

Poco ruido y muchas nueces

Posted on 3 de octubre de 2014 by Juan José Gómez Cadenas

Natalia X acaba de llegar a casa, tras una semana de viaje de negocios que ha terminado antes de lo previsto. Su marido, Adrián Z, no la espera hoy y Natalia saborea la sorpresa que va a llevarse, ya que planea celebrar su regreso anticipado invitando a Adrián cenar a algún sitio romántico. Animada por la perspectiva, se levanta de un salto y se dirige a la ducha. Después de refrescarse, y decidida a que la noche salga redonda, rebusca en el cajón donde guarda su lencería, queriendo seleccionar su combinación más seductora.

Sorpresa. De repente, frente a sus incrédulos ojos aparecen unas braguitas rojas, descocadísimas, *que no son suyas*.

Pero, ¿no era esta una reseña sobre un libro llamado *La señal y el ruido*? ¿Qué tienen que ver la estadística con las tribulaciones de la pobre Natalia? Más de lo que usted se imagina, asegura Nate Silver, el autor de esta extensa, divertida, algo desorganizada y muy interesante obra.

Debo confesar que comencé la lectura del libro con cierto escepticismo, sabiendo que me las veía con un *bestseller* (¿desde cuándo un libro de matemáticas que merezca su nombre arrasa en las listas de superventas?) escrito por un jovencito espabilado (a Silver le viene al pelo la palabra inglesa *hotshot*) cuyo sospechoso pedigrí combina las actividades de

futurólogo político, exjugador profesional de póquer y diseñador de software (supuestamente) capaz de predecir los resultados de variadas quinielas deportivas.

Y, sin embargo, la lectura del primer capítulo, dedicado a la economía, no tardó en engancharme. El autor sabe ir al grano y, después de una somera introducción, en la que nos recuerda los desvaríos cometidos por las agencias de *rating* durante los años del *boom* inmobiliario, identifica certeramente la razón subyacente a sus errores, a saber, la incapacidad de distinguir entre dos conceptos, a simple vista similares y en la práctica muy diferentes: *incertidumbre* y *riesgo*.

El *riesgo*, nos apunta, es algo a lo que puede ponerse precio. Supongamos que, mientras Natalia estruja, con lágrimas en los ojos, el trapito carmesí que acaba de aparecer en sus cajones, Adrián está echando una partida de póquer con los amigos. Examina las cartas y se encuentra con que tiene cuatro ases. Es decir, tiene la mano ganada, a no ser que otro de los jugadores saque escalera de color. Un cálculo elemental arroja que la escalera de color sólo sale en una de cada cien mil manos (para ser exactos, la probabilidad de acertar con ella en un mazo de cincuenta y dos cartas es de 0,0000138). Por tanto, el *riesgo* que corre Adrián es muy pequeño, de modo que resulta razonable que ponga todos sus ahorros encima de la mesa.

La *incertidumbre* se da cuando no se sabe calcular el riesgo. Es posible realizar una estimación de la variable en cuestión (digamos la evolución de un determinado activo en Bolsa), pero no es posible realizar una estimación fiable de la desviación típica respecto a ese valor medio. Ese fue, precisamente, el problema de las agencias de *rating*. Eran capaces de realizar predicciones muy precisas (en el sentido de que se atrevían a ponerle varias cifras decimales a sus estimaciones) y a la vez totalmente infundadas (ya que no tomaban en cuenta correlaciones y variables ocultas que afectaban extraordinariamente a los valores medios que arrojaban sus modelos, lo que se traducía en variaciones con respecto a estos de varios órdenes de magnitud). El capítulo concluye con una interesante reflexión: «Uno de los riesgos más generalizados de la era de la información es que, a medida que crece la cantidad de conocimiento existente en el mundo, se ensancha la brecha entre lo que sabemos y lo que creemos saber».

Tras la economía, vienen las predicciones políticas. Dado que el éxito del libro se debe en buena parte a que Silver acertó a predecir todos y cada uno de los Estados en que el presidente Obama ganó las elecciones, el capítulo prometía ser uno de los más interesantes. Curiosamente, me resultó uno de los más pesados. El autor divaga bastante, metiéndose con los analistas políticos y tertulianos televisivos que infestan los medios de comunicación norteamericanos (y los nuestros) y, de paso, se dedica una buena cantidad de

autopropaganda que acaba por resultar un poco cargante. Tampoco el capítulo 3, dedicado a las predicciones deportivas, en concreto a mundo de las apuestas en torno al rendimiento de los jugadores de béisbol, me resultó muy atractivo. Mi consejo: si se les hace cuesta arriba la lectura, pasen de largo. Lo mejor está todavía por llegar.

