

## 4. Modelos probabilísticos

### 4.1. Una central

El siguiente modelo se ha mencionado seguramente en Probabilidad I y se ve con detalle en Probabilidad II, pero es tan importante e interesante que no está de más insistir aquí. Con él podremos explicar por qué la normal de Gauss es una distribución central en las aplicaciones prácticas de la Estadística.

La situación que vamos a analizar es bastante genérica: digamos que hacemos un experimento en el que obtenemos una medición que en condiciones ideales es exacta, pero que en la práctica está alterada por una cantidad ingente de errores incontrolables. No hace falta ponerse muy serios, vestirse con bata blanca y pensar en un laboratorio lleno de matraces, basta imaginar algo tan lúdico como tirar dardo a una diana e intentar atinar en el centro. Aunque siempre apuntemos hacia él, con suerte lo más que conseguimos es quedarnos más o menos cerca. Si para fijar ideas y no complicarnos la vida desde el principio, suponemos vivir en planilandia o en un relieve del antiguo Egipto, de modo que la superficie de la diana sea unidimensional; se puede considerar que la desviación viene dada por un número real que indica la distancia con signo al origen (arriba +, abajo -). Ése es el número que medimos.

¿Por qué dicho número no es cero?, o lo que es lo mismo, ¿por qué no acertamos? Le podemos echar la culpa a muchas cosas y estaremos en lo cierto. Aquí y en otros experimentos es natural suponer que hay un montón de factores, digamos  $N$  con  $N \rightarrow \infty$ , que contribuyen a provocar una desviación infinitesimal. No podemos explicar biológica ni físicamente cómo nos va a fallar el pulso y es más fácil echarle la culpa a muchos factores biológicos o físicos que no sabemos controlar. Matemáticamente los representaremos por variables aleatorias  $\xi_1, \xi_2, \dots, \xi_N$  que supondremos uniformemente distribuidas en  $[-\epsilon, \epsilon]$  con  $\epsilon$  muy pequeño (más adelante veremos que esta hipótesis se puede rebajar mucho).

También supondremos que todos estos factores no están relacionados, es decir, las  $\xi_i$  son variables aleatorias independientes. Nuestra intención es dar en el blanco, pero por la influencia de todas las perturbaciones (la “mala suerte”), los errores se acumulan haciendo que la posición final en la diana del dardo que hemos lanzado sea la variable aleatoria  $\xi = \xi_1 + \xi_2 + \dots + \xi_N$  de esperanza nula (es igual de fácil desviarse arriba o abajo) y varianza  $\sigma^2$  que podemos hallar experimentalmente.

Diccionario:

- Factores de error (muchos)  $\longrightarrow$  Variables aleatorias  $\xi_1, \xi_2, \dots, \xi_N$  con  $N \rightarrow \infty$ ,  
y  $\xi_i \sim U([- \epsilon, \epsilon])$ .
- No relacionados  $\longrightarrow \xi_1, \xi_2, \dots, \xi_N$  independientes.
- Desviación = suma de errores  $\longrightarrow \xi = \xi_1 + \xi_2 + \dots + \xi_N$  con  $E(\xi) = 0$ ,  $V(\xi) = \sigma^2$ .

Ahora podemos plantear el problema de encontrar la distribución de  $\xi$  en estas condiciones cuando  $N \rightarrow \infty$ . Si tal distribución límite existe debe aparecer muchas veces en el *mundo real*, siempre que haya una acumulación de errores incontrolables como los indicados. Antes de seguir, nótese que por la fórmula para la varianza de la suma de variables aleatorias independientes

$$\sigma^2 = \sigma_1^2 + \dots + \sigma_N^2 = \frac{N}{2\epsilon} \int_{-\epsilon}^{\epsilon} x^2 dx = \frac{1}{3} N \epsilon^2.$$

De modo que dados  $\sigma$  y  $N$  el valor de  $\epsilon$  queda unívocamente determinado por  $\epsilon = \sigma \sqrt{3/N}$ .

Si llamamos  $f$  a la función de densidad de  $\xi_i$ , es decir,  $f(x) = (2\epsilon)^{-1}$  si  $|x| \leq \epsilon$ , y  $f(x) = 0$  en otro caso; la función de densidad de  $\xi_1 + \xi_2$  será para cada  $x$  la integral de  $f(x_1)f(x_2)$  sobre los valores que verifican  $x_1 + x_2 = x$ . Esto es,

$$g(x) = \int_{-\infty}^{\infty} f(x_1)f(x - x_1) dx_1 = (f * f)(x).$$

La última igualdad es sólo notación. Es decir, en general se define la *convolución*  $a * b$  de  $a = a(x)$  y  $b = b(x)$  como  $\int_{-\infty}^{\infty} a(t)b(x - t) dt$ . Análogamente, la función de densidad de  $\xi = \xi_1 + \xi_2 + \dots + \xi_N$  vendrá dada por la fórmula:

$$g_N(x) = (f * \overset{N \text{ veces}}{\dots} * f)(x).$$

Lo que queremos ver es que  $g_N$  se acerca a “algo”. Eso es lo que indica el siguiente

teorema, que es una versión en pequeño del teorema central del límite [Fe] VIII.4.

**Teorema 4.1.** Sea  $\sigma > 0$  y  $f$  como antes con  $\epsilon = \sigma\sqrt{3/N}$ , entonces

$$\lim_{N \rightarrow \infty} g_N(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}.$$

Aquí vemos aparecer mágicamente la famosísima campana de Gauss, que es familiar para cualquiera que haya seguido un curso básico de Estadística (en Alemania ni siquiera se exigía este requisito, porque la campana, su ecuación y su autor, aparecían en los antiguos billetes de diez marcos).

DEM.: De la definición de la transformada de Fourier, es fácil deducir que  $\widehat{a * b} = \widehat{a} \widehat{b}$ , y un cálculo muestra  $\widehat{f}(t) = \text{sen}(2\pi\epsilon t)/(2\pi\epsilon t)$ . Por tanto

$$g_N = f * \overset{N \text{ veces}}{\dots} * f \Rightarrow \widehat{g}_N(t) = (\widehat{f}(t))^N = \frac{\text{sen}^N(2\pi\epsilon t)}{(2\pi\epsilon t)^N}.$$

Como  $g_N$  es de soporte compacto y suficientemente regular cuando  $N$  es grande, se puede aplicar la fórmula de inversión, obteniéndose

$$g_N(x) = \int_{-\infty}^{\infty} \frac{\text{sen}^N(2\pi\epsilon t)}{(2\pi\epsilon t)^N} e(tx) dt = I_1 + I_2,$$

donde  $I_1$  es el valor de la integral sobre  $A = [-(2\pi\epsilon)^{-1}, (2\pi\epsilon)^{-1}]$  e  $I_2$  sobre su complementario  $\mathbb{R} - A$ . Obviamente  $|I_2| \leq 2 \int_{(2\pi\epsilon)^{-1}}^{\infty} (2\pi\epsilon t)^{-N} dt = (\pi(N-1)\sigma)^{-1} \sqrt{N/3}$  que tiende a cero con  $N$ . De hecho se podría comprobar que el decaimiento de  $|I_2|$  es exponencial. Sea  $h_N(u) = (u - u^3/6)^{-N} \text{sen}^N u$ , entonces

$$I_1 = \int_A (1 - (2\pi\epsilon t)^2/6)^N h_N(2\pi\epsilon t) e(tx) dt.$$

Sustituyendo  $\epsilon = \sigma\sqrt{3/N}$ , se tiene que para cada  $t$  fijado

$$\left(1 - \frac{(2\pi\epsilon t)^2}{6}\right)^N = \left(1 - \frac{2\pi^2\sigma^2 t^2}{N}\right)^N \xrightarrow{N \rightarrow \infty} e^{-2\pi^2\sigma^2 t^2}.$$

