R.A. Fisher: su contribución a la Ciencia Estadística

FCO.JAVIER GIRÓN*
MIGUEL A. GÓMEZ VILLEGAS**

Wilde atribuye la siguiente broma a Carlyle: una biografía de Miguel Angel que omitiera toda mención de las obras de Miguel Angel. Tan compleja es la realidad, tan fragmentaria y tan simplicada la historia, que un observador omnisciente podría redactar un número indefinido, y casi infinito, de biografías de un hombre, que destacan hechos independientes y de las que tendríamos que leer muchas antes de comprender que el protagonista es el mismo. Simplifiquemos desaforadamente una vida: imaginemos que la integran trece mil hechos. Una de las hipotéticas biografías registraría la serie 11,23,33,...; otra, la serie 9,13,17,21,...; otra, la serie 3,12,21,30,39,... . No es inconcebible una historia de los sueños de un hombre; otra, de los órganos de su cuerpo; otra, de las falacias cometidas por él; otra, de todos los momentos en que se imaginó las pirámides; otra de su comercio con la noche y con las auroras. Lo anterior puede parecer meramente quimérico; desgraciadamente, no lo es. Nadie se resigna a escribir la biografía literaria de un escritor, la biografía militar de un soldado; todos prefieren la biografía genealógica, la biografía económica, la biografía psiquiátrica, la biografía quirúrgica, la biografía tipográfica. Setecientas páginas en octavo comprende cierta vida de Poe; el autor, fascinado por los cambios de domicilio, apenas logra rescatar un paréntesis para el Maelstrom y para la cosmogonía de Eureka. Otro ejemplo: esta curiosa revelación del prólogo de una biografía de Bolívar: "En este libro se habla tan escasamente de batallas como en el que el mismo autor escribió sobre Napoleón."

Jorge Luis Borges (1952), p.250-251. *Otras inquisiciones*

1. Algunos aspectos biográficos de R.A. Fisher

Ronald Aylmer Fisher (1890-1962) nació en East Finchley cerca de Londres, fue el más joven de siete hermanos, era gemelo de otro, pero sólo sobrevivió él. Su familia se dedicaba a los negocios pero él rompió la norma. Acudió a la escuela en Stanmore y posteriormente estudió en Harrow. Desde sus comienzos tuvo problemas con su vista. Algunos de sus discípulos han atribuido su habilidad para analizar mentalmente situaciones complicadas a este hecho. En su juventud tuvo prohibido leer con luz artificial y se le recomendó no fijar la vista demasiado.

* Académico Numerario. Catedrático de la Universidad de Málaga
** Catedrático de la Universidad Complutense de Madrid.

Cuando dejó Harrow las finanzas familiares no estaban muy bien; sin embargo, gracias a una beca pudo estudiar en Gonville en el Casius College (Cambridge) donde se graduó entre 1909 y 1912 y al año siguiente fue lector de Física Matemática. Durante el tiempo que estuvo en Cambridge también estudió Biometría y Genética.

En opinión de Kendall la manera de combinar observaciones en Astronomía fue lo que le llevó a interesarse por las distribuciones de probabilidad. Su primer artículo (1912) utiliza el valor absoluto para ajustar curvas de frecuencias siguiendo las ideas de K. Pearson.

Entre 1913 y 1915 trabajó en una compañía de inversiones pero sin ningún tipo de vocación. Durante la primera guerra mendual le dispensaron de hacer el servicio militar por su vista y entre 1915 y 1919 se dedicó a la enseñanza en escuelas públicas, trabajo este último que simultaneó con la investigación, pues en 1916 escribió un artículo demostrando que las teorías de Mendel no se ven rechazadas por los datos; lo referencian K. Pearson como estadístico biométrico y Punnet en sus aspectos genéticos.

En 1917 se casó con Ruth E. Guinnes con la que tuvo dos hijos y seis hijas. En 1919 se le ofrece trabajar bajo K. Pearson dirigiendo el laboratorio Galton, o bien asesorar estadísticamente en la Rothamsted Experimental Station. Optó por esta segunda salida, más en consonancia con su propia filosofía, ya que pensaba que, puesto que no podía contribuir al esfuerzo de guerra directo, el mantenimiento y mejora de las granjas inglesas podría ser su particular contribución.

En 1912 escribe *On the Mathematical Foundation of Theoretical Statistics* donde introduce la noción de modelo estadístico y los conceptos de consistencia, eficiencia, precisión, validación, verosimilitud e información. Al enviar este artículo a Gosset (Student), se consolida su relación con éste, que ya se había iniciado en 1912. Precisamente es Gosset quien le da la idea de representar las observaciones como un vector de n dimensiones, esto junto a la intuición geométrica de Fisher le va a permitir llegar a obtener la distribución del coeficiente de correlación bajo la hipótesis de normalidad. Por cierto que en la carta de contestación a Fisher, Gosset le comenta sobre que el *evidently* para Fisher se traducía en dos horas de duro trabajo para él. Inútilmente Student, gran amigo de Karl Pearson, trató de limar diferencias entre ambos. En su última época y relacionado con el concepto de aleatorización en diseño de experimentos estuvo a punto de romper con Gosset, sin embargo el obituario de éste fue realizado por Fisher en términos muy cariñosos.

Fisher se unió a Rothamsted en octubre de 1919 y allí desarrolló *el análisis de la varianza* (1921) y los principios del *diseño de experimentos* (1923,1924); en el primer artículo de 1921, todavía no conoce la distribución del cociente de los cuadrados (hoy conocida como distribución F de Fisher-Snedecor) por lo que aproxima su logaritmo por una distribución normal.

En 1925 aparece su primer libro *Statistical Methods*, que da la impresión de ser un manual para aprendices más que un libro de texto. No obstante, en esto radicó su éxito, ya que a lo largo del mismo anima a los lectores a trabajar los ejemplos. Problemas prácticos, técnicos, teóricos y filosóficos se discuten en el texto principalmente a través de ejemplos numéricos y en él se aleja de los matemáticos diciendo que en Estadística hay que hacer razonamiento inductivo en lugar de deductivo, para lo cual es necesaria, sin embargo, una gran formación matemática pero aplicada a los datos con los que se trabaja.

