Algunas curiosas integrales

CON INTERPRETACIÓN PROBABILISTA

JOSÉ ANTONIO LÓPEZ DÍAZ. AEMET, MADRID

Uno de los aspectos más fascinantes de las matemáticas es la aparición de relaciones, en apariencia insospechadas, entre áreas distintas de la matemática. La teoría de la probabilidad es fecunda en producir este tipo de extraños maridajes. Un ejemplo clásico es el del problema de la aguja de Buffon, propuesto por este matemático en el siglo XVIII. El problema consiste en calcular la probabilidad de que al dejar caer una aguja de longitud l sobre una superficie plana con líneas rectas (infinitas) paralelas equidistantes separadas una distancia t, la aguja toque una línea. Lo curioso de este problema es que la solución contiene la constante π, y por tanto podría uno en principio calcular el valor de π a base de arrojar una aguja muchas veces sobre un patrón de líneas rectas equidistantes y contar. En concreto, la probabilidad del problema de Buffon, para el caso más simple de que $l < t$, vale $2\,l / t\,\pi$.

Georges-Louis Leclerc, conde de Buffon

Hay otro tipo de probabilidades que se relacionan con determinadas integrales que a mi me parecen también muy sugestivas. Empecemos planteando el problema de calcular la integral múltiple:

$$I_1 = \iiint_{0<u_1<u_2<\ldots<u_n<1} du_1 du_2 \ldots du_n$$

donde el dominio de integración es todo el hipercubo $[0, 1]^n$ con la condición indicada bajo el signo integral, de forma compacta. En dos dimensiones n = 2, por ejemplo, la integral sería

$$\int_{u_1=0}^{1} du_1 \int_{u_2=u_1}^{1} du_2$$

Esta integral vale 1/ n!, lo que se puede ver por integración directa usando la inducción, o también con razonamientos de tipo geométrico e inducción (n = 2 área de un triángulo, n = 3 área de un tetraedro,…).

Pero también la podemos hallar con un razonamiento estrictamente probabilístico. Para ello reparemos en que si tenemos n variables aleatorias (v.a. en lo sucesivo) U_i, i = 1, 2, ..,n, uniformes en [0, 1], independientes entre sí, la cantidad infinitesimal del integrando de I_1 es la probabilidad de que la primera v.a. esté en du_1, la segunda en du_2, y así sucesivamente hasta la n-ésima en du_n (recordemos que para una función uniforme [0, 1] su función densidad vale 1 en el intervalo y 0 fuera de el, lo que significa que la probabilidad de que caiga en un intervalo [a, b] cualquiera contenido en [0, 1] es b-a). Por tanto, dada la forma del dominio de integración, la integral múltiple es igual a la probabilidad del suceso $O = \{U_1 < U_2 < \ldots < U_n\}$. Esta probabilidad la podemos calcular aplicando la simetría del problema, que nos dice que la probabilidad de cualquier suceso de la forma $U_{i1} < U_{i2} < \ldots < U_{in}$ con {i1, i2, .., in} una permutación cualquiera de {1, 2, …, n} tiene que ser la misma que la de O (y la suma de todas esas probabilidades obviamente tiene que ser la unidad). Por tanto, como hay n! permutaciones, concluimos que $pr(O) = I_1 = 1/n!$.

Antes de pasar a casos más complicados, me parece interesante mencionar el caso de la integral elemental

$$\int_0^1 u^n \, du$$

que rápidamente vemos que vale 1/(n+1). Por supuesto esto lo sabemos porque, tras varios siglos desde el desarrollo del cálculo diferencial, sabemos que una primitiva de x^n es $x^{n+1}/(n+1)$. Pero si imaginamos a alguien que no supiera cálculo diferencial, pero sí probabilidades, podríamos darle el siguiente argumento: el integrando se puede interpretar como la probabilidad de que, dadas n+1 v.a. U[0, 1] independientes, las n primeras estén en [0, u] (la probabilidad de que una v.a. U[0, 1] esté en [0, u] vale lógicamente u, la probabilidad de que las n primeras lo estén vale por tanto u^n) y la última en du. Por tanto la integral, como suma límite de todas esas probabilidades elementales (tendríamos que explicarle lo que es la integral), representa la probabilidad de que dadas las mismas n+1 v.a., la última sea la mayor de todas. De nuevo por simetría, la probabilidad de que la última sea la mayor de todas es la misma que la probabilidad de que la primera sea la mayor, igual a su vez a la probabilidad de que la segunda sea la mayor, etc. Por tanto como hay n+1 posibilidades para ser la mayor, la probabilidad de que la última lo sea es 1/(n+1). Haciendo el cambio de variable u = t /a en la integral, con a un real cualquiera, tenemos una demostración para nuestro interlocutor que no sabe cálculo diferencial, de que una primitiva de x^n es $x^{n+1}/(n+1)$ sin haber usado el álgebra (por ejemplo con el binomio de Newton, que supongo es en lo que se basó Newton para hallar la derivada de x^n) y el teorema fundamental del cálculo.

Pero sigamos explorando las posibilidades de la probabilidad en este tipo de integrales, y planteemos el cálculo de la constante normalizadora de la función de distribución Beta de la estadística. Es decir, intentemos calcular

$$I_2 = \int_0^1 u^{a-1} \, (1-u)^{b-1} \; du$$

con a y b positivos. Si restringimos a y b a enteros positivos la integral se puede hallar integrando por partes sucesivamente hasta reducir uno de los exponentes de u o de (1-u) hasta 0. El resultado es $[(a+b-1)!/(a-1)!(b-1)!]^{-1}$.