Por otra parte  $|h_N| < 1$  y no es difícil ver que por ejemplo para  $|t| < N^{0.1}$ ,  $h_N(2\pi\epsilon t) \rightarrow 1$ . Usando los teoremas habituales (convergencia dominada, convergencia uniforme), queda justificado introducir en  $I_1$  el límite bajo el signo integral, obteniéndose [Gr-Ry] 17.23.13

$$\lim_{N \rightarrow \infty} I_1 = \int_{-\infty}^{\infty} e^{-2\pi^2\sigma^2 t^2} e(tx) dt = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}.$$

Como se quería demostrar. ■

La consecuencia de este resultado es que típicamente en la práctica el error viene dado por una variable aleatoria  $\xi$  de media cero y desviación típica  $\sigma$ , de forma que

$$\text{Prob}(\xi > \sigma X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\sigma X}^{\infty} e^{-t^2/(2\sigma^2)} dt = \frac{1}{\sqrt{2\pi}} \int_X^{\infty} e^{-t^2/2} dt.$$

Esta última función, llamada  $\text{Erfc}(X)$ , se puede aproximar de diferentes formas y aparece en las tablas estadísticas. Algunos valores (redondeados hasta cuatro decimales) útiles para los ejercicios de la sección son  $\text{Erfc}(0'1) = 0'4602$ ,  $\text{Erfc}(0'4630) = 0'3217$ ,  $\text{Erfc}(0'6481) = 0'2585$ ,  $\text{Erfc}(1) = 0'1587$ .

Epílogo: Dadas  $N$  variables aleatorias  $\eta_i$  equidistribuidas e independientes de media  $\bar{\eta}$  y varianza  $\sigma$ , se puede considerar que  $\xi_i = (\eta_i - \bar{\eta})/(\sigma\sqrt{N})$  son pequeños errores, además  $\xi = \xi_1 + \dots + \xi_N$  cumple  $E(\xi) = 0$  y  $V(\xi) = 1$ . Si confiamos que el teorema central del límite es aplicable, para cada  $X$  se tiene

$$(4.1) \quad \lim_{N \rightarrow \infty} \text{Prob}((\eta_1 + \eta_2 + \dots + \eta_N - N\bar{\eta})/(\sigma\sqrt{N}) > X) = \frac{1}{\sqrt{2\pi}} \int_X^{\infty} e^{-t^2/2} dt = \text{Erfc}(X).$$

En el caso de variables continuas, sea cual sea la función de densidad  $f$  de  $\xi_i$ ,  $E(\xi_i) = \int f = 0$  y  $V(\xi_i) = \int x^2 f = 1/N$  implican por Taylor que  $\hat{f}(t) \approx 1 - (2\pi t/\sqrt{N})^2$  para  $t$  “pequeño”. Y esto, con una leve condición de regularidad para que  $I_2 \rightarrow 0$ , permite copiar la demostración del teorema y dar una prueba rigurosa de (4.1). En definitiva, una vez normalizados los errores infinitesimales, da igual la distribución que tengan, siempre se llega a una normal [**Ze-Ra-So**]. En el caso de variables discretas, (4.1) sigue siendo cierto, pero hay que tratar con funciones de distribución o con probabilidades en vez de con funciones de densidad, que en este caso no existen.

Como se ha indicado, para las variables continuas en principio se necesita un mínimo de regularidad que asegure el decaimiento de la transformada de Fourier, pero esto es gratis con algunos trucos sucios del análisis [**Dy-Mc**] §2.7. Uno siempre puede huir de todos estos tecnicismos cayendo en otros que no requieren en absoluto la transformada de Fourier, como se muestra en [**Fe**] VIII.4.

Quien desee conocer con rigor muchas de las variantes que conducen a la ubicua campana de Gauss, que mire con cuidado todos los teoremas del curso de Probabilidad II que terminan diciendo “converge ... a una distribución normal” [**Fe**], [**Ko**].

## Ejercicios

1) Sin mirar la teoría: a) Hallar la varianza de  $\xi_1\xi_2 + \dots + \xi_N$  con  $\xi_i$  independientes y uniformemente distribuidas en  $[-\epsilon, \epsilon]$ . b) Indicar por qué se pide la independencia de las variables en el modelo estudiado.

2) En Estadística I se prueba que la suma de dos variables aleatorias independientes con una distribución normal también tiene una distribución normal. Explicar este hecho interpretando las normales como acumulación de errores infinitesimales independientes.

3) Se dice que una señal recibida tiene *ruido gaussiano* si la diferencia con la señal emitida se comporta como una variable aleatoria con distribución normal de media cero y varianza típicamente pequeña. Y se dice que tiene *ruido de sal y pimienta* (*salt & pepper noise*) si la diferencia es nula salvo en cierta proporción típicamente pequeña de puntos, en los que se comporta como una distribución uniforme en cierto intervalo no necesariamente pequeño. Explicar por qué el primer tipo de ruido es el habitual en las transmisiones analógicas y el segundo en las digitales.

4) Sean  $\eta_1, \eta_2, \dots, \eta_{10}$  las variables aleatorias que dan las sucesivas puntuaciones de un dado al lanzarlo diez veces. La probabilidad de que la suma de puntuaciones sea 32 se puede escribir evidentemente como  $\text{Prob}(31'5 \leq \eta_1 + \dots + \eta_{10} < 32'5)$ . A partir de esta expresión y aplicando el teorema central del límite a  $\xi_1 + \dots + \xi_{10}$  con  $\xi_i = (\eta_i - 3'5)\sqrt{6/175}$ , aproximar dicha probabilidad (y compararla con el valor exacto 3801535/60466176).

5) Aproximar la probabilidad de que al tirar una moneda un millón de veces, la diferencia entre el número de caras y de cruces sea mayor que mil.

6) Si al tirar dardos en una diana apuntando al centro, la desviación típica es  $\sigma = 5 \text{ cm}$ , calcular la probabilidad de acertar en el círculo central que tiene un radio de  $0'5 \text{ cm}$ .

7) Con los datos del problema anterior hallar la probabilidad de acertar al menos una vez tras cinco intentos.

8) Dada  $f(x) = e^{x^{-2}} \int_0^x u^{-2} e^{-u^{-2}} du \in C^\infty$ , hallar su polinomio de Taylor de grado tres alrededor de cero (puede ser útil notar que  $x^3 f'(x) + 2f(x) = x$ ). Con el cambio  $t = \sqrt{2}/u$ , deducir la aproximación  $\text{Erfc}(x) \approx e^{-x^2/2}(x^{-1} - x^{-3})/\sqrt{2\pi}$  para  $x$  grande.