Fisher fue, en este sentido, un matemático aplicado dentro de la mejor tradición inglesa en este campo -con un gran dominio de la geometría n-dimensional, la integración múltiple, un dominio de lo que hoy se llaman funciones especiales (gamma, beta, hipergeométrica, etc.), la combinatoria, junto con una gran facilidad para los cálculos, la construcción de tablas y el uso de gráficos-, como lo fue también, p. ej., el astrónomo y estadístico H. Jeffreys.

A este respecto, debemos tener en cuenta que Mahalanobis decía de Fisher:

> "Trabajar mecánicamente en la técnica de la afirmación rigurosa era detestable para él, en parte por considerarlo pedante y, por otra parte, porque suponía un desperdicio del uso activo de la mente. Fisher opinaba que era más importante pensar desarrollando las cosas, incluso a expensas de errores ocasionales que una inteligencia despierta sería capaz de descubrir, que recorrer caminos bien conocidos al paso de la tortuga por senderos trillados con la ayuda de las herramientas más sofisticadas.
>
> ... Fisher pensaba de si mismo que era sencillamente un joven muy impetuoso e impaciente. Sin duda esto es cierto, pero era impaciente no porque era joven sino porque fue un genio creativo".

En 1929 es elegido miembro de la Royal Society. Opinaba que era un error elegir un alto porcentaje de personas por encima de los 50 años; él era partidario de la admisión de gente más joven.

En 1930 formula *The Genetical Theory of Natural Selection*, en la que apoya y modifica la teoría de la evolución de las especies de Darwin. Al año siguiente realiza su primer viaje a EE.UU., concretamente a Iowa, invitado por Snedecor.

En 1913 acepta la cátedra de eugenesia en el University College de Londres, trabajando a fondo en genética.

En 1938 viaja a la India invitado por Mahalanobis. Pasa en 1943 a ser profesor Balfour de genética en la Universidad de Cambridge, dedicándose durante esta época a desarrollar aspectos de la inferencia inductiva.

En 1943 viaja a EE.UU. por segunda vez como profesor visitante de la Universidad de Carolina del Norte. En 1947 funda la Sociedad Biométrica Internacional.

Durante 1953 y 1954 es presidente de la Royal Statistical Society y dedica su intervención presidencial a glosar las contribuciones de los primeros estadísticos.

En 1956 publica *Statistical Methods and Scientific Inference*.

Se retira en 1957 a la edad de 67 años y se marcha a Australia como investigador senior en el CSIRO (Commonwealth Scientific and Industrial Research Organization) que dirigía Cornish, donde muere en 1962 de un cáncer de boca.

Por su carácter y por lo innovador de sus descubrimientos Fisher mantuvo distintas polémicas. Quizás las más importantes fueron las que mantuvo con K. Pearson, con Neyman y E. Pearson y con casi todo el mundo respecto a la necesidad de la aleatorización.

Como hemos dicho, en su primera época trabajó sobre la teoría de la evolución de Darwin, lo que le llevó a escribir el ya citado artículo sobre las leyes de Mendel, que trató de publicar en Biometrika, revista dirigida por K. Pearson, y que fue rechazado. Tampoco lo logró en la Royal Society, por lo que por sugerencia del mayor Darwin (hijo de Charles Darwin) lo sometió a la Royal Society de Edimburgo, siendo aceptado allí en 1918). Los problemas con K. Pearson continuaron, pues en 1920 Fisher escribió sobre el error probable del coeficiente de correlación, que también sometió a Biometrika y fue rechazado, por lo que decidió no volver a someter nada a esta revista. Más tarde Fisher comentaba que había sido referenciado por un biólogo que no sabía Estadística y por un estadístico que no sabía Biología. La enemistad con K. Pearson aumentó cuando Fisher le corrigió los errores que había cometido en el número de grados de libertad de la χ^2 para el problema de ajuste entre unos datos y una distribución con parámetros desconocidos, señalando también que el método de los momentos, que había sido introducido por K. Pearson para estimar los parámetros, no bastaba para asegurar la convergencia del estadístico a la χ^2_{k-1-r} siendo r el número de parámetros estimados.

E. Pearson y J. Neyman eran admiradores del trabajo de Fisher. Gosset hizo la presentación de Neyman a Fisher en 1927 diciendo que "Neyman era la única persona a la que había oído hablar de *máxima verosimilitud* aparte de al mismo Fisher".

Sobre 1930, Neyman había desarrollado en Polonia el germen de la teoría de los intervalos de confianza, al mismo tiempo que Fisher desarrollaba la aproximación fiducial a la inferencia. En 1935 coincidieron Fisher, E. Pearson y Neyman bajo el mismo techo del departamento de Estadística del University College. Fisher empezó a pensar que las ideas de la estimnación por intervalos eran una copia de su aproximación fiducial. Para terminarlo de arreglar, Neyman contestó de forma no muy delicada a las críticas que Fisher hizo en aquel momento al trabajo de K. Pearson. Esto desencadenó las hostilidades, que pasaron a un ámbito más amplio cuando Neyman se trasladó a Berkeley (California) para hacerse cargo del departamento de Estadística. Como es conocido Fisher pensaba en $f(x|\theta)$ como una función en las dos variables y explotaba su consideración en función de θ para tratarla prácticamente como una distribución de probabilidad en θ.