En cambio la vía probabilística nos ahorra engorrosas integraciones por partes con aplicación de la inducción. Razonamos que, dadas nuestras a+b-1 v.a. U[0, 1] independientes entre sí, el integrando es la probabilidad de que las primeras a-1 caigan en [0, u], la siguiente en du, y las b-1 últimas en [u, 1]. Por tanto la integral I_2 es la probabilidad de que dadas las a + b -1 v.a. U[0, 1] las a- 1 primeras estén a la izquierda de la a-ésima, y las b-1 últimas estén a la derecha de la a-ésima. De nuevo aplicando la simetría en la ordenación relativa de las v.a. vemos que la probabilidad buscada es la misma que la del problema siguiente: dadas a+b-1 bolas idénticas en un urna, con a-1 blancas, una roja, y b-1 negras, hallar la probabilidad de que en una extracción aleatoria las bolas blancas salgan primero, después la roja, y por último las negras. Las posibilidades en la ordenación de las a+b-1 bolas serán las permutaciones con repetición de a+b-1 bolas de las que se repiten a-1 y b-1, es decir, (a+b-1)!/(a-1)!(b-1)!. Solo una ordenación es la buscada, su probabilidad es por tanto el inverso de la cantidad anterior, que coincide con I_2[1] . Aunque los razonamientos son transparentes, a mi no me deja de causar admiración que una integral como I_2 sea lo mismo que una probabilidad relativa a un determinado orden de extracción de bolas idénticas de una urna.

Una generalización de la anterior integral, que se presta a un enfoque probabilístico rápido, es :

$$I_3 = \iint\limits_{0<u_1<u_2<1} u_1^{\,a} (u_2 - u_1)^b (1-u_2)^c \, du_1 du_2$$

con a, b y c enteros positivos. Con la práctica de las anteriores integrales no es muy difícil reducir I_3 a un problema de "bolas": dadas a+b+c+2 bolas idénticas en una urna, con a blancas, b rojas, c negras, una azul y otra verde, hallar la probabilidad de extraer las bolas en el orden, primero las blancas, luego la azul, luego las rojas, luego la verde y por último las negras. Las posibilidades en el orden relativo son las permutaciones con repetición de a+b+c+2 elementos de los que se repiten a, b y c, en total (a+b+c)!/(a!b!c!). Por tanto la probabilidad buscada, igual a I_3, es el inverso del número anterior.

El último ejemplo de integral con correlato probabilístico lo he encontrado este verano leyendo un libro de estadística, *The Statistical Analysis of Failure Data de J.D. Kalbfleisch y R.L. Prentice, 2ª ed., Wiley Series in Probability and Statistics*. En el capítulo 7 aparece una integral que, tras un cambio de variable, se reduce a

$$I_4 = \iiint\limits_{0<u_1<u_2<\ldots<u_k<1} \Pi_{i=1}^{k} \, du_i \; (1-u_i)^{m_i}$$

Los autores dicen que, por integración directa, el valor de la integral es

$$\Pi_{i=1}^{k} n_i^{-1} \qquad \text{con} \qquad n_i = \sum_{j=i}^{k} (1+m_j)$$

Propongo al lector que aplicando los esquemas anteriores transforme la integral en un problema de orden de extracción de bolas idénticas adecuado y pruebe la solución (los m_j supuestas enteras positivas)[2].

[1] Es interesante que la transformación de la integral I_2 al hacer una integración por partes también se puede probar con razonamientos combinatorios. Notando a I_2 como Q (a-1, b-1) para hacer aparecer explícitamente los exponentes, una integración por partes conduce a la igualdad Q (a-1, b-1) = (b-1)/a Q(a, b-2). Con la analogía probabilista expuesta esto equivale a establecer una relación análoga entre el número de ordenaciones con a-1 bolas blancas y b-1 negras y el número de ordenaciones con a blancas y b-2 negras. Consideremos una ordenación cualquiera con (a-1, b-1) blancas/negras, podemos convertir una negra en blanca de b-1 formas, de esa forma generando todas las ordenaciones posibles con (a, b-2) blancas/negras, pero cada ordenación con (a, b-2) puede venir de a ordenaciones con (a-1,b-1), por tanto los números de ordenaciones están en la relación
N(a ,b-2) = (b-1)/a N(a-1, b-1), que equivale a la relación entre las integrales anterior (recordar que la integral es el inverso del nº de ordenaciones)

[2] Solución: el problema se reduce al siguiente: Dadas n1 bolas idénticas, con k de ellas marcadas con los números del 1 al k, m_1 de ellas marcadas M1, m_2 de ellas marcadas M2, …, y m_k de ellas marcadas Mk, extraerlas en un orden determinado por las condiciones:
- C1: primero la bola 1
- C2: la bola 2 antes de las bolas 3,..,k y las M2, M3, .., Mk (n_2 -1 en total)
- C3: la bola 3 antes de las bolas 4,..,k y las M3, M4,…,Mk (n_3 -1en total)
- ….
- Ck: La bola k antes de las m_k bolas Mk

(Nótese que las bolas M1 no aparecen explícitamente en las condiciones porque la condición C1 implica que las bolas M1 se extraen después de la 1).
La probabilidad de obtener una ordenación con esas condiciones la calculamos por medio de :
pr(C1) = $1/n_1$
pr(C2|C1) = $1/n_2$
pr(C3|C1∩C2) = $1/n_3$
….
pr(Ck|C1∩C2∩…∩Ck-1)=1/nk Finalmente la probabilidad de C1∩C2∩….∩Ck es el producto de las anteriores.