9) Demostrar que si  $f$  y  $g$  son suficientemente regulares  $\widehat{f * g} = \widehat{f} \widehat{g}$ .

10) Sea  $f$  la función que vale uno en  $[-1/2, 1/2]$  y cero en el resto. Hallar explícitamente  $f * f * f$  como función definida a trozos, y comparar su gráfica con la de  $e^{-2x^2} \sqrt{2/\pi}$  explicando la similitud.

11) Hallar la función de densidad en coordenadas cartesianas  $x$  e  $y$  cuando se apunta al centro de una diana ilimitada (que se supone  $\mathbb{R}^2$ ) con varianza  $\sigma^2 = 1/2$ . Esto es, la función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  tal que la probabilidad de que un dardo caiga en  $A \subset \mathbb{R}^2$  sea  $\int_A f(x, y) dx dy$ .

## SECCIÓN 4.1

### **Trabajos sugeridos a sólo 3 centauros** (el precio de esta fotocopia)

#### De la sección:

- El teorema ergódico y sus aplicaciones.
- Interpretación de Copenhague de la Mecánica Cuántica.

#### Generales:

- Teoría de Juegos y sus aplicaciones.

### **Y por el mismo precio, las palabras del sabio:**

*Aun más, cuando realizo una experiencia debo hacer algunas correcciones en el resultado, porque sé que he debido cometer errores. Estos errores son de dos clases: unos son accidentales y los corregiré tomando el valor medio, otros son sistemáticos y no podré corregirlos más que por un estudio profundo de sus causas. [Po] p. 143.*

## 4.2. Simple visitante

Un aburrido domingo salimos de casa, y en nuestra indecisión lanzamos una moneda para saber si vamos hacia la derecha o hacia la izquierda. Después del primer paso repetimos el procedimiento, y así sucesivamente. La pregunta que se plantea es si es probable que en un largo paseo aleatorio de estas características volvamos muchas veces a casa. Recordando a nuestra hermana menor, la Modelización I, esto es algo muy parecido a las cadenas de Markov allí estudiadas, pero la gran diferencia es que esta vez los pasos se dan sobre un conjunto discreto infinito, digamos  $\mathbb{Z}$  (puede que el mundo sea redondo, pero nos moriremos de viejos, de dolor de pies, o nos ahogaremos, antes de que demos la vuelta completa; por lo que es natural considerar un conjunto infinito). Se puede representar convencionalmente cada paso a la derecha con un signo “+” y cada paso a la izquierda con un signo “-”, lo que corresponde a sumar o sustraer una unidad en  $\mathbb{Z}$ . De este modo, un paseo es un conjunto ordenado de mases y menos.

En dos dimensiones la situación es análoga, pero ahora hay que considerar  $\mathbb{Z}^2$  y podemos dar pasos en las direcciones norte  $N$ , sur  $S$ , este  $E$  y oeste  $O$ . De modo que un paseo queda representado como una tira con estos símbolos.

En general, en dimensión  $D$  se consideran  $2D$  símbolos indicando los  $2D$  posibles sentidos en  $\mathbb{Z}^D$  y las listas formadas con ellos corresponden a paseos aleatorios en  $\mathbb{Z}^D$ .

Sea  $N_n$  el número de paseos aleatorios de  $n$  pasos que terminan en el punto de partida (en el origen, en casa). Evidentemente hay  $(2D)^M$  posibles paseos de  $M$  pasos en  $\mathbb{Z}^D$  y  $N_n(2D)^{M-n}$  de ellos pasarán por casa exactamente en el paso  $n$ -ésimo, de modo que el número medio de visitas a casa de un camino aleatorio de  $M$  pasos es

$$(2D)^{-M} (N_1(2D)^{M-1} + N_2(2D)^{M-2} + \dots + N_M(2D)^0) = \sum_{n=1}^M (2D)^{-n} N_n.$$

Para  $D = 1$ ,  $N_n$  no es más que el número de listas ordenadas de longitud  $n$  e igual número de mases que de menos; análogamente para  $D = 2$ , debe haber igual número de enes que de eses y de es que de oes, para así poder acabar en el origen. En general  $N_n = 0$  si  $n$  es impar, y con el lenguaje de la combinatoria  $N_n$  no es más que el número total de permutaciones con repetición de  $2D$  símbolos tomados de  $n$  en  $n$ , de manera que el número de repeticiones de los símbolos sea igual por parejas.

Diccionario:

- Paseos aleatorios en  $\mathbb{Z}^D \longrightarrow$  Listas ordenadas formadas con  $2D$  símbolos.
- Número de paseos que después de  $n$  pasos  $\longrightarrow$  Número total de permutaciones con repetición de  $2D$  símbolos tomados de  $n$  vuelven a casa. repetición de  $2D$  símbolos tomados de  $n$  en  $n$  con repeticiones iguales dos a dos.
- Número medio de vueltas a casa  $\longrightarrow \sum_{n=1}^{\infty} (2D)^{-n} N_n$ .

La fórmula para las permutaciones con repetición de  $2D$  símbolos con  $r_1 + r_2 + \dots + r_{2D} = n$  repeticiones es

$$PR_{r_1, r_2, \dots, r_{2D}}^n = \frac{n!}{r_1! r_2! \dots r_{2D}!}.$$

(Hay  $n!$  formas de permutar  $n$  elementos distintos, si  $r_1$  de los elementos son iguales se reducen a  $n!/r_1!$ , si otros  $r_2$  son iguales, a  $n!/(r_1!r_2!)$ , etc.). Por consiguiente

$$(4.2) \quad N_n = \sum_{2k_1+2k_2+\dots+2k_D=n} \frac{n!}{(k_1!)^2(k_2!)^2 \dots (k_D!)^2}.$$

El sorprendente resultado debido a Pólya es que  $\sum (2D)^{-n} N_n$  diverge si y sólo si  $D = 1, 2$ . Es decir, que para dimensión uno o dos, en media se vuelve a casa infinitas veces, mientras que en dimensiones mayores sólo un número finito. Un curioso efecto de la dimensión sobre los paseos aleatorios.

**Teorema 4.2.** *La serie  $\sum (2D)^{-n} N_n$  diverge para  $D = 1$  y  $D = 2$ , y converge para  $D \geq 3$ . De hecho en este caso se tiene la fórmula:*

$$\sum_{n=1}^{\infty} (2D)^{-n} N_n = -1 + \frac{D}{2\pi^D} \int_0^\pi \int_0^\pi \dots \int_0^\pi \left( \text{sen}^2 \frac{u_1}{2} + \text{sen}^2 \frac{u_2}{2} + \dots + \text{sen}^2 \frac{u_D}{2} \right)^{-1} du_1 \dots du_D.$$



En la demostración se necesitará evaluar una integral suficientemente sencilla e ingeniosa como para que hagamos el cálculo aparte.