E. Pearson y J. Neyman mediante el concurso de las variables pivotales desarrollaron la noción de región de confianza. Al pasar estas ideas a América, bajo la influencia de Neyman y sobre todo de A. Wald, se considera a la inferencia como una rama de la Teoría de la Decisión. Fisher mantenía que la inferencia en la ciencia no era una materia de decisión y que por tanto criterios basados en pagos de cualquier tipo no debían de ser utilizados. Por cierto que a pesar de que la distribución fiducial puede utilizarse en muchos casos como una distribución de probabilidad a posteriori respecto a una priori no informativa, Fisher no estaba de acuerdo con esta aproximación. Decía del libro de Harold Jeffreys *Theory of Probability*: "Tiene un error lógico en la primera página que invalida las restantes 395, y es que adopta el postulado de Bayes". Se refería al postulado que supone que si no se sabe nada sobre A y B, dos sucesos posibles, habría que suponer que $P(A)$ y $P(B)$ eran iguales. Esto le hizo decir posteriormente a L.J. Savage que "la aproximación fiducial era un intento de hacer la tortilla bayesiana sin romper los huevos bayesianos". A pesar de todo, en su libro de 1956 *Statistical Methods and Scientific Inference*, Fisher se muestra partidario de la aproximación bayesiana, cuando la información sobre θ sea lo suficientemente extensa para venir dada a través de una distribución de probabilidad, calculando la distribución a posteriori mediante el teorema de Bayes. En otro caso era partidario del argumento fiducial basado en un estadístico suficiente. Todo esto teniendo en cuenta que era un gran amigo personal de Jeffreys.

La polémica con Neyman y E. Pearson también se trasladó al contraste de hipótesis, ya que Fisher, como veremos más adelante, no suponía hipótesis alternativa a la hora de plantear un contraste; simplemente afirmaba que una observación acreditaba o no un valor de la hipótesis nula $H_0 : \theta = \theta_0$ sin ninguna referencia a cual pudiera ser H_1; eso sí, trabajando de forma condicional al valor observado. En el problema de contrastar la diferencia de medias en poblaciones normales $X \sim N(\theta_1, \sigma_1^2)$, $Y \sim N(\theta_2, \sigma_2^2)$ con varianzas desconocidas y distintas (problema de Behrens-Fisher), Fisher obtuvo una distribución fiducial para la distribución de $\theta_1 - \theta_2$ que dependía del parámetro $\delta = \sigma_1^2 / \sigma_2^2$ y que ponía de manifiesto que desde el punto de vista de la distribución en el muestreo, no se podía determinar el nivel de significación para todo valor de δ como afirmaba la teoría de Neyman-Pearson.

Finalmente, como muestra de su carácter, recogemos de W.G. Cochran una demostración hecha por Fisher:

"En una de sus clases citó sin demostrar un resultado. Tras de varios intentos sin que me saliera le pedí, en su despacho, si podía hacerme la demostración. Me dijo que en algún sitio la tenía archivada; abrió varios cajones y decidió que era mejor obtenerla de nuevo. Nos sentamos y escribió la misma expresión de la que yo había partido. El camino obvio va en esta dirección dijo, y escribió una expresión de dos líneas. Ahora supongo que hay que desarrollar esto, y puso una ecuación que ocupaba tres líneas. Miró la expresión y comentó: el único camino parece ser éste, y obtuvo una expresión de 4 líneas y media. Hubo un silencio de unos 45 segundos y dijo, el resultado se debe seguir de esto, y escribió debajo la expresión que yo le había preguntado. La clase había terminado".

Gracias a Fisher el estudio de la estadística se introdujo en las Universidades. No olvidemos que salvo en el *University College London (U.C.L.)*, donde ya había cursos de estadística gracias a la personalidad e iniciativa de Karl Pearson, la estadística como tal no aparecía en los programas de las diversas disciplinas científicas.

Curiosamente, como ya hemos mencionado, nunca llegó a ser catedrático de Estadística sino de Genética en el U.C.L. y en Cambridge.

2. Contribuciones a la inferencia estadística

En opinión de L.J. Savage (1976,1981) lleva menos tiempo decir a qué partes de la Estadística Fisher no contribuyó, que a las que sí lo hizo. No obstante vamos a resumir las principales aportaciones de Fisher a los métodos estadísticos y a los fundamentos de la inferencia estadística.

Para ello partimos del estado en que se encontraba la Estadística que podríamos denominar prefisheriana, dominada por la figura de K. Pearson, que resumimos en la Tabla 1.

LA ESTADÍSTICA ANTES DE FISHER

1.- **Estadística descriptiva**

- 1.1- Histogramas, diagrama de barras, diagramas de nube de puntos, etc.
- 1.2- Cálculo de medidas de centralización, dispersión y asociación.
- 1.3- No se distinguía entre población y muestra.

2.- **Ajuste de curvas**

- 2.1- La ley normal, introducida por Gauss y Laplace para ajustar toda clase de datos.
- 2.2- Sistema de curvas de frecuencia de Pearson, que tienen en cuenta la asimetría y la curtosis.
- 2.3- Test de la χ^2 de Pearson. Uno de los 20 descubrimientos más importantes del siglo XX en los campos de la ciencia y la tecnología.

3.- **Contrastes de hipótesis**

- 3.1- Prácticamente inexistente hasta la introducción del test de la *t* de Student en 1908.

4.- **Estimación**

- 4.1.- método de los cuadrados mínimos de Gauss y Legendre
- 4.2.- método de las desviaciones absolutas mínimas de Laplace
- 4.2.- método de los momentos de K. Pearson

5.- **Inferencia inductiva (teorema de Bayes)**

Tabla 1. Resumen de la situación de la Estadística antes de Fisher

Hemos dividido esta sección en cuatro partes o subsecciones: la primera corresponde a los libros sobre estadística; la segunda recoge prácticamente todas las aportaciones de Fisher a la metodología estadística; la tercera se dedica a los fundamentos de la inferencia estadística y, finalmente, la cuarta a los temas más controvertidos de la misma.

2.1. Libros sobre estadística

Statistical Methods for Research Workers fue escrito entre los veranos de 1923 y 1924 y fue reeditado en catorce ocasiones. Esta última edición contiene algunos añadidos tomados de unas notas dejadas por el autor con este fin, justo antes de su fallecimiento. Contiene: una introducción, estudio de diagramas, distribuciones, tests de ajuste, independencia y homogeneidad, tests de significación de medias, diferencia de medias y coeficientes de regresión, el coeficiente de correlación, correlaciones intraclásicas y el análisis de la varianza, aplicaciones del análisis de la varianza y los principios de la estimación estadística. Como puede verse, no desmerece en nada de lo que puede ser un curso actual de inferencia.

