

Aplicación de la teoría de la información a la investigación clínica

RAÚL A. BORRACCI^{MTSAC, 1}, CARLOS D. TAJER^{MTSAC, 2}

Una información es un conjunto de datos que nos permiten aclarar algo sobre aquello que nos es desconocido. Como siempre partimos de un conocimiento *a priori*, datos que confirmen hechos que se saben con seguridad no contienen ni aportan información alguna. La información aumenta en la medida en que el conjunto de datos indique hechos cuya ocurrencia consideramos improbable. En definitiva, podemos enunciar que un conjunto de datos contendrá mayor cantidad de información cuanto menor sea la probabilidad *a priori* de su ocurrencia. Un ejemplo clásico (1) dice así: si un habitante de la Antártida pregunta por el pronóstico del tiempo, el servicio meteorológico (en este caso, la fuente de información) emitirá alguno de los siguientes mensajes:

- 1^{er} mensaje: día muy frío y nublado.
- 2^{do} mensaje: día muy frío y soleado.
- 3^{er} mensaje: día frío y nublado.
- 4^{to} mensaje: día templado y soleado.

De acuerdo con las características climáticas habituales de la Antártida (conocimiento *a priori*), el primer pronóstico contiene mucha menos información que el segundo y así sucesivamente hasta el último mensaje, habida cuenta de que en esta región el 95% de las veces el clima es muy frío y nublado y el 4% de los días son muy fríos y soleados. Por el contrario, un mensaje que sugiera que el clima será frío y nublado o templado y soleado tendrá mucha más información por ser estas situaciones mucho más improbables (históricamente 0,9% y 0,1%, respectivamente). A través de este ejemplo resulta claro que existe una relación inversa entre la cantidad de información que puede entregar un mensaje y la probabilidad de ocurrencia del suceso.

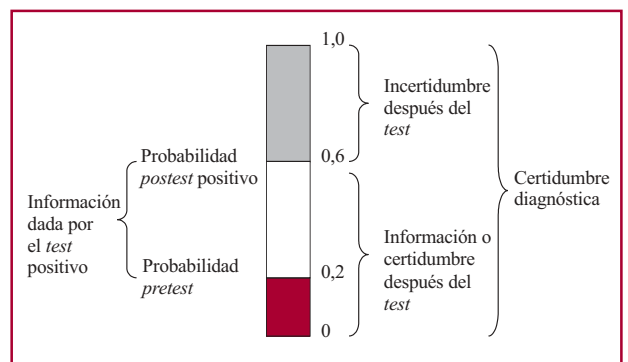
La teoría de la información desarrollada por Claude Shannon estudia, entre otros aspectos, cómo se mide la información, (2) utilizando por analogía el concepto de la entropía tomado de la física. La segunda ley de la termodinámica afirma que en un sistema aislado, es decir, que no intercambia materia ni energía con su entorno, la entropía (desorden en un sistema) siempre irá aumentando (nunca disminuyendo; como mucho, se mantiene) desde que ésta se mide por primera vez hasta otra segunda vez en un momento distinto. Los sistemas termodinámicos tienden a la entropía, tendencia natural a la pérdida del orden. Si colocamos, por ejemplo, dos metales con diferente temperatura en un líquido a baja temperatura, y pudié-

ramos aislar ese sistema del medio circundante, en un momento inicial tendríamos tres niveles de temperatura diferentes. Sin embargo, si dejamos al sistema sin intervenir, tiende a la igualación absoluta de la temperatura, es decir, al absoluto desorden. La tendencia al orden, la inversa de la entropía, se denomina negentropía o entropía negativa. La postulación conocida como entropía de Shannon es la traslación de este concepto a la teoría de la información. En este campo, el orden de un sistema es igual a la cantidad de información necesaria para describirlo. En resumen, toda medición (medida), es decir, toda adquisición de información, supone una interacción consumidora de energía. En relación con la *información* se puede decir que a mayor desorden o entropía, mayor es la cantidad de información necesaria para recuperar un mensaje. La entropía en información es de esta manera un índice del nivel de incertidumbre.

Este último teorema es una de las fuentes de la concepción moderna que demuestra la correlación positiva entre información y entropía negativa (negentropía o, lo que es igual, disminución de la incertidumbre). Es decir, a mayor información, menor incertidumbre.

En medicina, la aplicación de este análisis de la información tiene su ámbito de aplicación en la interpretación de métodos diagnósticos y puntajes pronósticos.