**Lema 4.3.** *Sea  $m$  un entero no negativo, entonces*

$$\int_{-1}^1 \frac{x^m}{\sqrt{1-x^2}} dx = \begin{cases} 0 & \text{si } m \text{ es impar} \\ 2^{-2k} \pi (2k)! / (k!)^2 & \text{si } m = 2k \end{cases}$$

DEM.: Para  $m$  impar el resultado es trivial porque el integrando es una función impar. Si  $m = 2k$ , con el cambio  $x = \cos t = (e^{it} + e^{-it})/2$  se sigue

$$\int_{-1}^1 \frac{x^{2k}}{\sqrt{1-x^2}} dx = \frac{1}{2} \int_{-\pi}^{\pi} \cos^{2k} t dt = 2^{-2k-1} \sum_{j=0}^{2k} \binom{2k}{j} \int_{-\pi}^{\pi} e^{(2k-2j)it} dt.$$

Y basta notar que la última integral es nula excepto si  $j = k$ . ■

DEM.(del teorema): Al igual que los números combinatorios vienen generados por la potencia de un binomio, las permutaciones con repetición lo están por la de un multinomio. Concretamente

$$(x_1 + x_2 + \dots + x_D)^n = \sum_{m_1+m_2+\dots+m_D=n} \frac{n!}{m_1!m_2!\dots m_D!} x^{m_1} x^{m_2} \dots x^{m_D}.$$

La suma de los coeficientes tiene un aspecto similar a la fórmula (4.2) para  $N_n$ , pero a fin de que coincida exactamente hay que reemplazar cada  $m_i!$  por  $(k_i!)^2$  con  $m_i = 2k_i$ . Esto se consigue gracias al lema, que a través de (4.2) y la fórmula anterior, implica

$$N_n = \frac{2^n}{\pi^D} \int_{-1}^1 \int_{-1}^1 \dots \int_{-1}^1 \frac{(x_1 + x_2 + \dots + x_D)^n}{\sqrt{(1-x_1^2)(1-x_2^2)\dots(1-x_D^2)}} dx_1 dx_2 \dots dx_D.$$

Con el cambio de variable  $x_i = 1 - 2 \sin^2(u_i/2) = \cos u_i$ , se tiene

$$N_n = \frac{(2D)^n}{\pi^D} \int_0^\pi \int_0^\pi \dots \int_0^\pi \left(1 - \frac{2}{D} (\sin^2 \frac{u_1}{2} + \sin^2 \frac{u_2}{2} + \dots + \sin^2 \frac{u_D}{2})\right)^n du_1 du_2 \dots du_D.$$

Si se conviene que  $N_0 = 1$ , esta igualdad también es cierta para  $n = 0$ . De esta forma,  $\sum_{n=1}^\infty (2D)^{-n} N_n = -1 + \sum_{n=0}^\infty (2D)^{-n} N_n$ . Sustituyendo  $N_n$  por la fórmula integral ante-

rior y empleando que  $1 + r + r^2 + r^3 + \dots = 1/(1 - r)$  para  $|r| < 1$ , se tiene

$$-1 + \frac{D}{2\pi^D} \int_0^\pi \int_0^\pi \cdots \int_0^\pi \left( \operatorname{sen}^2 \frac{u_1}{2} + \operatorname{sen}^2 \frac{u_2}{2} + \dots + \operatorname{sen}^2 \frac{u_D}{2} \right)^{-1} du_1 du_2 \dots du_D.$$

Cuando los  $u_i$  tienden a cero,  $(\operatorname{sen}^2 \frac{u_1}{2} + \operatorname{sen}^2 \frac{u_2}{2} + \dots + \operatorname{sen}^2 \frac{u_D}{2}) / (u_1^2 + u_2^2 + \dots + u_D^2) \rightarrow 1/2$ , de modo que la integral converge si y sólo si  $\int_B \|\vec{x}\|^{-2} < \infty$  con  $B$  un entorno del origen, por ejemplo la bola unidad en  $\mathbb{R}^D$ . Pasando a esféricas (generalizadas), esta integral es, salvo un factor constante,  $\int_0^1 r^{-2} \cdot r^{D-1} dr$ , que claramente converge para  $D > 2$  y diverge en otro caso. ■

## Ejercicios

1) Sin mirar la teoría: a) Explicar por qué en  $\mathbb{Z}^D$  hay  $(2D)^M$  paseos de longitud  $M$ .  
b) Explicar por qué hay  $N_n(2D)^{M-n}$  paseos en  $\mathbb{Z}^D$  que vuelven a casa en el paso  $n$ -ésimo.

2) Hallar la probabilidad de estar a diez pasos de distancia de casa después de haber dado 20. ¿Cuál es la respuesta si se dan 21?

3) En el caso  $D = 1$  escribir el número medio de vueltas a casa como una serie que involucre números combinatorios. Sabiendo que  $\lim_{n \rightarrow \infty} 2^{-2n} \sqrt{\pi n} \binom{2n}{n} = 1$ , demostrar la divergencia de la serie.

4) Repetir el problema anterior si la probabilidad de dar un paso a la derecha es  $p > 0.5$ , pero demostrando ahora la convergencia de la serie. Indicar por qué este resultado es natural.

5) Calcular la varianza de la variable aleatoria que indica la posición tras  $n$  pasos. Concluir que es muy raro llegar a una distancia mucho mayor que  $\sqrt{n}$ .

6) Si  $B_R$  es el número de puntos de  $\mathbb{Z}^D$  en la bola de radio  $R$ , probar que  $\lim_{R \rightarrow \infty} n/B_{\sqrt{n}} = 0$  si y sólo si  $D > 2$ . A partir del problema anterior, explicar por qué es lógico que justamente para  $D > 2$  un camino a la larga no vuelva a visitar un punto y por tanto el número de regresos al origen se deba fundamentalmente a lo que ocurre con caminos cortos y por tanto sea finito.

7) Utilizando que  $x \geq \sin x$  en  $[0, \pi/2]$ , y que  $[0, \pi]^3$  incluye al primer octante de la bola de radio  $\pi$ , dar una cota inferior para la integral del teorema de esta sección cuando  $D = 3$ .

8) Generalizar el resultado del problema anterior para  $D > 3$ .

9) En [Dy-Mc], después de concluir que el número medio de vueltas a casa es finito para  $D \geq 3$ , se afirma: “Como el origen no es de ningún modo especial, lo mismo debe ocurrir para cualquier punto de  $\mathbb{Z}^D$ . Pero esto significa que para cualquier  $R < \infty$  la partícula [el paseante] acaba dejando de visitar la bola  $\|\vec{x}\| < R$ , y esto es lo mismo que decir  $\text{Prob}(\lim_{n \rightarrow \infty} |s_n| = \infty) = 1$  [donde  $s_n$  es la posición tras  $n$  pasos]”. Explicar este argumento con todo el rigor que sea posible.

10) A partir de la varianza de la posición, indicar intuitivamente, en el caso  $D = 1$ , por qué típicamente cada vez se tarda más en volver al origen. Dar con ello una explicación de las rachas de mala o buena suerte que mencionan muchos jugadores.

11) Supongamos un circuito en forma de polígono regular tal que en cada vértice hay probabilidades no nulas  $p$  y  $1 - p$  de ir a la derecha y a la izquierda (no necesariamente las mismas en diferentes vértices). ¿Es siempre infinito el número medio de retornos al punto de partida?