Cada sección empieza con un conjunto de datos con los cuales él se había encontrado en alguna investigación y lleva al lector a través de las etapas del análisis estadístico hasta que lo introduce en los conceptos necesarios para su solución. Es notable señalar que es el único libro de texto avanzado, que enfatiza la importancia de las técnicas gráficas en el análisis de datos. De hecho Fisher dice:

> "El examen preliminar de la mayor parte de los datos se ve facilitado por el uso de diagramas. Los diagramas no prueban nada, pero hacen que nos demos cuenta de los tests que debemos aplicar a los datos, sin sustituir a estos, y nos permiten explicar las conclusiones de los tests basados en los diagramas".

De hecho hasta la aparición, inmediatamente después de la segunda guerra mundial, del influyente libro de H. Cramér *Métodos Matemáticos de la Estadística*, era el único libro de estadística superior, aunque iba dirigido a un público mucho más amplio que el de los matemáticos. El éxito del mismo, sobre todo entre aquellos científicos que no tenían una sólida base matemática, es que su metodología les permitía resolver de inmediato muchos de los problemas estadísticos que tenían.

No obstante, Kendall opinaba de esta obra que "no es un libro fácil", y que alguien dijo una vez que "ningún estudiante debería intentar leerlo a menos que ya lo hubiera leído antes".

Sus trabajos sobre genética están contenidos en su libro *The Genetical Theory of Natural Selection* (1930) donde explica cómo se produce el ocaso de las civilizaciones con un sofisticado tratamiento matemático y demuestra la modificación de la dominancia a la hora de explicar la herencia, que hasta entonces se creía que era un concepto estático.

Su otro grupo de contribuciones importantes, que obtuvo a partir de su trabajo en Rothamsted, está contenido en su publicación *The Design of Experiments* (1935). Uno de los conceptos más importantes en el libro fue el concepto de aleatorización, a la hora de asignar tratamientos a los bloques. Hasta entonces esta asignación se realizaba sistemáticamente, con lo que se podían obtener conclusiones erróneas. Él propuso aleatorizar esta asignación ya que así además podía utilizar la distribución normal para los contrastes. Este procedimiento, uno de los más controvertidos de la estadística, le llevó a nuevas polémicas, y estuvo a punto de terminar con su amistad con Gosset.

El libro introduce, por vez primera, una serie de conceptos que hoy son de uso común en esta disciplina. Una lista de los términos más importantes aparece en la subsección 22.4.

También el libro convirtió el tomar el té en algo histórico. Fisher tenía la costumbre, desde sus tiempos de Rothamsted de tomar el té con todos los miembros del laboratorio de Estadística. Un día al dar la taza a la doctora B. Muriel Bristol ésta declinó diciendo que prefería que la leche se vertiera primero. Fisher contestó diciendo que aquello era irrelevante. William Roach, otro miembro del laboratorio, quien después se casó con ella, propuso realizar el experimento y ella acertó en un porcentaje muy alto de los casos. Posteriormente en 1935 incluyó en su libro, al principio del capítulo II, este problema, aprovechándolo para plantear una serie de preguntas relevantes sobre el experimento como ¿cuántas tazas deberían servirse?, ¿en qué orden?, ¿cuántas se deberían acertar?... y lo utiliza como guía de acción para enfrentarse a otros experimentos.

Su último libro se titula *Statistical Methods and Scientific Inference*, publicado por vez primera en 1956 (3ª edición 1973), tiene un nivel matemático mayor que los otros dos, especialmente los capítulos V y VI, y ahonda más en los aspectos filosóficos y matemáticos de la inferencia estadística. El segundo capítulo -el primero es el prólogo- resume brevemente parte de la historia de la estadística, comenzando con Thomas Bayes hasta Galton y K. Pearson, y finaliza con una breve digresión sobre el significado de la probabilidad. El libro repasa aquellos aspectos más controvertidos de la inferencia como son los contrastes de significación, en particular los referentes a las tablas de contingencia 2×2 y el que hoy denominamos problema de Behrens-Fisher, tests o contrastes de aleatorización, que incluso hoy son objeto de controversia. El argumento fiducial, tan caro a Fisher, en contraposición al método *clásico [sic]* de Bayes, los problemas de predicción, el estudio de la función de verosimilitud y, por último, los principios de la estimación.

Los tres libros de Fisher sobre estadística se han editado recientemente (1990) en un solo volumen por J.H. Bennet, precedidos por un sustancioso prólogo escrito por un estadístico que estuvo muy vinculado a Fisher y, de hecho, contribuyó en gran medida al desarrollo del Diseño de Experimentos, junto con su creador F. Yates.

Ya hemos comentado que a pesar de que Fisher rechazaba el argumento bayesiano en la inferencia estadística, Fisher fue uno de los promotores de la reedición del clásico artículo de Bayes en 1958.

2.2. Contribuciones a los Métodos Estadísticos

Basta examinar simplemente el contenido de esta subsección para darse cuenta de que realmente no hubo ninguna parcela de los métodos estadísticos -de los que muchos fue el creador indiscutible, como ocurre con el Diseño de Experimentos-, a los que Fisher no realizase contribución alguna. La lista que sigue, sin ser de todo exhaustiva, recoge la mayoría de los temas sobre los que Fisher trabajó, acompañados de una breve descripción o comentario sobre los trabajos y el año de su publicación.

1.- Distribuciones en el muestreo

1.1- t de Student (conjeturada por Gosset), F de Snedecor y z de Fisher.

1.2- Distribuciones no centrales. t de Student, F de Snedecor y z de Fisher. (Sorprende el que Fisher dedujera estas distribuciones ya que su génesis se debe al estudio de la *función de potencia* en la teoría de los contrastes de hipótesis como la consideran Neyman y Pearson, concepto que Fisher rechazaba).
1.3- Coeficiente de correlación muestral para variables correladas y coeficiente de regresión múltiple.