En el caso de las **pruebas diagnósticas**, la información que contienen (y la consiguiente incertidumbre que existe antes y después de aplicar dicha prueba diagnóstica) se puede determinar conociendo tres factores determinantes: la probabilidad *pretest* de la enfermedad o prevalencia, la tasa de verdaderos positivos y la tasa de falsos positivos, de acuerdo con el siguiente esquema:



MTSAC Miembro Titular de la Sociedad Argentina de Cardiología

¹ Bioestadística, Facultad de Ciencias Biomédicas, Universidad Austral

² Director del Grupo GEDIC. Jefe del Departamento Cardiovascular, Instituto Alexander Fleming

En este ejemplo gráfico, la prevalencia o probabilidad *pretest* de tener cierta enfermedad (20%) se incrementa al 60% cuando la prueba diagnóstica es positiva, lo que deja una incertidumbre diagnóstica del 40%. En forma clásica, estos métodos se describen en términos de sensibilidad y de especificidad, y permiten a través de la formulación bayesiana establecer la probabilidad *postest* sobre la base del *likelihood ratio* y del *odds* o chance previa (prevalencia *pretest*). (3, 4) Estas formulaciones son complejas. Aunque estamos menos familiarizados con su uso, la teoría de la información permite calcular dicha incertidumbre *preprueba* y *posprueba* expresada en dígitos binarios o *bits*, en forma más comprensible. Parte inicialmente de la siguiente ecuación: (5)

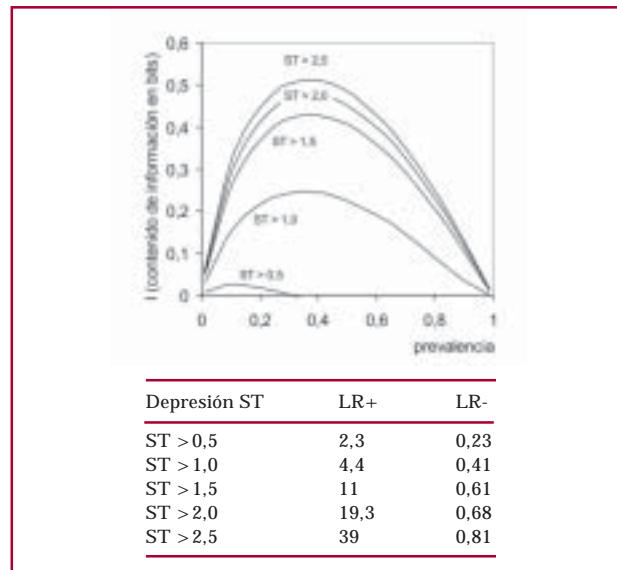
Contenido de información (I) =

$$= a.p.\log_2(a.p) + b.q.\log_2(b.q) - (a.p + b.q)\log_2(a.p + b.q) - p.\log_2 p$$

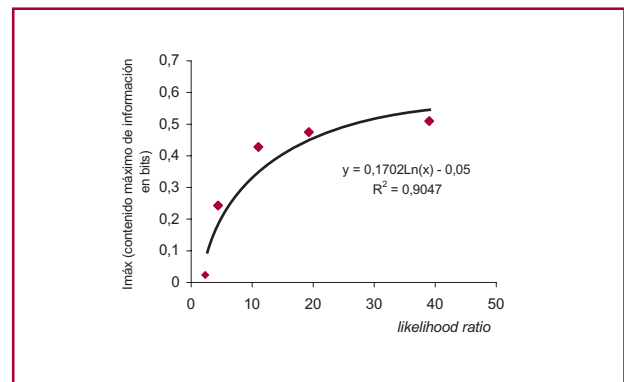
- donde a = tasa de verdaderos positivos
- b = tasa de falsos positivos
- p = prevalencia de la enfermedad o probabilidad *pretest*, y
- q = prevalencia de la no enfermedad (1 - p)

El resultado obtenido al aplicar la ecuación se hallaría entre 0 y 1 *bit*. Si la prueba diagnóstica provee toda la información necesaria para hacer el diagnóstico (certeza completa), entonces el contenido de información de la prueba será de 1 *bit*. En caso de que sólo dé información parcial, por ejemplo 0,5 *bit*, esto indicará que la prueba provee sólo la mitad de la información necesaria para establecer o excluir un diagnóstico.

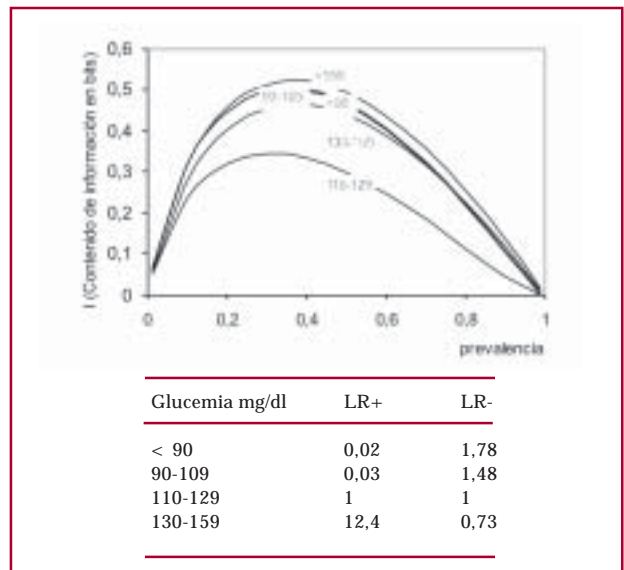
En el siguiente gráfico se calculó el contenido de información (I) que provee la prueba ergométrica de acuerdo con diferentes puntos de corte (depresión del segmento ST en mm) para diagnosticar enfermedad coronaria en una población con distintas tasas de prevalencia de la enfermedad (*gold standard*: angiografía con estenosis mayor del 50% en por lo menos un vaso): (5)



