

## DISEÑOS EXPERIMENTALES Y ENSAYOS CLÍNICOS

Pedro Saavedra

Usando la terminología de la teoría general de sistemas de Bertalanffy, los sistemas pueden clasificarse en sistemas cerrados y sistemas abiertos. Un sistema es un conjunto de elementos de la naturaleza que interactúan entre sí. Los sistemas que aparecen en la física clásica son generalmente sistemas cerrados. Esto significa que el investigador o manipulador del sistema puede conocer y/o controlar todos los elementos que lo constituyen. De esta forma, para una cierta entrada, puede predecir casi sin error la respuesta del sistema. Su información acerca de éste a menudo puede expresarla a través de sistemas de ecuaciones o sistemas de ecuaciones diferenciales. Un sistema muy elemental es un gas encerrado en un recipiente deformable. El volumen, la temperatura y la presión del mismo son elementos que interactúan entre sí. Para un valor de volumen y temperatura, podemos predecir la presión utilizando el conocido modelo dado por la ecuación de los gases perfectos.

En el campo de las biociencias, sin embargo, raramente podemos encontrar sistemas de esta naturaleza. Podemos intentar pensar en factores que influyen en el nivel de colesterol total de un individuo (alimentación a lo largo de la vida, antecedentes familiares, genotipos, sexo, edad, patologías que el individuo ha padecido, etc.). Por muy larga que hagamos la relación, sabemos con seguridad que cualquier predicción que hagamos de su nivel de colesterol a partir de tales factores va a estar sujeta a considerables errores. Esto se debe a que los seres vivos en general son extraordinariamente complejos, lo que supone que cualquier variable que podamos medir en ellos está afectada por un número indefinido de factores. En la práctica, la proporción de factores que el investigador puede medir o controlar es reducida. Tales sistemas reciben el nombre de sistemas abiertos. En algunos casos, podemos describir este tipo de sistemas en los siguientes términos; a saber: si  $Y$  es la variable que se quiere predecir,  $X_1, \dots, X_p$  son los factores controlables y  $Z_1, \dots, Z_q$  los no controlables, entonces:

$$(1) \quad Y = f(X_1, \dots, X_p) + g(Z_1, \dots, Z_q)$$

Ahora bien, dado el desconocimiento del segundo término, el modelo puede representarse como:

$$(2) \quad Y = f(X_1, \dots, X_p) + \mathcal{E}$$

A la variable  $\mathcal{E}$  que sustituye a la función  $g$  se le denomina frecuentemente error aleatorio. Si (2) se corresponde con un modelo para un sistema cerrado, el error aleatorio en líneas generales representa un error debido a los aparatos de medición. Sin embargo, aunque para sistemas abiertos  $\mathcal{E}$  puede recoger tales errores, su significado fundamental se refiere a los factores incontrola-

dos. Los métodos estadísticos de exploración de datos pueden por otra parte darnos una aproximación a la forma de la función  $f$ . El investigador puede tener ideas a priori sobre su forma, pero sin duda, los datos son esenciales, no sólo para realizar el ajuste si se elige un modelo *ad-hoc*, sino para explorar la propia forma del modelo.

Queda pues establecida la importancia de los datos en la construcción de modelos. Estos se obtienen en general a partir de la realización de estudios observacionales o experimentales. En los primeros, el investigador se limita a observar diversas variables sobre una muestra de objetos de la población correspondiente. En los estudios experimentales sin embargo, el investigador interviene sobre los objetos, asignándoles a tratamientos específicos.



Roland Fisher

R. A. Fisher fue el impulsor del uso de los métodos estadísticos en el diseño de experimentos. Durante algunos años estuvo a cargo de la estadística y del análisis de datos en la estación agrícola experimental Rothamsted en Londres. Fisher fue quien desarrolló y usó por primera vez el análisis de la varianza como herramienta para el análisis estadístico en el diseño experimental. En 1933 trabajó como profesor en la Universidad de Londres. Después formó parte del cuerpo docente de la Universidad de Cambridge, y fue profesor visitante en varias universidades del mundo. Las aplicaciones que realizó de los diseños experimentales y el análisis de datos fueron principalmente en el campo de la

agricultura. ¿Qué tipo de problemas trataba con esta metodología? Un ejemplo elemental sería el siguiente: comparar el efecto sobre el rendimiento en la producción de cereales de tres clases de semillas y cuatro marcas de fertilizantes. Cada combinación de semilla y fertilizante es un *tratamiento*, por lo cual hay doce tratamientos posibles. Estos tratamientos hay que aplicarlos a unidades experimentales que en este caso serían parcelas de terreno. En este caso, dado que hay doce tratamientos posibles, se requerirían al menos doce parcelas. Ahora bien, el rendimiento en la producción (variable respuesta), puede estar afectado por características inherentes a la parcela, a menudo no controlables. Por tal motivo, sería necesario que cada tratamiento se repitiera en diferentes parcelas. La repetición de los tratamientos en diferentes unidades muestrales sería además ineludible si se quiere evaluar adecuadamente la interacción entre semilla y fertilizante. ¿Cuántas réplicas de cada tratamiento deben realizarse? ¿Cómo asignar los tratamientos a las parcelas? ¿Cómo influyen los factores controlados (clase de semilla y marca de fertilizante) sobre el rendimiento en la producción? Esta última cuestión se corresponde con esta otra: ¿cómo es el modelo de la ecuación (2)?

