



CORRELACIÓN Y CAUSALIDAD

Jose Miguel Vega

Analista Senior de Investigación – Ipsos Marketing

E-mail: jose.vega@ipsos.com

De manera natural, casi inconsciente, tendemos a asignar una causa a los sucesos que vemos de manera cotidiana. Lo hacemos basándonos en la experiencia, la lógica o una combinación de ambas. Un ejemplo típico es asociar un movimiento sísmico con el “cambio de clima”. Ocioso es exponer las muchas veces que se ha dicho que no existe evidencia que sustente esta relación.

Como casi todo lo cotidiano, este tipo de relaciones causa-efecto son medibles. La forma más común de hacerlo es usando la **correlación**. La intención, al usarla, es conocer el efecto de una variable (sabor, olor, rendimiento, confiabilidad, precio) sobre el agrado general, la satisfacción o la intención de adquirir determinado bien o servicio.

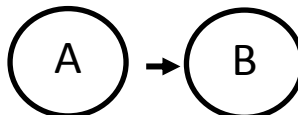
Sin embargo, no debemos olvidar que la correlación es un cálculo matemático frío y que, como tal, es susceptible a errores que están más allá del cálculo numérico. A la hora de tomar decisiones, la lógica, el sentido común y la experiencia que tenemos – tanto en Ipsos como investigadores y las empresas como desarrolladoras de productos y servicios – deben completar el dibujo de una realidad que la computadora que hace el cálculo no conoce. De aquí que se desprende la afirmación: **La existencia de correlación no implica necesariamente causalidad.**

Sin problema, se puede demostrar numéricamente que *“comer chocolate aumenta tu probabilidad de ganar un premio Nobel”*. Así lo hizo el médico Franz Messerli, al comparar el consumo per cápita de chocolate de los países y la proporción de personas en estos que han ganado un Nobel. No solo la conclusión del estudio es caprichosa, sino que el cálculo matemático para llegar a ella también lo es. Sin duda, el cálculo está bien hecho, los números bien ingresados y las fórmulas perfectamente usadas; pero el sentido común nos dice que no podemos dar esa atribución al chocolate (sin importar cuánto nos guste).

Otro error común al usar correlaciones, es asignar la causa de un suceso a una variable que, aunque sí está relacionada, es en realidad causada por una tercera, que es la causante real de las dos primeras. Este enredo de palabras se puede explicar mejor gráficamente

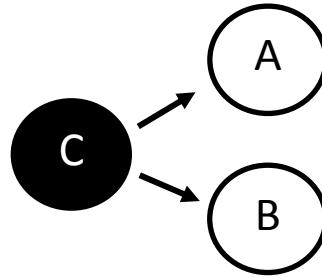
“Comer helados (A) causa más ahogamientos (B): Se demuestra que el aumento en el consumo de helados, está directamente relacionado con el aumento en el número de ahogamientos”.

Imagen 1:



¿Podemos demostrar esta afirmación? Matemáticamente, sí. Pero, ¿son los helados la causa de los ahogamientos? ¿Debemos prohibir los helados? La respuesta a estas dos últimas preguntas es no. Estamos olvidando una tercera variable que influye en las dos anteriores: el calor (C).

Imagen 2:

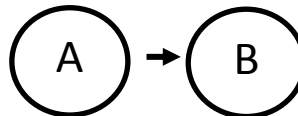


Es fácil notar que es el calor lo que hace que aumente el consumo de helados. También, el calor alienta que las personas vayan a playas y piscinas a bañarse mucho más, por lo que aumenta el número de ahogamientos. Por lo tanto, no es que el consumo de helados y el número de ahogamientos tengan alguna relación, lo que sucede es que ambas se ven influenciadas en un mismo sentido por una tercera variable originaria.

Le llamo originaria en este caso pues, es una variable que es anterior a las planteadas. Pero hay casos en que esa tercera variable no es originaria, sino intermedia.

“Ver televisión más de 3 horas diarias (A) acorta la vida 5 años (B)”.

Imagen 3:



En este caso, el error que se comete es omitir una variable en el circuito lógico. ¿Podemos atribuir a la televisión la reducción en nuestra expectativa de vida? Quizás sí, pero no de manera directa. El ver televisión más de 3 horas diarias es solo el reflejo de otra variable: el sedentarismo (C). Si vemos televisión 3 horas diarias o más, lo más probable es que seamos personas sedentarias, sin actividad física que nos dé una vida más saludable y por lo tanto, mayor expectativa de vida.

Imagen 4:



Con la afirmación inicial, estábamos atribuyendo a una variable un efecto que, seguramente, puede ser compartido con muchas otras pero que al final decantan en una, en este caso, el sedentarismo.



Así pues, al indagar por **correlaciones** y **causalidades**, debemos tener siempre en cuenta que existen muchas variables que pueden estar jugando en conjunto. Se hace necesario entonces tener hipótesis planteadas desde el conocimiento del producto o servicio que es materia de estudio, y con ello tomarnos el tiempo de encontrar – a priori – las variables que nos ayuden a explicar las causas de nuestro problema de investigación, y a diferenciar – a posteriori – aquellas que sean realmente relevantes de las que no lo son. En este sentido, para Ipsos es fundamental desarrollar un *partnership* con nuestros clientes a fin de compartir experiencias y conocimientos para entregar resultados accionables. No olvidemos nunca que el sentido común usa mucho el razonamiento lógico pero se basa en la experiencia y el conocimiento.