2.- Modelos lineales (1922,1938)

2.1.- Métodos numéricos de cálculo de los estimadores de los coeficientes de regresión. (Eliminación gaussiana, etc.).
2.2.- Test de la bondad de ajuste de una función de regresión y contraste de significación de los coeficientes individuales.
2.3.- Selección de variables en regresión.
2.4.- Estudio de los residuos.

3.- Análisis multivariante

3.1.- Función discrimante.
3.2.- Distribución de los valores propios de las componentes principales.

4.- Diseño de Experimentos (1926, ..., 1966)

Fisher fue el creador indiscutible de lo que hoy llamamos **Diseño de Experimentos**, en el:

- sentido amplio de introducir consideraciones estadísticas al planificar experimentos.

> "... Un examen cuidadoso del proceso de recogida de datos, o diseño experimental, puede incrementar la precisión de los resultados en diez o doce veces, empleando el mismo tiempo y esfuerzo. Consultar a un estadístico después de que se haya concluído un experimento es, muy a menudo, pedirle que realice un examen post-mortem. Quizás le pueda decir de qué murió el experimento".

- sentido más restringido de explotar los patrones combinatorios en el diseño de experimentos.

Fisher luchó denodadamente contra la práctica o máxima, hasta entonces utilizada, de variar un solo factor en cada ocasión.

Una idea de la aportación de Fisher a esta nueva disciplina es la enorme cantidad de nuevos términos por él acuñados; a continuación, enumeramos los más importantes.

Conceptos estadísticos introducidos por Fisher en el Diseño de Experimentos:

- Aleatorización
- Diseños aleatorizados
- Diseños factoriales
- Diseños combinatorios
- Replicación parcial
- Confundido
- Bloques incompletos
- Control estadístico local
- Análisis de la Varianza
- Análisis de la Covarianza

Estas ideas tuvieron mucha importancia en el desarrollo de la **Matemática combinatoria** (geometrías finitas, cuadrados latinos ortogonales, demostración de la falsedad de la conjetura de Euler sobre la no existencia de cuadrados latinos ortogonales de orden $4n+2$, teoría de la codificación, etc.).

Los apartados 5 y 6 recogen aquellos trabajos que originaron la controversia con Karl Pearson de la que Fisher salió completamente victorioso, como comentamos en las notas biográficas.

5.- Uso correcto de la distribución χ^2

(En tablas de contingencia y en el test de la bondad del ajuste).

6.- Ineficiencia del método de los momentos como método de estimación puntual.

7.- Versión simplificada de la cota de Fréchet-Cramér-Rao.

(Véanse, en 2.3, los comentarios sobre los conceptos de Consistencia y Cantidad de Información).

8.- Familias exponenciales

Fisher, al parecer según la opinión de L.J. Savage, se adelantó a la idea de las familias de Pitman-Koopman-Darmois.

9.- Métodos no paramétricos

9.1.- Introdujo el test de los signos como alternativa al test de la t.
9.2.- Introdujo los denominados tests exactos, como sustitutos de los tests de la t y de la F, cuando la hipótesis de normalidad no es válida.
9.3.- Estudió las rachas (1926) y estadísticos de orden como límites fiduciales de los percentiles poblacionales (1939,1945).
9.4.- Tabuló los valores esperados de los estadísticos de orden de una población normal y la suma de sus cuadrados para aplicarlos al problema de las dos muestras y al análisis de la varianza basada en rangos.

10.- Análisis secuencial (1922,1941,1952)

11.- Teoría de Juegos

Desarrolló, de modo independiente, las ideas de estrategia minimáx y estrategia aleatorizada aplicadas a la resolución del juego de le Her (1934).

12.- Estadística Espacial

Inició el estudio (1953) del análisis de datos direccionales aplicado a determinar, por el método de la máxima verosimilitud, la dirección del magnetismo residual en las emisiones de lava volcánica.

13.- Valores extremos

En 1928, Fisher y Tippett, iniciaron el estudio de la distribución asintótica de los valores extremos, continuado por Gumbel (1934).

14.- Procesos de ramificación (1930)

Aunque se considera a Watson y Galton (1874) como los que iniciaron el estudio de estos procesos, fue Fisher el que avanzó en su estudio al introducir como herramienta las funciones generatrices.

15.- Los *k*-estadísticos (estimadores insesgados de los cumulantes) (1929, 1930,1931 y 1937)

Llama la atención el que Fisher dedicara parte de sus esfuerzos al estudio de los *k*-estadísticos, que son simplemente estimadores insesgados de los cumulantes, puesto que el concepto de estimador insesgado era irrelevante para Fisher.

16.- Análisis del modelo probit (1935)

Aplicó el método de la máxima verosimilitud, junto con Bliss, a la estimación de los parámetros de este modelo.

2.3. Fundamentos de la Inferencia Estadística

La aparición en 1922 del artículo de Fisher *On the Mathematical Foundation of Theoretical Statistics* marca una nueva era (que podríamos llamar postfisheriana) de la estadística, ya que en este artículo se introduce, por vez primera, una larga serie de términos estadísticos que, desde entonces forman parte de la literatura estadística, y pone de manifiesto la importancia de la noción de modelo estadístico y establece la distinción, por primera vez en la literatura, entre **Población** y **Muestra aleatoria**, que anteriormente estaban difuminadas y, a veces, entremezcladas. Pasamos, como hicimos en la subsección anterior, a enumerar y comentar brevemente toda la terminología que Fisher introdujo en ese artículo y en trabajos sucesivos.

Enunciado de los problemas fundamentales de la Estadística

- **i) Especificación** (modelo poblacional representado por una familia de distribuciones de probabilidad P_θ).
- **ii) Estimación** (elección del valor del parámetro θ más apropiado, basado en la muestra).
- **iii) Distribución en el muestreo** (que permite calcular la precisión del estimador de θ o medir la incertidumbre en la elección del parámetro).

