea de regresión, siendo el primero en explicar el fenómeno de la regresión a la media, fue pionero en el uso de la distribución normal e introdujo el concepto de correlación. Walter Weldon cristaliza los trabajos de Galton ideando un método para calcular los coeficientes de correlación. Aquí, la velocidad de las máquinas de calcular ampliaron las aplicaciones de los métodos estadísticos permitiendo estudiar, por ejemplo, los fenómenos de regresión.

Karl Pearson fue el que estableció la disciplina de la estadística matemática. Desarrolló una intensa investigación sobre la aplicación de los métodos estadísticos introduciendo técnicas matemáticas nuevas en diferentes campos. Introdujo el método de los momentos y definió la curva normal y la desviación normal, popularizó el criterio de la “ji-cuadrado” midiendo la discrepancia entre una distribución observada y otra teórica llamada bondad de ajuste. En 1935, en una comunicación oral Pearson dijo: “La probabilidad está tan conectada a la estadística que, a pesar de ser posible ser enseñadas por separado, separarlas sería algo forzado.” En 1901 Galton, Weldon y Pearson crean *Biometrika*, revista dedicada a la biometría definida por ellos como “la ciencia de la medida de la vida”.

Es necesario mencionar que, la teoría de probabilidades y la estadística han estado a lo largo del siglo XX, no solamente estrechamente unidas entre sí, sino también a otro campo que constituye una de las características diferenciales más notables de nuestra época, la de la dependencia cada vez mayor de las ciencias de la computación.

Aunque la ji-cuadrada de Pearson puede ser considerada como su comienzo, la obra de Ronald Fisher le otorga a la estadística el carácter de ciencia independiente. Sus aportes fueron muchos y destacan la creación de la genética biométrica, la metodología del análisis de varianza, considerándola superior a la de la correlación. En uno de sus artículos mostraba que la herencia de rasgos, medibles por variables continuas, era consistente con los principios mendelianos. Fisher acuñó gran parte de los términos que usamos hoy en estadística: parámetro, estadístico, varianza, hipótesis nula, verosimilitud, test y nivel de significación, entre otros. Un punto clave de sus avances fue establecer claramente una diferencia entre valores muestrales y poblacionales que se reflejó en la notación utilizada al usar las letras griegas para la población y las latinas para la muestra, que usamos en la actualidad.

■ Las pruebas de hipótesis irrumpen en el mundo científico

Las pruebas de hipótesis fueron creadas en el período entre 1915 y 1933 como el resultado de dos visiones; la de Fisher por un lado y la de Jerzy Neyman y Egon Pearson por otro. Las dos visiones tuvieron su origen en la famosa prueba de Ji -cuadrado presentada por Karl Pearson. Neyman trabaja con Pearson

desarrollando el punto de vista pearsiano hacia una teoría de decisiones en contraposición a la visión de Fisher centrada en el análisis de datos. Estas dos posiciones se desarrollaron a partir de posiciones filosóficas contrapuestas. Esta historia ha tenido, controversias, disputas personales y discordancias científicas. Resumiendo, la prueba propuesta por Fisher es un test de significación mientras que la propuesta por Neyman-Pearson es una regla de decisión (Ponteville, 2014).

El enfoque teórico de Fisher para abordar una prueba de hipótesis se apoya en la realización de una inferencia inductiva. Se plantea una hipótesis que denomina nula, consistente en suponer que la muestra proviene de una población hipotética infinita, con distribución muestral conocida. Se calcula un estadístico a partir de la muestra, que posee una distribución conocida bajo H_0 . Para dicho fin se calcula un estadístico a partir de los resultados de una muestra, cuya distribución de probabilidad queda fijada cuando se asume la hipótesis nula como verdadera. La misma es rechazada cuando la estimación muestral difiere de la media de la distribución con una probabilidad menor al nivel de significación. La prueba de significación de Fisher se realiza para evaluar la fuerza que tenemos en contra de la hipótesis nula. El hecho de no rechazar la hipótesis nula es interpretado como que es aceptada por el momento sobre una base provisoria.

Pearson y Neyman plantean un proceso de decisión de tipo deductivo, en un contexto de muestreo repetido, diseñado a priori. Su objetivo es obtener pruebas de significación óptima. Se fijan las dos hipótesis: hipótesis nula (H_0) e hipótesis alternativa (H_a). Se calcula el estadístico con los datos. Se decide rechazar o no la hipótesis nula comparando con el nivel de significación elegido. De esta manera, el rechazo de la hipótesis nula no implica que esta sea falsa. Este contraste de hipótesis toma en consideración los errores que pudieran cometerse al rechazar o no rechazar una hipótesis nula: error de tipo I y error de tipo II. Por un lado el error de tipo I, que consiste en rechazar una hipótesis nula que es verdadera, y el error de tipo II a su vez, que plantea no rechazar una hipótesis nula que es falsa con probabilidades llamadas α y β respectivamente. En esta teoría el valor P es reemplazado por una regla de decisión R basada en la noción de nivel de significación de la prueba (α). El problema de probar hipótesis se traduce entonces en elegir entre dos hipótesis, de acuerdo con los datos muestrales, pero con conocimiento de los errores que pueden estarse cometiendo.

