Michel Talagrand y las desigualdades estocásticas en sistemas de muchas dimensiones

JOSÉ ANTONIO LÓPEZ DÍAZ

El matemático francés Michel Talagrand (1952, Béziers) ha sido galardonado este año con el premio Abel, uno de los más preciados galardones en la comunidad matemática. Su área de trabajo ha sido fundamentalmente el análisis funcional y los procesos estocásticos. En esta última área sus resultados más importantes se refieren a desigualdades estocásticas (probabilísticas). En este artículo haré un bosquejo de alguna de sus más conocidas aportaciones en este campo, empezando por comentar el destacado papel que juegan las desigualdades probabilísticas en la teoría de la probabilidad.

Al ser una probabilidad básicamente una medida de la incertidumbre acerca de la ocurrencia de un determinado suceso, al igual que con otras medidas de carácter físico como volumen, peso, etc., a menudo basta una cota aproximada, inferior o superior, para satisfacer nuestras necesidades. De hecho en la teoría de la probabilidad y estadística las desigualdades desempeñan un papel muy importante. Algunas son de carácter elemental, pero no por ello menos útiles en muchos casos. Una de las más básicas es que la probabilidad de la unión de dos sucesos (o sea, que ocurra al menos uno de los sucesos) es menor o igual a la suma de las probabilidades de ambos. Algo menos evidente que esta es la que se obtiene pasando a los sucesos complementarios, que afirma que la probabilidad de la intersección de dos sucesos (que ambos ocurran) es mayor o igual a la suma de las probabilidades menos uno[1]. Si por ejemplo sabemos que mañana hay una probabilidad de 60 % de que llueva y de 70 % de que el viento supere los 80 km/h, entonces sabemos que hay al menos una probabilidad de un 30 % (60 + 70 -100) de que se den ambos fenómenos[2].

Numerosos teoremas en probabilidad se han podido probar usando desigualdades estocásticas. Una de las más célebres es la desigualdad de Chebyshev, que establece una cota para las desviaciones grandes en valor absoluto de una variable aleatoria respecto de la media. En concreto dice que la probabilidad de que el cociente entre la desviación absoluta respecto de la media y la desviación típica de la variable[3] sea mayor que un número k es menor o igual que $1/k^2$. Esta desigualdad por tanto nos permite acotar las desviaciones grandes de una variable aleatoria con tal de conocer su varianza, lo cual puede tener mucho interés. Pensemos en los extremos climáticos sin ir más lejos, aunque para este caso no es muy efectiva y suele haber mejores alternativas si se puede aproximar la función densidad de probabilidad (pdf) de los extremos[4]. En la teoría de las probabilidades uno de los resultados más importantes, la ley de los grandes números (débil), se puede demostrar fácilmente usando la desigualdad de Chebyshev. Esta ley fue probada por Jacobo Bernoulli a principios del siglo XVIII, fue el primero de los teoremas de esta clase que sirven para cimentar la interpretación frecuentista de las probabilidades [5].

Pasemos ahora a comentar la aportación de Talagrand. El fenómeno de concentración de la medida (o probabilidad) es muy importante en cuanto el número de variables aleatorias que conforman un proceso estocástico crece. En este caso de dimensión grande del proceso nuestra intuición a menudo falla clamorosamente. Un ejemplo típico es calcular el número mínimo de alumnos en una clase para que haya una probabilidad del 50 % o más de que al menos dos de ellos nazcan el mismo día del año. Sorprendentemente, pese a que la probabilidad de que una pareja concreta de alumnos cumpla años el mismo día es muy baja, 1/365, basta que haya 22 alumnos en la clase para que se cumpla lo pedido. Es la dimensionalidad del problema la que hace que la probabilidad muestre estas aparentes paradojas. De hecho la Física Estadística se fundamenta en la enorme concentración de la probabilidad en sistemas de un número enorme de partículas. En concreto, si prácticamente toda la probabilidad se concentra en la capa externa del espacio probabilístico[6].

Unos de los resultados más importantes de Talagrand es una desigualdad de concentración de la medida. Viene a decir que en un espacio probabilístico producto, o sea, que describe los valores posibles de un vector estocástico compuesto por una sucesión de variables aleatorias, el producto de la probabilidad de un subconjunto del espacio por la probabilidad de estar a una distancia del mismo al menos t (según una distancia que introdujo Talagrand), está acotada por arriba por $\exp(-t^2 / 4)$. Esta disminución exponencial con t de la cota superior indica que se vuelve rápidamente más improbable estar fuera de un entorno de cualquier región del espacio probabilístico a una distancia al menos t, implicando por tanto una alta concentración de la pdf de los estados del vector estocástico formado por variables aleatorias independientes. De esta forma se pueden acotar eficientemente las fluctuaciones de procesos estocásticos.

[1] p(A y B) = 1 – p ((no A) o (no B)) en virtud de la ley de Morgan >= 1 – p (no A) – p (no B) = p (A) + p (B) -1. Si p(A) + p(B) <1 la cota es trivial pues cualquier probabilidad es >=0.

[2] Naturalmente podemos tener más información que nos haga elevar la probabilidad de la intersección, como que exista una correlación positiva importante entre ambos sucesos, pero esta cota inferior del 30 % se cumple necesariamente (incluso con la mayor anticorrelación). Por otra parte es claro que la probabilidad de la intersección puede ser como máximo igual a la menor de las dos probabilidades, en este ejemplo 60 %.

[3] Si existe, no todas las variables aleatorias tienen varianza finita, aunque sí la mayoría en las aplicaciones.

[4] Es conocido que una variable normal estándar tiene un 5 % de probabilidad de superar en valor absoluto 1.96. Aplicando la desigualdad de Chebyschev obtenemos una cota máxima para esta probabilidad de $1/1.96^2 \sim 25$ %, no muy aquilatada. Pero la utilidad de la desigualdad de Chebyschev radica en que es válida independientemente de la pdf de la variable aleatoria.

[5] Esta ley afirma que la media aritmética de variables aleatorias independientes e idénticamente distribuidas converge en probabilidad a la esperanza matemática de las mismas.

[6] Si tomamos una esfera de radio 1 en un espacio de N dimensiones, la proporción de volumen en un casquete esférico de espesor muy pequeño d crece con N. En dos dimensiones tenemos que esa proporción vale $2 \pi d / \pi = 2d$; en tres dimensiones $4 \pi d / (4/3 \pi) = 3 d$. En general, al ser V proporcional a r^N se tiene $dV/V = d \ln V = N d$.