Conceptos estadísticos introducidos por Fisher:

- Parámetro
- Estadístico y estimador
- Consistencia
- Eficiencia
- Suficiencia
- Cantidad de información
- Estadístico ancilario o subsidiario
- Función de verosimilitud
- Máxima verosimilitud
- Probabilidad e inferencia fiducial o fiduciaria

Estadístico-Tal como la conocemos hoy. No necesariamente tiene la función de estimar un parámetro.

Fisher (1922)

Estimador-Un estadístico cuya función sea estimar algún parámetro. No dio una definición precisa de él.

Fisher (1922)

No le gustaba el término **estimador puntual** ya que lo asociaba con una estimación sin precisar su acuracidad o precisión.

Fisher ve el problema de la estimación como el de reducción de los datos. Así, si $x = (x_1, \ldots, x_n)$ es una muestra, la idea de estimador es encontrar un vector $\mathbf{T} = (T_1(x), \ldots, T_k(x))$, con $k << n$, tal que **T** *contenga la mayor cantidad, idealmente toda, de información relevante.*

¿Cuáles son los criterios apropiados para elegir un estadístico T que sustituya a toda la muestra? Fisher introdujo los tres criterios siguientes:

Suficiencia-Un estadístico es suficiente si contiene toda la información relevante, es decir, toda la que contiene la muestra.

La función de verosimilitud del estadístico es igual a la de todas las posibles muestras a partir de las cuales el estadístico se ha obtenido.

(Véase Función de verosimilitud)

Fisher (1920,1922,1925)

En las dos últimas publicaciones aparece el conocido lema de factorización.

(Véase Función de verosimilitud).

Consistencia (en el sentido de Fisher)-Un estadístico satisface el criterio de consistencia si, cuando se calcula a partir de toda la población, es igual al parámetro requerido.

Fisher (1921), SM (1928) 2ª edición y siguientes y SI (1956)

Esta definición es perfectamente válida, y así aparece recogida en los textos sobre muestreo en poblaciones finitas, cuando la población es finita.

En alguna ocasión se acerca a la definición habitual de convergencia en probabilidad hacia el verdadero valor del parámetro. Pero posteriormente rechazó esta definición. Quizás lo que Fisher tenía en mente, según Rao (1922), es, por una parte, la convergencia en probabilidad y por otra, que el estimador (considerado como función de las frecuencias observadas), tome el verdadero valor cuando las frecuencias relativas se sustituyen por las frecuencias esperadas.

Parece ser que Fisher siempre tenía presente el caso de tomar muestras de una multinomial y que, además, consideraba cualquier otra distribución como una multinomial infinita.

En la terminología actual, y refiriéndonos al caso multinomial, la definición de consistencia de Fisher se podría escribir:

Si $p_1,\dots,p_k$ son las frecuencias relativas en una muestra, de los sucesos $A_1,\dots,A_k$ de una multinomial cuyas probabilidades son $\pi_1(\theta),\dots,\pi_k(\theta)$ (funciones de un cierto parámetro θ), un estimador $T(p_1,\dots,p_k)$ es consistente si

$$T(\pi_1(\theta),\dots,\pi_k(\theta)) \equiv \theta$$

sin hacer referencia a ninguna propiedad límite. De esta definición Fisher deduce:

La desigualdad de información de Fisher: Si T es consistente, entonces

$$\mathrm{Var}(T) > 1/I$$

donde I es la cantidad de información esperada o información de Fisher. De modo que si T es consistente en el sentido de Fisher la distribución asintótica de $\sqrt{n}\left(T(p_1,\ldots,p_k)-\theta\right)$ es $N\left(0,\operatorname{Var}(T)\right)$.

Un defecto de la definición de Fisher es que es demasiado restrictiva y no se puede aplicar, p.e., a los procesos estocásticos.

Eficiencia (en el sentido de Fisher). Un estadístico satisface el criterio de eficiencia si su distribución en el muestreo tiende hacia una distribución normal, cuando el tamaño muestral aumenta, y tiene la varianza más pequeña posible.

Fisher (1921), SM (1925) 1ª edición y siguientes y SI (1956)

Fisher lo utiliza para elegir, dentro de la clase de los estimadores consistentes, el que tenga mayor precisión o acuracidad.

Cantidad de información (esperada)

Fisher la define, para un modelo estadístico dado por $f(x\,|\,\theta)$, como es habitual:

$$I = E\left[\left(\frac{\partial \log f(x\,|\,\theta)}{\partial\theta}\right)^2\right] = E\left[-\frac{\partial^2 \log f(x\,|\,\theta)}{\partial^2\theta}\right]$$

Propiedades importantes:

i) Es aditiva para procesos independientes.
ii) Se reduce, si se agrupan datos o se sustituyen estos por un estadístico.
iii) No hay pérdida de información si, y solo si, un estadístico es suficiente.
iv) Relación con la teoría de la información (Shannon, Kullback and Leibler, etc.)
v) Interpretación para muestras grandes: La precisión asintótica de los estimadores máximo-verosímiles es la inversa de la varianza.
vi) Interpretación para muestras pequeñas. Los estimadores que no sean eficientes pierden una fracción de la información que proporcionan los datos.

Fisher daba mucha importancia a que las inferencias fueran condicionales a cierta información que ya se conoce y se puede suponer constante (como ocurre, de modo automático con la inferencia bayesiana), acuñando el concepto de estadístico ancilario o subsidiario.

Estadístico ancilario o subsidiario. Aquél cuya distribución no depende del parámetro sobre el que se quiere hacer inferencias ... *no suministran información sobre el problema (de inferencia) en cuestión ... y se pueden considerar como constantes.*

Si $(\mathcal{X},\mathcal{B},\Omega)$ es un modelo estadístico y $T(X)$, con $X\in\mathcal{X}$ es un estimador de θ, Fisher sugiere -en vez de lo que se hace habitualmente en la estadística frecuentista de calcular la distribución en el muestreo de $T(X)$ para hacer inferencias sobre θ - que la distribución en el muestreo de $T(X)$ debe calcularse no respecto del modelo estadístico original sino respecto del modelo restringido $(\mathcal{X}_A, \mathcal{B}_A, \Omega)$, donde

$$\mathcal{X}_A = \{X \in \mathcal{X} : A(X) = A(x)\},$$

x es el valor observado de $X, A(X)$ es un estadístico ancilario de θ y $A(x)$ es el conjunto de referencia.