## SECCIÓN 4.2

### **Trabajos sugeridos a sólo 3 centavos** (el precio de esta fotocopia)

#### De la sección:

- Teoría de Colas y sus aplicaciones.
- Modelos del tráfico.

#### Generales:

- Generación de números aleatorios.

#### **Y por el mismo precio, las palabras del sabio:**

*El demonio imaginario de Maxwell, que puede entresacar las moléculas una a una, bien podría constreñir al mundo a volverse atrás. ¿Puede volver allí por sí mismo? Esto no es imposible, no es más que infinitamente poco probable; hay probabilidades de que deberíamos esperar mucho tiempo el concurso de las circunstancias que permitieran el retroceso, pero tarde o temprano ellas se realizarán después de tantos años que para escribir su número serían menester millones de cifras. [Po] p. 119.*

### 4.3. Vienen o van

En 1827 el botánico R. Brown observó el *movimiento browniano* consistente en que pequeñas partículas de polen suspendidas en una disolución se trasladan siguiendo caminos caóticos. Sus contemporáneos (y en parte él mismo) pensaron que esto era un signo de vida primaria, pero más tarde el desarrollo de la teoría atómica probó que representaba los empujones que dan las moléculas a las partículas de polen en direcciones aleatorias. De la misma forma, podemos estudiar los fenómenos de difusión de un gas ocultando nuestro desconocimiento submicroscópico diciendo que las moléculas que lo conforman se mueven totalmente al azar, ya que las colisiones entre ellas las hacen cambiar continuamente de dirección.

Para simplificar vamos a restringirnos al caso unidimensional, esto es, como si las partículas de un gas estuvieran metidas en un tubo largo y delgado y sólo pudieran ir a la derecha o a la izquierda. Si hiciéramos fotos de las partículas cada  $h$  segundos durante cierto periodo de tiempo, sólo ocuparían un conjunto discreto de valores (las veríamos saltar a trompicones, como ocurre con las luces estroboscópicas en las discotecas). Por ello no es descabellado suponer que cada una describe un paseo aleatorio en  $\epsilon\mathbb{Z}$  donde  $\epsilon > 0$  es un número muy pequeño. En los instantes  $0, h, 2h, 3h$ , etc. cada partícula puede trasladarse  $\epsilon$  unidades (una casilla) a la derecha o hacia la izquierda con la misma probabilidad (el 50%).

Fijado un tiempo  $t_k = kh$ ,  $k \in \mathbb{Z}^+ \cup \{0\}$ , habrá cierta densidad (porcentaje) de partículas  $p(x_n, t_k)$  en el punto  $x_n = n\epsilon$ ,  $n \in \mathbb{Z}$ . Esto es,

$$p(x_n, t_k) = \frac{\text{número de partículas en } x_n}{\text{número total de partículas}}.$$

Se puede entender  $p$  como una probabilidad (la de encontrar una partícula en  $x_n$ ) y evidentemente  $\sum_n p(x_n, t_k) = 1$ , lo que manifiesta la conservación del número de partículas.

El problema que consideramos es predecir la evolución de esta densidad o probabilidad suponiéndola conocida sólo en el tiempo inicial  $t_0 = 0$ . A modo de ilustración, imaginemos que se dejan de golpe en el origen de coordenadas un montón de hormigas rastreadoras. Cada una de ellas seguirá un camino aleatorio y aunque no seamos capaces de saber dónde

estará al cabo de un rato la hormiga  $j$ -ésima, desde lejos veremos una mancha negra que se expande. Lo que queremos es capturar esa idea y deducir el comportamiento a gran escala a partir de la distribución inicial, sin importarnos las partículas u hormigas individuales, haciendo, como se dice en [Va 2], “predicciones sobre lo no exacto”.

Con esta idea, debemos considerar una cantidad innumerable de partículas y, para que el modelo represente fenómenos reales, la densidad de probabilidad  $p/\epsilon$ , debe acercarse a una función suave al pasar al caso continuo en el que  $\epsilon$  y  $h$  tienden a cero (de forma adecuada). Es posible concretar más la forma en la que  $\epsilon$  y  $h$  se deben hacer pequeños para que el modelo discreto tienda a uno continuo con sentido. Considerando la variable aleatoria que da la posición de una partícula que parte del origen en un paseo aleatorio de  $k$  pasos en  $\epsilon\mathbb{Z}$ , su desviación típica es  $\epsilon\sqrt{k}$ . Esto implica que típicamente en  $kh = 1$  segundo una partícula se ha desplazado  $\epsilon\sqrt{k} = \epsilon h^{-1/2}$  metros de su posición inicial. Si no queremos que el conjunto de partículas “explote” o que permanezca inmóvil, deberemos hacer que esta velocidad media en el primer segundo,  $\epsilon h^{-1/2}$ , sea una constante positiva. Volviendo a las hormigas, puede haber unas que se alejen más por seguir caminos más rectos y otras que se alejen menos, pero no queremos que en promedio toda la nube avance infinito o cero en el primer segundo (aunque no descartamos que alguna lo haga), sino cierta cantidad positiva.

Parece muy complicado controlar una infinidad de paseos aleatorios, sin embargo hay una ecuación muy sencilla que regula la evolución de  $p$ , simplemente conviniendo que cada partícula que llega a  $x_n$  en  $t = t_{k+1}$  tiene un 50% de posibilidades de provenir de  $x_{n-1}$  (de la izquierda) en el tiempo anterior  $t = t_k$ , y otro 50% de provenir de  $x_{n+1}$  (de la derecha). Es decir,

$$(4.3) \quad p(x_n, t_{k+1}) = \frac{1}{2}(p(x_{n-1}, t_k) + p(x_{n+1}, t_k)).$$

#### Diccionario:

- Posiciones posibles  $\longrightarrow x_n = n\epsilon, n \in \mathbb{Z}$ .
- Tiempos posibles  $\longrightarrow t_k = kh, k = 0, 1, 2, \dots$
- Concentración (densidad, probabilidad)  $\longrightarrow p(x_n, t_k) \geq 0$  con  $\sum_n p(x_n, t_k) = 1$ .
- Al encontramos con una partícula hay la misma probabilidad de que el instante previo estuviera a derecha o izquierda  $\longrightarrow p(x_n, t_{k+1}) = (p(x_{n-1}, t_k) + p(x_{n+1}, t_k))/2$ .
- Velocidad media finita  $\longrightarrow \epsilon h^{-1/2} = \text{cte}$ .
- Discreto  $\rightarrow$  continuo  $\longrightarrow p/\epsilon \rightarrow u =$  función suave.

