1. ARTÍCULOS DE ESTADÍSTICA

¿POR QUÉ LA TEORÍA DE LA INFORMACIÓN?

Pedro Gil Álvarez
Departamento de Estadística e Investigación Operativa y Didáctica de la Matemática
Universidad de Oviedo

Fueron los ingenieros los que obligaron a los matemáticos a desarrollar la "Teoría de la Información". Las necesidades creadas para la transmisión de la información a distancia por cualquier medio (teléfono, fax, radio, televisión, ordenador) forzaron y fuerzan al desarrollo de nuevos modelos cada vez más poderosos y, cuando es posible, más sencillos de construir.

Es C.E. Shannon quien puede considerarse el "padre" de la teoría cuando, a mediados del pasado siglo, define la entropía, la cantidad de información y otros muchos conceptos que se aplicarían a la obtención de líneas telefónicas más fiables y con menos ruido. Las ideas son inicialmente sencillas: medir la incertidumbre antes y después de haberse realizado una experiencia aleatoria. La medida de la incertidumbre para un sistema probabilístico $\{p_1, p_2, \ldots, p_n\}$ la proporciona la conocida *entropía de Shannon*

$$H = -\sum p_i log p_i$$

Pero el verdadero motor de la teoría se encontró en las matemáticas: no sólo se formularon propuestas teóricas sino que se dieron las pautas para aplicarlas. Los matemáticos entran primeramente en una guerra de caracterizaciones axiomáticas de la entropía, hasta extraer la esencia de la misma, esencia que será considerada como inherente a cualquier medición de la incertidumbre.
Son entonces legión los investigadores que construyen nuevas medidas (Rényi, Oniçescu, Sharma, Mittal, Havrda, etc.) entre las cuales debería destacarse la idea genial de la "energía informativa" de Oniçescu (que funciona en sentido contrario a la entropía) y su posterior incorporación a la "entropía cuadrática" en el esquema general de Havrda & Charvat, por su simplicidad y operatividad

$$H^2 = 2\left(1 - \sum p_i^2\right)$$

Un paso más lleva a considerar la Información como entidad primaria, no necesitada de la probabilidad, y plantearse la cuestión clásica "¿cuál es el concepto previo: la probabilidad o la información?". Hace ya muchos años, el que luego sería creador de la moderna teoría de fractales, Benoit Mandelbrot, decía que la Teoría de la Información podía aspirar a ser la gran ciencia unificadora de todas las demás, ya que el concepto de información estaba presente en todas ellas. Surgen así los trabajos de Kampé de Fèriet y Forte sobre la información definida de modo axiomático en contextos con probabilidad o sin ella, continuados posteriormente por muchos investigadores.

Como hemos señalado, es la transmisión de la información la que genera los problemas iniciales. Los comienzos son menos ambiciosos que en la actualidad: se trata de transmitir mensajes suponiendo que los símbolos que se emiten se reciben sin modificación. La aplicación de las técnicas conseguidas permitió, por ejemplo, mejorar los tiempos empleados por el famoso "código Morse" empleado en la transmisión telegráfica.

En seguida se observa que en la transmisión de los mensajes aparecen interferencias, "ruidos" que se añaden durante la transmisión, haciendo que lo recibido no coincida con lo emitido, dificultando así la reconstrucción del mensaje original. Puesto en marcha el aparato matemático consigue los grandes teoremas de codificación con resultados como los de Wolfowitz, que garantizan la existencia de codificaciones que permiten reconstruir los mensajes con un margen de error tan pequeño como se desee. Formidable desde el punto de vista matemático, pero decepcionante desde el punto de vista práctico: los resultados no indican cómo construir tales códigos.

Surgen así los códigos capaces de detectar y corregir errores producidos durante la transmisión y es entonces cuando cobra toda su importancia el "alfabeto binario" formado exclusivamente por los símbolos 0 y 1 que cambia nuestras vidas desde mediados del pasado siglo.

Una vez que se conoce que los códigos detectores y correctores de errores son las soluciones de un sistema de ecuaciones (en un espacio un tanto especial) entran en juego los algebristas que producen gran cantidad de resultados, buena parte de ellos aplicables en la práctica, que dan lugar a nuevas tecnologías (telefonía móvil, MP3, radio digital, etc.)

Pero los códigos detectores y correctores de error, que tan fiable hacen la comunicación pueden ser "vistos" por observadores no deseados. Surge entonces el desarrollo moderno de la Criptografía. La búsqueda de códigos "encriptados" es algo que ha preocupado a la colectividad humana desde casi siempre. Recuérdense a este respecto las escrituras cifradas de las tumbas del antiguo Egipto los "escítalos" espartanos, los "números" y la "caja" de Julio César o, más recientemente las máquinas "enigma" de los alemanes en la segunda guerra mundial (época que coincide, como es bien conocido, con el nacimiento moderno y gran cantidad de desarrollos de la Investigación Operativa).

Hoy los matemáticos más puros, los algebristas, aparentemente los menos aplicados, son los que contribuyen a la búsqueda de nuevos códigos más fiables y rentables para ser aplicados a todas las nuevas tecnologías, muy en particular en lo que afecta a la "seguridad de la información" (tarjetas de crédito, banca on-line, comercio electrónico, etc.). Es la moderna criptografía con bases en la teoría de números, para la que, tristemente, se dedican más esfuerzos en el mundo a la ruptura de claves que a la mejora de las mismas.

En el aspecto más puro de la Teoría (que no de la Transmisión) de la Información se encuentran las aplicaciones a la Inferencia Estadística. Iniciada esta aplicación a finales de los años cincuenta del siglo pasado por Kullback en su clásico texto, con el empleo de la cantidad definida por el mismo y por Leibler, ha tenido un auge espectacular en los últimos años, debido sobre todo a los trabajos de Cressie & Read y en especial a los del grupo liderado por el Prof. Leandro Pardo que, con el empleo de divergencias generalizadas, ha atacado con éxito los problemas de estimación y contraste de hipótesis, proporcionando nuevos métodos que pueden competir ventajosamente con los clásicos.

Para terminar, quisiera expresar (aunque soy consciente de que ahora intervienen a partes iguales el corazón y la razón) mi convicción de que la Teoría de la Información seguirá obteniendo éxito en sus aplicaciones científicas y tecnológicas, a lo largo del presente siglo al menos.

Referencias

[1] ASH, R.B. (1965). *Information Theory*, John Wiley.

[2] GIL, P. (1981). *Teoría matemática de la información*, I.C.E.

[3] KAMPÉ DE FÉRIET, J., and FORTE, B. (1967). Information et probabilité. *C.R.A.S. París*, **265**, 110-114, 142-146, 350-353.

[4] KULLBACK, S. (1959). *Information Theory and Statistics*, John Wiley.

[5] PARDO, L (1997). *Teoría de la Información Estadística*, Hespérides.

[6] PARDO, L. (2006). *Statistical Inference Based on Divergence Measures*, Chapman & Hall.

[7] RÉNYI, A. (1966). *Calcul des Probabilités*, Dunod.

[8] SHANNON, C.E. (1948). A mathematical Theory of Communication. *Bell Syst. Tech. J.*, **27** 379-423, 623- 656.

[9] WELSH, D. (1990). *Codes and Cryptography*, Oxford Science.

[10] WOLFOWITZ, J. (1964). *Coding Theorems of Information Theory*, Springer.