Ejemplos:

i) El estudio de los coeficientes de regresión cuando se muestrea de una normal bivariante se realiza condicionado a los valores de una de las variables como si estos fueran prefijados y no aleatorios.

ii) (Rao (1971.) Supóngase que se quiere estimar la población del estado de Bengala Occidental mediante una muestra aleatoria del total de las N ciudades que lo constituyen, y no se tiene información de sus tamaños. La teoría nos dice que si $x_1 \ldots, x_k$ son los tamaños de la población de las k unidades muestrales, entonces la estimación de la población de Bengala Occidental es $N\bar{x}$. Sin embargo, si en la muestra observada hay un valor, p. ej. x_3, (quizás la población de Calcuta), mucho mayor que el resto, entonces un estimador mejor sería $x_3 + (N-1)\bar{x}'$, donde $\bar{x}'$ es el promedio calculado excluyendo a x_3. Aquí, el conjunto de referencia es el de las muestras que contienen a x_3.

A pesar de la importancia de este concepto, hay problemas con la unicidad de los estadísticos ancilarios que, además, pueden dar origen a inferencias diferentes.

Función de verosimilitud (1922). La verosimilitud asociada a una muestra es la probabilidad o densidad de la observación como función del parámetro módulo una constante multiplicativa, es decir, es la clase de funciones del parámetro proporcional a la probabilidad de la observación dado el parámetro

$$L(\theta; x_1, \ldots, x_n) = \prod_{x=1}^{n} f(x_i; \theta).$$

Como puede observarse la anterior es básicamente la definición actual de función de verosimilitud. Fisher fue, desde el primer momento, consciente de su importancia.

La verosimilitud representa el estado de nuestra información con respecto a los parámetros de las poblaciones hipotéticas (1922).

La verosimilitud es una medida del orden de preferencia entre las diferentes poblaciones posibles (1954).

La verosimilitud es una medida de creencia racional distinta de la probabilidad y que no obedece las reglas del cálculo de probabilidades (1956).

Basado en esta última consideración A.W.F. Edwards en 1972 publicó su libro *Likelihood* en el que desarrolla una teoría de la inferencia estadística basada exclusivamente en el concepto de verosimilitud.

Fisher demostró sus propiedades más importantes, que enumeramos a continuación:

Propiedades:

i) Es multiplicativa para observaciones independientes y, por consiguiente, el logaritmo de la función de verosimilitud es una función aditiva.
ii) Es un estadístico minimal suficiente (1925).

Máxima verosimilitud (1912,1922). El método de estimación de la máxima verosimilitud, Fisher (1930) lo atribuye a Gauss, aunque hay precedentes en los trabajos de Lambert (1760) y D. Bernoulli (1788) y en Edgeworth (1908-9). Sin embargo Fisher fue mucho más lejos que sus predecesores en promocionar su uso como un método universal de estimación y estudiar sus propiedades.

Hoy se sabe que muchas de las proposiciones que Fisher *demostró* sobre las propiedades de los estimadores de máxima verosimilitud no son universalmente ciertas. Sin embargo, muchos de estos resultados se pueden demostrar rigurosamente bajo ciertas condiciones. Fisher fue, sin embargo, consciente de los defectos de sus demostraciones:

> ... Por mi parte, gustósamente habría retrasado su publicación hasta que hubiera formulado una demostración completamente rigurosa; pero la cantidad y la variedad de resultados nuevos que este método revela me empujó a publicarlo.
>
> Fisher (1922)

Propiedades de los estimadores máximo verosímiles

Bajo ciertas condiciones de regularidad se sabe que los estimadores máximo verosímiles:

i) Son consistentes. En el sentido de Fisher, para los experimentos multinomiales y, en general, en el sentido habitual de convergencia en probabilidad hacia θ.
ii) Son eficientes, es decir, su distribución asintótica es normal centrada en θ y de varianza asintótica la inversa de la cantidad de información de Fisher.
iii) Son eficientes de segundo orden, desde el punto de vista de la teoría de la decisión, ya que minimizan el desarrollo en serie de la función de riesgo (hasta términos del orden $1/n^2$) cuando la función de pérdida tiene una cierta forma.

iv) Si el verdadero modelo estadístico Q del que se muestrea no pertenece a la familia paramétrica especificada P_θ y $\tilde{\theta}$ es el estimador de máxima verosimilitud de θ, entonces $P_{\tilde{\theta}}$ es la mejor aproximación, dentro de la familia especificada, al verdadero modelo Q, en el sentido de que minimiza la divergencia de Kullback-Leibler.

2.4. Temas controvertidos

Fisher, dado su carácter polémico y controvertido fue el centro de muchos debates, que incluso hoy persisten. Basta citar el ingente número de artículos que sobre el denominado *test exacto de Fisher* en tablas de contingencia existe y sobre el que hoy se sigue publicando, sin que los diversos autores se pongan de acuerdo con la metodología.

Fisher nunca estuvo demasiado interesado por los problemas de los fundamentos de la inferencia inductiva, probablemente por su carácter y por su inquietud por resolver problemas prácticos. En particular nunca mantuvo disputas sobre el concepto de probabilidad sino sobre la metodología estadística. Su concepto de probabilidad fiducial es oscuro, véase Girón (1989), y está unido a las ideas de inferencia y distribución fiduciaria, que han tenido su continuación en la denominada *Inferencia Estructural* de Fraser y en la consideración del principio de invariancia en inferencia y en la *Teoría de la Decisión.*

Aunque a Fisher se le considera un frecuentista, esta afirmación, tal como se deduce del análisis de sus trabajos por L.J. Savage, no está clara. Aunque negaba la utilización de la probabilidad inversa de Bayes como método inferencial cuando no se tenía información a priori, era partidario del enfoque bayesiano cuando se disponía de esa información, caso contrario se mostró partidario de su enfoque fiducial, a pesar de los problemas que conlleva.