La fórmula (4.3) es una relación de recurrencia que permite estudiar la evolución de nuestro modelo discretizado, sin embargo aspiramos a estudiar el límite cuando  $\epsilon \rightarrow 0$ . El razonamiento que vamos a hacer es realmente sencillo. Escribimos simplemente la relación (4.3) como

$$\frac{p(x_n, t_{k+1}) - p(x_n, t_k)}{h} = \frac{\epsilon^2}{2h} \frac{p(x_{n-1}, t_k) + p(x_{n+1}, t_k) - 2p(x_n, t_k)}{\epsilon^2}.$$

Usando el lenguaje de la asignatura de Cálculo Numérico II, éste es el método de diferencias finitas aplicado a  $\partial u / \partial t = \frac{1}{4} \alpha \partial^2 u / \partial x^2$  con  $\alpha = 2\epsilon^2 / h$ . Dicho de otra forma, el miembro izquierdo aproxima a la derivada respecto a la segunda variable cuando  $h \rightarrow 0$  mientras que en el segundo miembro aparece una derivada segunda. Por si esto último no se cubrió en el curso de Cálculo I, lo enunciamos a continuación:

**Lema 4.4.** Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Si  $f''(a)$  existe, entonces

$$f''(a) = \lim_{\epsilon \rightarrow 0} \frac{f(a + \epsilon) + f(a - \epsilon) - 2f(a)}{\epsilon^2}.$$

DEM.: Aplicando la regla de L'Hôpital,

$$\lim_{\epsilon \rightarrow 0} \frac{f(a + \epsilon) + f(a - \epsilon) - 2f(a)}{\epsilon^2} = \lim_{\epsilon \rightarrow 0} \frac{f'(a + \epsilon) - f'(a)}{2\epsilon} + \lim_{\epsilon \rightarrow 0} \frac{f'(a - \epsilon) - f'(a)}{-2\epsilon}.$$

Y basta aplicar la definición de derivada. ■

Retomando el argumento anterior, si  $\epsilon$  y  $h$  tienden a cero con  $\alpha = 2\epsilon^2/h$ , constante, y  $p/\epsilon$  tiende en un sentido apropiado a una función suave  $u$ , ésta debe cumplir  $\partial u / \partial t = \frac{1}{4} \alpha \partial^2 u / \partial x^2$ . Partiendo de una concentración inicial  $u(x, 0) = f(x)$  (una función de densidad suficientemente regular), para estudiar la posterior evolución del sistema hay que resolver la *ecuación del calor*

$$(4.4) \quad \begin{cases} \frac{\partial u}{\partial t} = \frac{\alpha}{4} \frac{\partial^2 u}{\partial x^2} & x \in \mathbb{R}, t > 0 \\ u(x, 0) = f(x) \end{cases}$$

En la sección correspondiente, ya vimos cómo emplear la transformada de Fourier para obtener la solución general de esta ecuación bajo hipótesis adecuadas de regularidad

sobre  $f$ . El coeficiente  $\alpha/4$  no cambia el aspecto de la solución, que es

$$u(x, t) = (\pi\alpha t)^{-1/2} \int_{-\infty}^{\infty} e^{-\alpha^{-1}t^{-1}(x-y)^2} f(y) dt.$$

El caso  $\alpha = 4$  corresponde a la solución de la ecuación del calor habitual,  $\partial u/\partial t = \partial^2 u/\partial x^2$ . Nótese que cuando  $t \rightarrow +\infty$  la función  $u$  tiende a cero, lo que indica que las partículas están cada vez distribuidas de manera más uniforme en  $\mathbb{R}$ . El aumento de la difusión hace que la densidad se aproxime puntualmente a cero. Esto deja de ser cierto si se plantea (4.4) en un dominio acotada en lugar de en  $\mathbb{R}$ , pero siempre la densidad tenderá a su valor promedio.

Epílogo: En cierto modo en la fórmula anterior para  $u(x, t)$  lo único que se hace es “sumar” (integrar) todas las campanas de Gauss correspondientes a aplicar el teorema central del límite a los paseos aleatorios de cada partícula (la interpretación de ciertas integrales similares a ésta como sumas sobre “todos los caminos aleatorios” [Ze-Ru-So] es muy importante en Física Cuántica [Yn] y no del todo fundamentada matemáticamente). Por simplicidad, aquí sólo hemos tratado el problema en una dimensión; pero el método se extiende a dimensiones superiores con fórmulas similares simplemente cambiando  $\partial^2 u/\partial x^2$  por  $\Delta u$ . Una “pega” que se puede poner al modelo en cualquier dimensión es que si  $f$  tiene soporte compacto,  $u(x, t)$  no lo tiene para ningún  $t > 0$  ya que  $u(x, t) > 0$ , lo que implica que todo funciona como si inicialmente las partículas viajasen arbitrariamente rápido. Aunque esto sea mecánicamente imposible (relatividad especial), el decaimiento exponencial de  $u$  cuando  $x \rightarrow \infty$  provoca que  $u$  sea prácticamente indistinguible de una función de soporte compacto. Sólo en condiciones extremas debemos modificar el modelo reemplazando la ecuación del calor por la *ecuación de los medios porosos*.



## Ejercicios

1) Sin mirar la teoría: a) Explicar el significado de la fórmula de recurrencia básica  $p(x_n, t_{k+1}) = (p(x_{n-1}, t_k) + p(x_{n+1}, t_k))/2$ . b) Indicar qué representa que la solución de la ecuación del calor tienda a cero cuando  $t \rightarrow +\infty$ .

2) Explicar por qué la ecuación del calor  $\partial u/\partial t = \frac{1}{4}\alpha \partial^2 u/\partial x^2$ ,  $x > 0$ ,  $t > 0$  con  $u(x, 0) = f(x)$  y  $u(0, t) = 0$  corresponde al caso en que el origen hay un “agujero” que absorbe las partículas.

3) Si en el problema anterior en  $x = 0$  hay una barrera que impide que las partículas pasen hacia la izquierda, tratar de justificar por qué se debe imponer  $\partial u/\partial t(0, t) = 0$  en lugar de  $u(0, t) = 0$ .

4) Traducir la relación  $\sum_n p(x_n, t_k) = 1$  para todo  $k$  en alguna ley de conservación para la ecuación del calor en  $\mathbb{R}$  (con  $f$  de decaimiento rápido) y demostrarla. Estudiar si tal ley se sigue cumpliendo para la ecuación del calor en  $[0, \infty)$  bajo las condiciones especificadas en los dos problemas anteriores.

5) Supongamos que cada partícula puede con igual probabilidad moverse a la derecha, a la izquierda, o quedarse inmóvil. Indicar los cambios en el modelo y estudiar si hay diferencias cuando se pasa al límite.

6) La concentración de partículas en los cuatro vértices de un cuadrado es del 12'5%, 12'5%, 37'5% y 37'5%. Calcular la concentración esperada después de tres unidades de tiempo, sabiendo que en cada una de ellas cada partícula se dirige aleatoriamente a uno de los dos vértices adyacentes.

7) Si  $f$  es continua de soporte compacto probar que  $u \in C^\infty$ . El proceso de reemplazar una señal  $f = f(x)$  por  $u(x, t)$  con  $t$  pequeño se emplea habitualmente para reducir ruidos. Explicar por qué. (En teoría de la señal se llama a esto un *filtro gaussiano*).

8) Escribir la ecuación de recurrencia para  $p$  si la probabilidad de una partícula de ir a la derecha y a la izquierda no coinciden. Argumentar por qué en este modelo se observa transporte (traslación) más que difusión.

9) Generalizar el modelo de esta sección al caso bidimensional escribiendo la relación de recurrencia para  $p$  y la ecuación límite.