En el criterio de Fisher, el valor de P se establece a posteriori, es decir sobre la base de los datos; en el de Neyman y Pearson, los datos se obtiene con una confiabilidad dada a priori por los errores α y β .

Entre los autores de las dos teorías, se generó en su época una controversia que involucraba algo más profundo que la manera de realizar los cálculos y las decisiones a tomar. En la mirada de Neyman y Pearson, la prueba de hipótesis es una regla de decisión en el contexto de muestreo repetido. Para Fisher, la finalidad de las pruebas de hipótesis no consistía en tomar decisiones irrevocables, sino provisionarias y sólo aceptaba la visión de Neyman y Pearson para problemas comerciales y tecnológicos, pero no para la validación de hipótesis científicas, pues en su opinión, en estos contextos el muestreo repetido genera confusión y puede no existir una toma de decisión final. Esta controversia se mantuvo durante la vida de Fisher. Posteriormente, la teoría de pruebas de hipótesis evolucionó realizando una fusión de estas dos visiones (Ponteville, 2015).

■ Reflexiones finales

Analizando la evolución de las diversas ideas que dieron nacimiento tanto a la estadística como a las probabilidades y su posterior desarrollo como áreas de conocimiento científico tanto interrelacionadas como en su desarrollo individual, podemos decir que las ideas germinales de las pruebas de hipótesis se encuentran establecidas por tres áreas de pensamiento humano:

- La teoría de probabilidades, producto de la matematización de los juegos de azar
- La ciencia experimental del siglo XIX, buscando explicaciones científicas a hechos observados
- La aritmética política, que comprende todo tipo de datos relacionados con los estados

La unión de la aritmética política junto con la Teoría de las Probabilidades, influenciadas por los científicos experimentales genera el concepto de prueba de hipótesis del cual disponemos y hacemos uso en la actualidad. Resumiendo, la genética era discreta y la evolución asumía continuidad, la primera trabaja con muestras pequeñas y la otra con muestras grandes. Estas diferencias en términos estadísticos enmarcan el trasfondo de la conocida como la controversia de Pearson-Fischer.

A través de esta mirada de las pruebas de hipótesis inmersas en la enseñanza de la estadística y esta como forma de alfabetización estadística podemos observar la necesidad de instalar una mirada sobre la enseñanza de la estadística actual en función de los objetivos buscados por el curso impartido, trascendiendo el esquema actual de formación tradicional (Crespo Crespo, 2012). Como es necesario que ocurra, los resultados obtenidos nos dejan preguntas de cómo identificar diferentes planos de análisis tanto dentro como fuera de las aulas: análisis de la inclusión de influencias externas propias del área científica estudiada en el aula, creencias sociales respecto de cómo se toman las decisiones en diferentes ámbitos de la sociedad, validación de resultados científicos a través de pruebas de hipótesis, prácticas sociales asociadas al valor P como elemento de decisión, influencia de modelos informáticos para enseñar pruebas de hipótesis en diferentes escenarios, entre muchas otras.

■ Referencias bibliográficas

- Batanero, C. (2002). *¿Hacia dónde va la educación estadística?* Disponible en: <http://www.ugr.es/~batanero/ARTICULOS/CULTURA.pdf>
- Bell, E.T. (1995). *Historia de las matemáticas*. México: Fondo de Cultura Económica.
- Boyer, C. (1994). *Historia de la matemática*. Madrid: Alianza Editorial.
- Crespo Crespo, C. (2012). Socioepistemología. M. Pochulu y M. Rodríguez (comp). *Educación matemática. Aportes a la formación docente desde diferentes enfoques teóricos* (pp. 91-114). Villa María y Los Polvorines: Universidad de General Sarmiento y Editorial Universitaria de Villa María.
- Ponteville, Ch. (2014) ¿Para qué enseñamos estadística? En P. Lestón (Ed), *Acta Latinoamericana de Matemática Educativa* 27, 517-525. México: Comité Latinoamericano de Matemática Educativa.
- Ponteville, Ch.. (2015). *El rol de las Argumentaciones Estadísticas: Pruebas de Hipótesis*. Tesis de maestría no publicada. Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada del IPN. México.