Esos dos temas son los que comentamos brevemente en esta subsección.

Probabilidad en inferencia fiduciarias (1922)

La idea de Fisher parece ser la de llegar a realizar inferencias análogas a las de la inferencia bayesiana, mediante algún prodeso que él denomina el argumento fiducial, pero sin utilizar la distribución a priori.

Distribuciones fiduciarias obtenidas por Fisher

i) Para muestras de una normal, la individual de la media, la desviación típica y su conjunta.
ii) Para muestras de una normal, la de una combinación lineal de la media y la desviación típica y la media y la desviación típica de una observación futura.
iii) Para muestras de una normal bivariante, la de todos los parámetros, la del coeficiente de correlación y la del cociente de las medias.

iv) Para el problema de las dos muestras de poblaciones normales, la que hoy conocemos por distribución de Behrens-Fisher.

Kolmogoroff, discutiendo el argumento fiduciario, escribe:

> En este caso deberíamos señalar que hace falta un nuevo axioma de la teoría de la probabilidad según sigue: Si la probabilidad condicional, $P(A|\theta_1,\ldots,\theta_s)$, de un suceso A es igual a ω para todos los valores posibles de los parámetros $\theta_1,\ldots,\theta_s$, entonces la probabilidad no condicional de A existe y es igual a ω.
>
> Deberíamos señalar que tal axioma es innecesario cuando el problema se trata desde el punto de vista clásico ya que, en este caso, se supone una distribución a priori sobre los parámetros y el resultado se sigue de las reglas de la teoría de la probabilidad.
>
> Kolmogoroff (1942)

> Recientemente en París... me enteré de que un eminente ruso, llamado Kolmogoroff, ha postulado un axioma, que pretende justificar axiomáticamente cierto tipo de argumentos fiduciarios de los que soy responsable... De modo que cuando me encuentro en dificultades puedo asegurar con confianza ¡por el axioma de Kolmogoroff!.
>
> Kolmogoroff (1942)

Contraste de hipótesis (1922)

Fisher ve los contrastes de la manera siguiente: Dada una hipótesis nula H_0, se elige un estadístico de contraste y se calcula su percentil al 95% y al 99%. Si el valor observado del estadístico excede este valor, se rechaza H_0. La decisión se basa en la disyunción lógica: *O ha ocurrido un suceso intrínsecamente improbable, o la hipótesis nula no es correcta (1926). El nivel de significación en tal caso satisface las condiciones de una medida del sustrato racional de la incredulidad que engendra (1956).*

- Diferencias sustanciales con la teoría de Neyman-Pearson.
- No considera hipótesis alternativas. El test, caso de rechazarse la hipótesis nula, no indica cual pueda ser la alternativa.
- No considera el error de tipo II, por consiguiente, la función de potencia.
- No recomienda un nivel de significación fijo.
- Importancia de lo que los tests sean exactos (a veces condicionando los estadísticos de contraste a uno ancilario y, en otras ocasiones, que se pueda determinar el tamaño del test).
- Fisher parte de la idea de que se parta de un cierto nivel de significación ya que el cálculo del p-valor asociado al estadístico de contraste era en esa época difícil- sobre todo en los tests exactos- por la falta de medios eficientes de computación.

Agradecimientos

Este trabajo se ha subvencionado, en parte, con la ayuda de la *Dirección General de Investigación Científica y Técnica* (DGICYT) correspondiente a los proyectos números TD91-0411 and PB93-1154 y la *Consejería de Educación de la Junta de Andalucía.*

Referencias

[1] BOX, J.F. (1978) *R.A. Fisher, The Life of a Scientist.* New York. Wiley.

[2] EDWARDS, A.W.F. (1972). *Likelihood.* Cambridge University Press: Cambridge.

[3] FISHER, R.A. (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edimb.*, **52**, 399-433.

[4] FISHER, R.A. (1921) On the "probable error" of a coefficient of correlation deduced from a small sample. *Metron*, **1**, 3-32.

[5] FISHER, R.A. (1921) Studies in crop variation. I. An examination of the yield of dressed grain from Broadbalk. *J. Agric. Sci.*, **11**, 107-135.

[6] FISHER, R.A. (1922) On the mathematical foundations of theoretical statistics. *Phil. Trans.*, A, **222**, 309-368.

[7] FISHER, R.A. (1922) On the interpretation of χ^2 from contingency, and the calculation of P.*J. Roy. Stat. Soc.*, **85**, 87-94.

[8] FISHER, R.A. (1925,1970) *Statistical Methods for Research Workers.* Edimburgo. Oliver and Boyd (Hay edición en español).

[9] FISHER, R.A. (1930) *The Genetical Theory of Natural Selection.* Oxford. University Press y (1958) New York. Dover.

[10] FISHER, R.A. (1930) Inverse Probability. *Proc. Camb. Phil. Soc.*, **26**, 528-535.

[11] FISHER, R.a. (1935) The fiducial argument in statistical inference. *Am. Eugen.*, **6**, 391-398.

[12] FISHER, R.A. (1935,...,1966) *The Design of Experiments.* Edimburgo. Oliver and Boyd (Hay edición en español).

[13] FISHER, R.A. (1956,1959) *Statistical Methods and Scientific Inference.* Edimburgo. Oliver and Boyd. (1973) New York. Hafner.

[14] FISHER, R.A. (1990) *Statistical Methods, Experimental Design and Scientific Inference.* Edited by Bennett, J.M. with a foreword by Yates, F. Oxford. Oxford University Press.

[15] GIRÓN, F.J. (1989). El concepto de probabilidad. En *Historia de la Ciencia Estadística.* Ed. Real Acad. Cien. Exac. Fís. Nat., pp. 15-44.

[16] RAO, C.R. (1922). R.A. Fisher: The founder of modern statistics. *Statist. Sci.*, **7**, 34-48.

[17] SAVAGE, L.J. (1970) On rereading R.A. Fisher. Fisher Memorial Lecture published posthumously. J.W. Pratt, Ed. *Ann. Math. Stat.*, (1976), 4, 441-500.