Cuando la prueba es positiva, se observa que a medida que se elige un punto de corte más alto aumenta el contenido de información de la prueba para realizar el diagnóstico de coronariopatía. De la misma forma, el gráfico ofrece información sobre todo del espectro de la prevalencia o probabilidad *preprueba*. Esta última característica es una de las ventajas que tiene sobre el uso del *likelihood ratio*, habida cuenta de que éste no considera la prevalencia de la enfermedad en los individuos estudiados, aunque esta deficiencia se corrige al calcular el *odds* *posprueba*. La relación entre *likelihood ratio* y el contenido máximo de información ($I_{\text{máx}}$), definido este último como el valor más alto obtenido dentro del rango de prevalencia, se correlaciona en el siguiente gráfico para el mismo ejemplo del diagnóstico de enfermedad coronaria con la prueba de esfuerzo:



En realidad, el contenido de información presenta una relación con el *likelihood ratio* (LR) más compleja, que tiene en cuenta a su vez el valor del LR+ y del LR-. En el gráfico del siguiente ejemplo se considera el valor de la glucemia basal para el diagnóstico de diabetes: (4)



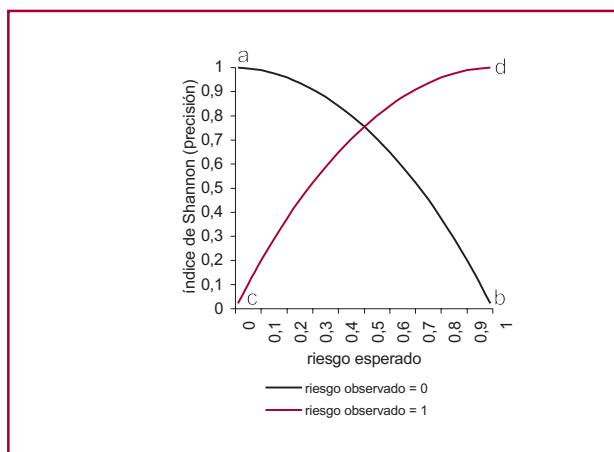
A diferencia del ejemplo de la depresión del ST, en esta ocasión el LR- alcanza valores “multiplicadores” mayores que el LR+, cuando la glucemia se halla por debajo de 110 mg/dl; por eso, las curvas de contenido de información demuestran que los niveles de glucemia por encima de 159 y por debajo de 110 mg/dl son los que proveen mayor información para confirmar o descartar un diagnóstico.

Como en el método ROC (*receiver operating characteristic*), la integración (integral matemática) de las curvas de contenido de información (I) permite obtener las áreas para cada valor de corte, equivalentes a cada $I_{m\acute{a}x}$.

La **precisión (accuracy)** de una prueba diagnóstica también puede determinarse en base al modelo de la teoría de la información. En este caso se evaluará la estimación del riesgo realizada por la prueba diagnóstica y se cotejará con la presencia o no de la enfermedad, de acuerdo con la siguiente fórmula: (6-8)

$$\text{Precisión (S)} = [(1 + X_i) \ln(1 + p_i) + (2 - X_i) \ln(2 - p_i) - \ln 2] / \ln 2$$

donde p_i es la probabilidad estimada de que el paciente “i” tenga la enfermedad, y X_i indica la presencia real de la enfermedad ($X_i = 1$) o su ausencia ($X_i = 0$). Este valor de precisión se conoce también como índice de Shannon y presenta la siguiente relación cuando se compara con el riesgo esperado y observado:



En este caso se observa que, cuando el riesgo observado es 0 (ausencia del evento), el índice disminuye a medida que el riesgo esperado crece (el riesgo esperado indica mayor probabilidad de ocurrencia del evento). Por ejemplo, en (a) el riesgo observado es 0 (no ocurre el evento) y el riesgo pronosticado por la prueba (riesgo esperado) también es 0; por lo tanto, la precisión de acuerdo con el índice de Shannon es máxima (uno). En (b), en cambio, el riesgo observado sigue siendo 0, pero la prueba indica un riesgo esperado máximo (uno); por lo tanto, la precisión de este pronóstico se acerca a un índice de Shannon igual a 0. Por el contrario, cuando el riesgo observado es 1 (presencia del evento), el índice aumenta a medida que el riesgo esperado también crece. En el mismo ejemplo

se observa que en (c) el riesgo observado vale 1 (ocurre el evento) y en cambio el riesgo esperado predicho por la prueba es 0; en consecuencia, la precisión de dicha prueba es mínima (índice de Shannon igual a 0). Por último, en (d), el riesgo observado en el paciente es 1 (presencia del evento) y el esperado también es 1 (la prueba también predice que ocurrirá el evento); por lo tanto, la precisión de la prueba para diagnosticar la ocurrencia del evento es máxima.

A modo de ejemplo, supongamos que una prueba diagnóstica determina los distintos riesgos de padecer una enfermedad coronaria en un grupo de 10 pacientes:

Paciente #	Riesgo de coronariopatía según prueba diagnóstica (p)	Presencia de la enfermedad (X _i)	Precisión (S _i)
1	0,05	0	0,99
2	0,12	1	0,24
3	0,35	1	0,59
4	0,02	0	0,99
5	0,63	1	0,86
6	0,07	0	0,99
7	0,15	0	0,98
8	0,40	1	0,65
9	0,32	1	0,55
10	0,01	0	0,99

Después de calcular la precisión para cada diagnóstico individual, se pueden obtener el promedio y la desviación estándar para el grupo de pacientes, que en este caso serán $0,79 \pm 0,266$. En consecuencia, esta modalidad permite la estimación de la precisión de la prueba diagnóstica, tanto a nivel del paciente individual como del grupo o población.

Finalmente, las curvas ROC también pueden evaluarse a partir de la teoría de la información. Los detalles de su método son similares a los expuestos y pueden consultarse en Metz y colaboradores. (9) De la misma forma, el análisis cuantitativo de las imágenes radiológicas aplicadas al diagnóstico puede revisarse en Canals. (10)

La aplicación de los métodos relacionados con la teoría de la información parte de un razonamiento matemático aún más complejo que el expuesto en este artículo, (11) y es quizás esta dificultad lo que hace que esta metodología no se aplique tan asiduamente en medicina como en otras áreas.

Sin embargo, con los recursos computacionales disponibles actualmente, el concepto de $I_{m\acute{a}x}$ puede competir con ventajas con el área de la curva ROC y el *likelihood ratio*, poco intuitivos para el médico práctico, y eventualmente transformarse en una herramienta adicional para el análisis de la precisión de pruebas diagnósticas y pronósticas.

BIBLIOGRAFÍA

1. Castro Lechtaler AR, Fusario RJ. Teoría de la información y de la codificación. En: Castro Lechtaler AR, Fusario RJ. *Teleinformática*. 2ª ed. Vol I. Barcelona, España: Editorial Reverté SA; 1999. p. 263-73.
2. Shannon CE, Weaver W. *The mathematical theory of communication*. Chicago: University of Illinois Press; 1949.
3. Doval HC. Acerca de la probabilidad, la probabilidad condicional y el teorema de Bayes. *Rev Argent Cardiol* 2004;72:293-8.
4. Doval HC, Tajer CD. Interpretación de ensayos clínicos. En: Doval HC, Tajer CD. *Evidencias en Cardiología IV*. Buenos Aires: GEDIC; 2005. p. 52-4.
5. Diamond GA, Hirsch M, Forrester JS, Staniloff HM, Vas R, Halpern SW, et al. Application of information theory to clinical diagnostic testing. The electrocardiographic stress test. *Circulation* 1981;63:915-21.
6. Hlatky M, Botvinick E, Brundage B. Diagnostic accuracy of cardiologists compared with probability calculations using Bayes' rule. *Am J Cardiol* 1982;49:1927-31.
7. Bobbio M, Deorsola A, Pistis G, Brusca A, Diamond GA. Physician perception of exercise electrocardiography as a prognostic test after acute myocardial infarction. *Am J Cardiol* 1988;62:675-8.
8. Pinna-Pintor P, Bobbio M, Colangelo S, Veglia F, Giammaria M, Cuni D, et al. Inaccuracy of four coronary surgery risk-adjusted models to predict mortality in individual patients. *Eur J Cardiothorac Surg* 2002;21:199-204.
9. Metz CE, Goodenough DJ, Rossmann K. Evaluation of receiver operating characteristic curve data in terms of information theory, with applications in radiography. *Radiology* 1973;109:297-303.
10. Canals M. Bases cuantitativas del análisis de imágenes médicas: información e inferencia bayesiana. *Rev Méd Chile* 2003;131:553-61.
11. Abramson N. *Teoría de la información y la codificación*. 6ª ed. Madrid: Paraninfo; 1986.