10) Hallar una función  $f : \mathbb{R} \rightarrow \mathbb{R}$  que no tenga derivada segunda en cero pero tal que exista el límite  $\lim_{h \rightarrow 0} (f(h) + f(-h) - 2f(0))/h^2$ .

11) Sea  $f \in C^4(\mathbb{R}^2)$  y  $L_h(x, y) = \Delta f(x, y) - (f(x+h, y) + f(x-h, y) + f(x, y+h) + f(x, y-h) - 4f(x, y))/h^2$ . Utilizando la fórmula de Taylor, demostrar que  $\lim_{h \rightarrow 0} L_h/h^\alpha = 0$  para todo  $\alpha < 2$ .

SECCIÓN 4.3

## Trabajos sugeridos a sólo 3 centauros (el precio de esta fotocopia)

De la sección:

- Procesos de difusión en Matemática Financiera.
- Relación entre el movimiento browniano y el número de Avogadro (puede ser interesante indagar los errores teóricos y prácticos que llevaron a Einstein a deducir en su tesis que el número de Avogadro era aproximadamente  $2'1 \cdot 10^{23}$  mientras que el valor real es casi el triple).
- Mecánica Estadística.
- Ecuaciones diferenciales estocásticas.

Generales:

- Métodos matemáticos en Astrofísica.

### Y por el mismo precio, las palabras del sabio:

*No podemos prever en qué sentido vamos a extendernos; quizás sea la teoría cinética de los gases la que se desarrollará y servirá de modelo a las otras. Entonces, los hechos que primeramente aparecían como simples, no serán más que las resultantes de un número muy grande de hechos elementales que sólo las leyes del azar harían concurrir a un mismo fin. La ley física, por lo tanto, tomaría un aspecto completamente nuevo. Ya no sería solamente una ecuación diferencial; adquiriría el carácter de una ley estadística. [Po] p. 136.*

## Puede que poder pudiera

### Material:

- Diez dados.
- Un cubilete.
- Material para dibujar una gráfica.

Consideremos la variable aleatoria que asigna la cara de puntuación  $n$  de un dado el número  $(n - 3.5)/\sqrt{175/6}$ . Se comprueba con un cálculo que tiene esperanza nula y varianza 0.1. El teorema central del límite sugiere que si consideramos el lanzamiento de 10 dados y la suma  $S$  de sus puntuaciones, entonces  $(S - 35)/\sqrt{175/6}$  tiene aproximadamente una distribución  $N(0, 1)$ . Equivalentemente,  $S$  tiene aproximadamente una distribución  $N(35, \sqrt{175/6})$ . Es decir, cabe esperar

$$\text{Prob}(S = n) \approx \frac{1}{\sqrt{175\pi/3}} e^{-3(n-35)^2/175}.$$

La cantidad de formas en que se puede obtener suma igual a  $n$ ,  $10 \leq n \leq 60$  al lanzar 10 dados está recogida en la siguiente tabla:

<b>10</b> → 1	<b>11</b> → 10	<b>12</b> → 55	<b>13</b> → 220	<b>14</b> → 715
<b>15</b> → 2002	<b>16</b> → 4995	<b>17</b> → 11340	<b>18</b> → 23760	<b>19</b> → 46420
<b>20</b> → 85228	<b>21</b> → 147940	<b>22</b> → 243925	<b>23</b> → 383470	<b>24</b> → 576565
<b>25</b> → 831204	<b>26</b> → 1151370	<b>27</b> → 1535040	<b>28</b> → 1972630	<b>29</b> → 2446300
<b>30</b> → 2930455	<b>31</b> → 3393610	<b>32</b> → 3801535	<b>33</b> → 4121260	<b>34</b> → 4325310
<b>35</b> → 4395456	<b>36</b> → 4325310	<b>37</b> → 4121260	<b>38</b> → 3801535	<b>39</b> → 3393610
<b>40</b> → 2930455	<b>41</b> → 2446300	<b>42</b> → 1972630	<b>43</b> → 1535040	<b>44</b> → 1151370
<b>45</b> → 831204	<b>46</b> → 576565	<b>47</b> → 383470	<b>48</b> → 243925	<b>49</b> → 147940
<b>50</b> → 85228	<b>51</b> → 46420	<b>52</b> → 23760	<b>53</b> → 11340	<b>54</b> → 4995
<b>55</b> → 2002	<b>56</b> → 715	<b>57</b> → 220	<b>58</b> → 55	<b>59</b> → 10
<b>60</b> → 1				

La probabilidad de  $S = n$  es por tanto el número asignado a  $n$  dividido por el número de casos posibles  $6^{10} = 60\,466\,176$ . Con ello se comprueba que la bondad de la aproximación anterior es increíble teniendo en cuenta que sólo usamos  $N = 10$  dados mientras que la teoría nos habla de lo que ocurre cuando  $N \rightarrow \infty$ . Si representamos en una gráfica ambos

miembros de la aproximación, no es posible detectar diferencias a simple vista, salvo quizá en los tres puntos centrales donde el error relativo es menor que el 2%.

Una vez que hemos visto lo extraordinariamente bien que una normal aproxima a la distribución de la suma de las puntuaciones de 10 dados, el experimento consistirá en comprobar que si estimamos las probabilidades estadísticamente tirando nosotros mismos los dados, nos cansaremos antes de ver una campana de Gauss decente. La moraleja es que debemos creer ciegamente en la Estadística pero no siempre en las estadísticas.

Concretamente, el experimento es muy simple y consiste en lanzar los dados con el cubilete un número de veces grande  $A$ , hasta que nos aburramos, y apuntar en cada caso la suma. Al terminar, tras desperezarnos, compararemos las gráficas obtenidas al representar los puntos con abcisa  $10 \leq n \leq 60$  y ordenadas

$$\frac{\text{n}^{\circ} \text{ de veces en que la suma es } n}{A}, \quad \frac{1}{\sqrt{175\pi/3}} e^{-3(n-35)^2/175}.$$

Por ejemplo, en un experimento real\* con  $A = 100$  se obtuvo

El error en el punto central  $n = 35$  es de más del 30% y en el punto anterior  $n = 34$  de casi el 60%.

Explicación: En principio no hay ninguna contradicción: la aproximación es tan buena como antes sólo si  $A$  es suficientemente grande (ley de los grandes números [Fe]). La pregunta natural es por qué 100 o 200 (donde sólo habrán llegado los más pacientes) no es un número *suficientemente grande*. Evidentemente con un ordenador podríamos simular

---

\* *N. del A.* Como el experimento es un poco largo, más vale hacerlo con comodidad. Lancé los datos en un barreño para que no se desperdigaran. Después de cada tirada los llevaba hacia el borde alineándolos y copiaba las puntuaciones en una hoja de cálculo que efectuaba las sumas en mi lugar. Con ello también quise recopilar datos sobre las frecuencias para tratar de desmentir la queja típica cuando se juega al parchís de que existe “el dado de los seises”.

el lanzamiento de los dados un millón de veces y entonces el resultado sería bastante aproximado, pero hacer el experimento 100 o 200 veces de verdad, sin delegar en las tripas de un ordenador, conlleva tanto esfuerzo que es descorazonadora la pobreza de la aproximación.