Después de la economía, la política y la pelota base, le toca el turno al tiempo. La evolución de las condiciones meteorológicas en un área determinada (esto es, las familiares predicciones de dónde va a hacer sol y dónde va a llover que cada noche nos ofrece el hombre del tiempo) es un problema muy complejo, dado que la atmósfera se comporta como un sistema caótico. En un sistema caótico es posible realizar predicciones a corto plazo (a partir de la observación de unas condiciones iniciales), pero, a medida que este evoluciona, se torna más y más impredecible. En otras palabras, si amanece soleado en Madrid el lunes, las probabilidades de que el martes haga bueno son altas, pero no tenemos ni idea de qué esperar la semana que viene.

El capítulo, muy brillante, explica cómo las predicciones meteorológicas se realizan, de hecho, a partir de una combinación de cálculos en supercomputadores, que simulan miles de posibles configuraciones para la atmósfera (a partir de las cuales deducen valores medios y desviaciones típicas) y una muy considerable cantidad de heurística, esto es, cuentas de la vieja y recetarios prácticos, desarrollados durante décadas de práctica en los equipos de «hombres del tiempo». Es todo un ejemplo de matemáticas prácticas. En el mismo capítulo se trata con autoridad la forma en que los servicios de meteorología pueden predecir fenómenos violentos, tales como huracanes. De hecho, el tristemente famoso Katrina, que devastó la ciudad de Nueva Orleans, no fue otra cosa que la crónica de un desastre anunciado.

De la predicción del tiempo pasamos a la de los terremotos, un problema que, a primera vista, podría parecernos similar, pero que resulta ser mucho más intratable. La diferencia es que, en este caso, no se dispone de un buen modelo matemático, como en el caso de la atmósfera, y, por tanto, resulta imposible realizar predicciones sensatas. En resumen, se trata de un caso clarísimo en el que el ruido enmascara irremediablemente a la señal.

De los terremotos, Silver vuelve a la economía para predicar con bastante acierto su ya bien establecida tesis. Lo que cuenta en la evolución de los activos bancarios o las políticas monetarias y laborales es la incertidumbre, esto es, la dificultad de cuantificar el riesgo, asociada con la naturaleza caótica del sistema. El autor lo resume muy bien. Una mariposa bate sus alas en Brasil y alguien pierde su trabajo en Texas.

Pero volvamos a Natalia. El consejo que Nate Silver ofrece a nuestra atribulada amiga es el

siguiente. Las braguitas rojas parecen inculpar a Adrián de un delito de infidelidad. Pero, ¿cuán segura estás de ello? La solución al dilema de nuestra protagonista nos la ofrece el capítulo ocho, clave del libro y también uno de los más divertidos. Silver nos presenta al reverendo Thomas Bayes y nos introduce a su famoso teorema. Revisémoslo considerando el caso de Natalia. La pregunta que esta debe hacerse es: «¿Cuál es la probabilidad de que Adrián esté engañándome con otra?» (caer en la cuenta de que la pieza de lencería implica una probabilidad de culpa y no una certeza absoluta es ya un salto de gigante). Para evaluar esa *hipótesis*, partimos del hecho (*condición*) de haber encontrado una prenda acusatoria en la cómoda. Pues bien, el teorema de Bayes nos propone una solución, siempre que podamos estimar tres cantidades, a saber:

> 1) En la hipótesis de que haya habido traición, ¿qué probabilidad existe de que la pelandusca haya olvidado las braguitas? Es decir, asumimos que Adrián es culpable y que la prenda de ropa interior ha sido olvidada por alguna pelandusca después del pertinente revolcón. Pero un momento. Adrián, como Natalia sabe bien, es un tipo bastante cabal. Realmente, si fuera culpable, ¿sería tan descuidado como para no revisar los cajones a la búsqueda de alguna prueba inculpatoria? No parece nada obvio, así que, ante la duda, asignamos un 50% a la probabilidad de que las braguitas hayan aparecido, *condicionadas al hecho de que Adrián le sea infiel*.
>
> 2) En la hipótesis de que no haya habido traición, ¿qué probabilidad existe de que hayan aparecido las braguitas? En otras palabras, ¿hay alguna otra explicación que justifique la aparición de la prenda si el marido es inocente? Sin duda es posible encontrar alguna, desde la simple equivocación en la lavandería hasta la posibilidad de que a Adrián le dé por vestir, a escondidas, combinaciones femeninas (habría que ver si esta última posibilidad preocuparía a Natalia más o menos que la potencial infidelidad).
>
> 3) Por último, pero quizás lo más importante. Natalia necesita lo que en la jerga bayesiana se conoce como un «prior», a saber: ¿qué probabilidad le habría asignado a la hipótesis de que Adrián la engañara *antes* de encontrar las braguitas? Cierto, a la pobre muchacha va a resultarle difícil ser objetiva mientras estruja la prenda acusatoria, pero la clave del prior es que seamos capaces de establecerlo de antemano. Y ella está demasiado afectada para razonar: podemos recurrir a los estudios estadísticos que aseguran, que, en promedio, «sólo» el 4% de los maridos engañan a sus esposas en un año determinado.

A partir de aquí, el teorema de Bayes nos proporciona una respuesta, la que llamamos «probabilidad a posteriori» y que, de hecho, es exactamente lo que nos interesa: «¿Cuán probable es que Adrián esté engañando a Natalia, dado que hemos encontrado las

braguitas?» El cálculo, utilizando la sencilla fórmula de Bayes, arroja una probabilidad del 29%, bastante más pequeña de lo que nos imaginaríamos a primera vista. Naturalmente, si cambiamos el valor del prior de un 4% a un número superior, la probabilidad posterior aumenta. Pero a Natalia le conviene no precipitarse. A no ser que tenga buenas razones para pensar que su marido la engaña (no relacionadas con el episodio que nos ocupa), lo mejor que puede hacer es ceñirse a las razonables estimaciones estadísticas.

La historia de Natalia y Adrián le sirve al autor para entrar en materia y demostrar la versatilidad del teorema de Bayes como herramienta estadística. De hecho, reformula inmediatamente el episodio de la putativa infidelidad conyugal, proponiendo el cálculo de la probabilidad de que una mujer de cuarenta años tenga cáncer de mama, *sabiendo que la mamografía ha arrojado un resultado positivo*. El resultado es sorprendentemente bajo por las mismas razones que antes. El prior (la probabilidad de que una mujer de cuarenta años desarrolle cáncer de mama, establecida como una media estadística sobre un grupo grande de población) es muy bajo, de apenas el 1,5 %. Por ello, a pesar de que la probabilidad de que la mamografía acierte es del 75%, la probabilidad posterior es tan solo de un 10%.

En este punto, y después de establecer la utilidad de la estadística bayesiana, Silver podía haber dado el libro por concluido. Nada más lejos de sus intenciones. Sigue un capítulo dedicado al ajedrez (y a las máquinas que nos superan ya ampliamente en el juego) y otro dedicado al póquer (actividad a la que se dedicó, con razonable éxito, durante una temporada), antes de exponernos cómo el teorema de Bayes puede ayudar a la hora de planificar inversiones en Bolsa y concluir con un capítulo dedicado al cambio climático. La excusa narrativa en toda la segunda parte del libro está clara (cómo aplicar la estadística bayesiana) y Silver no duda en repetirla, pero el hilo conductor, a veces, es bastante tenue y a ratos da la impresión de que el autor se recrea en temas en los que no es un experto (el ajedrez o el cambio climático, por ejemplo), aunque no por ello la lectura deja de ser amena e interesante.

La obra, en su conjunto, vale la pena, a pesar de sus eventuales incursiones en los cerros de Úbeda. Silver es lo bastante divertido, brillante e inteligente como para que se le perdone el eventual exabrupto, el ocasional desliz o el pecadillo, nada infrecuente, de estar encantado de haberse conocido. En resumen, aunque el libro no está exento de cierta cantidad de ruido, la señal, afortunadamente, lo supera con creces.

Juan José Gómez Cadenas es científico y profesor de investigación en el CSIC. Es autor de *Materia extraña* (Madrid, Espasa Calpe, 2008), *El ecologista nuclear* (Madrid, Espasa Calpe, 2009) y *Spartana* (Madrid, Espasa Calpe, 2014).

Autor: Nate Silver

Edicion: Barcelona, Península, 2014

Paginas: 656 pp. 27,90 €

Titulo: La señal y el ruido