Los problemas que estudió Fisher en el campo de la agricultura son similares a los que surgen en otros sistemas abiertos, particularmente en el campo de los ensayos clínicos. Así por ejemplo, nos puede interesar analizar la efectividad de un nuevo tipo de insulina (insulina experimental) frente a una insulina estándar para el tratamiento de la diabetes en orden a determinar cuál produce mejor efecto sobre la hemoglobina glucosilada. Dentro de los posibles diseños para este estudio, podemos considerar el diseño de grupos paralelos, consistente en la formación de dos grupos experimentales. Todos los sujetos de uno de los grupos reciben la insulina experimental y el resto (grupo control) la insulina estándar. ¿Qué criterio utilizamos entonces para la asignación de los tratamientos? Lo que aquí busca el investigador esencialmente es determinar si hay diferencia en la respuesta debida al tratamiento y de qué magnitud es tal diferencia. ¿Cómo asegurarnos en que la diferencia no se deba a factores incontrolados? Para ello, es indispensable que los grupos sean homogéneos, entendiéndose por ello que tanto los factores que puedan controlarse como los que no se distribuyan de forma análoga en ambos grupos. El método que mejor garantiza esta homogeneidad es la aleatorización, esto es, asignar los tratamientos al azar. Si para grandes tamaños muestrales observamos el comportamiento de factores medibles en ambos grupos, observaremos que difícilmente se producen diferencias significativas de éstos entre los grupos. Cabe entonces pensar que este comportamiento ocurre también con los factores no controlados. Los ensayos clínicos presentan además otras sutilezas. Por ejemplo, una vez formados los grupos por la aleatorización, puede producirse un mayor efecto placebo en el grupo experimental que en el control. Por efecto placebo entendemos una sugestión en la persona que le hace sentirse en mejor estado por estar, por ejemplo, recibiendo un nuevo tratamiento que piensa que es muy eficaz, sin que esto sea realmente así (por ejemplo, porque no hay realmente efecto farmacológico). El efecto placebo rompería entonces en última instancia la homogeneidad necesaria. Esto puede evitarse enmascarando el estudio; esto es, ocultando a los pacientes el tipo de tratamiento que están recibiendo. Este método (simple y doble ciego) es particularmente interesante en aquellos ensayos en los que la naturaleza de la variable respuesta tenga alguna dosis de subjetividad.

Hay sin embargo un aspecto esencial que diferencia a los ensayos clínicos de otros tipos de diseños experimentales, que es la naturaleza de la recolección de los datos. En los estudios agrícolas citados anteriormente, los tratamientos se asignan simultáneamente a todas las parcelas y también simultáneamente se evalúan las respuestas. Sin embargo, en los ensayos clínicos es muy difícil lograr esta simultaneidad y ocurre que, cuando aún están incorporándose sujetos al ensayo, ya hay datos evaluables de otros pacientes que fueron incluidos al principio del estudio. Hay que decir en este punto que la ética del ensayo clínico se fundamenta en el equilibrio entre el posible beneficio y riesgo del tratamiento experimental. Desde el momento en que cualquier información altere este equilibrio en cualquier sentido, el clínico está moralmente obligado a ofertar al paciente el tratamiento que se ha revelado más eficaz. Un análisis de los datos correspondientes a los primeros pacientes incluidos en el estudio podría aportar esta información. Aquí, por tanto, el tamaño de la

muestra está determinado por los resultados intermedios que van obteniéndose de modo secuencial. Estos ensayos en los que periódicamente, o cada cierto número de pacientes, se realizan análisis intermedios se denominan ensayos secuenciales.

El análisis estadístico de datos secuenciales, introducido por A. Wald en la segunda mitad de los años cuarenta, se basa esencialmente en el uso de adecuados test secuenciales que eviten la acumulación de los errores estadísticos que se producen al utilizar repetidamente los contrastes para muestras fijas. A partir de los años 70 surgen publicaciones sobre aplicaciones de los métodos secuenciales a los ensayos clínicos. Muchos de estos trabajos están recogidos en la excelente monografía de Whitehead (1997). Uno de los problemas de los análisis secuenciales radica en la invalidez de los métodos de análisis para tamaño muestral fijo. La estimación de una media a través de una media muestral no tiene las mismas propiedades estadísticas cuando el tamaño de la muestra es fijo que cuando el tamaño de la muestra es aleatorio. P. C. O'Brien señalaba hacia finales de los 70 que una de las razones del escaso uso de los ensayos secuenciales radica en la incertidumbre que muchas veces existe en la comunidad estadística acerca de la eficiencia de estos métodos. La finalidad de los trabajos que posteriormente han ido surgiendo ha sido precisamente resolver estos problemas matemáticos, para conocer del modo más rápido el mejor tratamiento de la enfermedad.

#### **Bibliografía**

29

Bertalanffy, L.: *General System Theory. Foundations, Development, Applications*. George Braziller, New York, 1968.

Fisher, R.A.: *The Design of Experiments*. Oliver and Boyd, 1935.

O'Brien, P.C. and Fleming, T.R.: A Multiple Testing for Clinical Trial. *Biometrics*, 35, 549-556, 1979.

Wald, A.: *Sequential Analysis*. Wiley, 1947.

Whitehead, J.: *The Design and Analysis of Sequential Clinical Trials*. Wiley, 1997.