Demos a nuestra pregunta una forma matemática un poco más concreta y calculemos por ejemplo de qué tamaño debe ser típicamente  $A$  para que el error en el punto central  $n = 35$  sea menor que el 10%. Para tal fin, considérese la variable aleatoria que al tirar los dados  $A$  veces cuenta el número de veces en que la suma es 35 (número de éxitos). Esta variable aleatoria claramente tiene una distribución binomial  $B(A, p)$  con  $p$  la probabilidad de obtener suma igual a 35. Según la tabla,  $p = 4395456/6^{10} \approx 0'074$ . La esperanza de esta binomial es  $pA$ , y la desviación típica  $\sqrt{p(1-p)A}$ , por tanto cuando hagamos el experimento  $A$  veces, lo normal es que en vez de obtener  $pA$  veces suma 35 la obtengamos  $pA + \text{error}$  veces con **error** una cantidad comparable a  $\sqrt{p(1-p)A}$ . Si queremos que el error relativo sea típicamente menor que el 10%, se debería cumplir

$$\sqrt{p(1-p)A} < \frac{10}{100}pA, \quad \text{o equivalentemente} \quad A > \frac{1-p}{(0'1)^2 p}.$$

Sustituyendo  $p$  por  $0'074$ , esto conduce a  $A > 1251$ .

El error cometido al efectuar nuestra estadística preguntando a muchos dados qué número se obtiene como suma, ha sido bastante burdo: simplemente no deberíamos haber preguntado a muchos, sino a muchísimos, a más de mil. Errores como éste no se producen en las estadísticas serias (que no son todas las que aparecen en los medios de comunicación), porque son de algún modo de naturaleza matemática. Aunque éstos pueden llegar a ser realmente sutiles [**Ju**], seguramente los errores más graves en las estadísticas y que posiblemente invalidan un número no desdeñable de ellas, están ligados a factores psicológicos. Por ejemplo, es muy fácil obtener un “no sabe/no contesta” o una mentira al preguntar sobre temas escabrosos. También la forma de las estadísticas está muchas veces influida por lo que se quiere demostrar o por los propios prejuicios. Por ejemplo, si la imagen  $I$  de un suceso trágico e impresionante ha aparecido muchas veces en televisión, las preguntas: “¿Cree usted que se ha emitido demasiadas veces  $I$ ?”, “¿Cree usted que se debería evitar la emisión de  $I$ ?” y “¿Cree usted que deberían prohibir emitir  $I$ ?”; arrojarían resultados desiguales. Si hiciéramos la primera pregunta la respuesta sería seguramente “sí”, pero si hiciéramos la segunda o la tercera, casi todos intentamos no involucrarnos en algo que sugiera escabullirse o prohibir, de modo que la respuesta tendería más al “no”. Las conclusiones que alguien sacara de los resultados podrían llegar a ser opuestas aunque

las preguntas no lo sean.

Los que hayan hecho la experiencia anterior con los dados, probablemente ya habrán notado una curiosa manifestación experimental de lo psicológico. Al tirar los 10 dados casi todas las veces parece que las puntuaciones obtenidas tienen algo de singular e improbable: hay muchos seises, hay varios dados seguidos con puntuación ascendente, casi todas las puntuaciones son menores que cuatro, etc. La mayoría de las veces pensamos que hemos tenido “buena” o “mala suerte”, sin saber reconocer lo rutinario.

## Todo por igual

### Material:

- Un montón de judías blancas crudas (sin cocinar).
- Un montón de judías pintas similares a las anteriores.
- Un programa para generar números aleatorios y dibujar gráficas (opcional).

¿Qué ocurre cuando dos empresas compiten lanzando al mercado productos similares e incompatibles? Estamos acostumbrados a ver que en esta situación (sistemas de vídeo, sistemas operativos de ordenadores), después de una pugna inicial con altibajos, la empresa que logra una ventaja significativa acaba con la otra, independientemente de la calidad del producto, ya que *el pez grande se come al chico*.

Lo que vamos a comprobar, gracias a un bello, interesante y sorprendente modelo conocido como *urna de Pólya*, es que el mundo matemático es menos violento y permite una coexistencia pacífica.

Metamos una judía de cada color en un bote. Éstas representarán los productos iniciales de cada empresa. No es descabellado suponer que los clientes eligen al azar entre los nuevos productos, por tanto si hay una desproporción en la oferta a favor de uno de ellos, lo elegirán más. Escojamos pues, una judía al azar del bote, y después de verla, repongámosla y añadamos otra judía del mismo color. Ahora habrá dos judías de un tipo y una de otro, con lo cual es más fácil escoger las primeras. Repitamos el procedimiento un número grande de veces\*.

Cabría esperar que una mayoría clara obtenida al azar, en unas cuantas iteraciones se convierte en aplastante. Pero el experimento nos muestra, prácticamente siempre, volcando el bote, que hay una proporción apreciable de la minoría que no sólo no tiende a desaparecer sino que parece estabilizarse.

---

\* *N. del A.* Repetí el proceso 200 veces anotando los resultados en una hoja de cálculo para poder representar la evolución del sistema.

Una simulación con ordenador nos muestra que éste es el caso.

Simulación

Ejemplo real

Explicación: Sea  $X_n$  la variable aleatoria que toma el valor 1 si en la  $n$ -ésima extracción la judía es blanca y 0 si es pinta. La propiedad importante de estas variables aleatorias es que aunque no son independientes, son *intercambiables*. Esto quiere decir que para cualquier vector de ceros y unos  $\vec{v} \in \{0, 1\}^N$ , se tiene la igualdad de probabilidades

$$P((X_1, X_2, \dots, X_N) = \vec{v}) = P((X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(N)}) = \vec{v})$$

donde  $\sigma$  es cualquier permutación en  $S_N$  (reordenamiento de  $1, 2, \dots, N$ ). Este hecho, muy poco intuitivo, es ridículamente sencillo de comprobar escribiendo las cuentas. Por ejemplo, las probabilidades de que las tres primeras extracciones sea BBP, BPB o PBB, son repectivamente (abajo se indica las que hay de cada tipo en el bote):

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{1}{4} \qquad \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{2}{4} \qquad \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{2}{4}$$

B	1	2	3	3
P	1	1	1	2

B	1	2	2	3
P	1	1	2	2

B	1	1	2	3
P	1	2	2	2

La probabilidad de que al extraer  $N$  judías, las primeras  $m$  sean blancas y las  $N - m$  restantes pintas, es

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \dots \cdot \frac{m}{m+1} \cdot \frac{1}{m+2} \cdot \frac{2}{m+2} \cdot \frac{3}{m+2} \cdot \dots \cdot \frac{N-m}{N+1} = \frac{m!(N-m)!}{(N+1)!}$$

Por la propiedad de intercambiabilidad, la probabilidad de que después de  $N$  extracciones haya exactamente  $m + 1$  judías blancas en el bote es, por tanto

$$P(X_1 + X_2 + \dots + X_N = m) = \binom{N}{m} \frac{m!(N-m)!}{(N+1)!} = \frac{1}{N+1}$$

Es decir, que todas las proporciones de judías blancas y pintas son equiprobables. La distribución de esta proporción es la uniforme (para la existencia y sentido de la “distribución límite”, véase [Fe]).