Nota a um problema de Markov

por J. Marques Henriques

O seguinte problema, originàriamente devido a A. A. MARKOV (1856-1922), embora de simples formulação, foi o ponto de partida de um dos capítulos centrais de toda a Teoria das Probabilidades e dos Processos Estocásticos: o das cadeias (e dos processos) de MARKOV. O objectivo desta Nota é, além de chamar a atenção para o seu interesse histórico, apresentar o seu enunciado e solução completa.

PROBLEMA: Um dado «ideal» (i. e., com iguais probabilidades — portanto todas iguais a $1/6$ — de saída de cada uma das faces) é lançado n vezes. Qual é a probabilidade $p_n(\nu)$ $(\nu = 0, 1, 2, 3)$ de que a soma das faces saídas nos n lançamentos dividida por 4 tenha ν por resto? Por outras palavras: se designarmos por $X = \sum_{\rho=1}^{n} X_\rho$ a soma dos valores saídos em cada um dos n lançamentos, qual é a probabilidade $p_n(\nu)$ tal que

$$p_n(\nu) = P\,(X \equiv \nu \bmod 4)\,?$$

Como se vê, a formulação deste problema é muito simples. O mesmo, porém, se não pode dizer da sua resolução, como veremos abaixo. Uma formulação mais geral deste problema ainda é possível (cf. RICHTER [2, A III. 5.10]), admitindo que o dado não é «ideal» mas antes «viciado», i. e. tem diferentes probabilidades de saída de cada uma das faces; este caso é de solução um pouco mais trabalhosa, mas análoga, e por isso não o apresentaremos aqui. Além disso, o caso do dado «ideal» permite a verificação imediata de que $\lim_{n\to\infty} p_n(\nu) = 1/4$ (limite tomado no sentido da Análise), o que é de todo intuitivo devido à igual probabilidade de saída de cada uma das faces e portanto de obtenção, no limite para um número infinito de lançamentos, de iguais probabilidades para as congruências das somas dos valores saídos em cada lançamento.

Para solucionar o nosso problema, comecemos por calcular, por métodos elementares, os vários valores de $p_n(\nu)$ para $n=1$ e $n=2$. No caso de um só lançamento do dado teremos:

$$P(X_1 = k) = 1/6$$

$(k = 1, \dots, 6)$ e portanto, da tabela seseguinte:

k	$P(X_1 = k)$	ν
1	1/6	1
2	1/6	2
3	1/6	3
4	1/6	0
5	1/6	1
6	1/6	2

concluímos (pela clássica fórmula de LAPLACE) que

$$(1) \qquad \begin{aligned} p_1(0) = p_1(3) = 1/6 \\ p_1(1) = p_1(2) = 1/3 \end{aligned}$$

(Evidentemente que $\sum_\nu p_1(\nu) = 1$). Agora, com $n=2$, teremos

$$P(X_1 = k_1 \;\&\; X_2 = k_2) = 1/36$$

$(k_1, k_2 = 1, \ldots, 6)$, e portanto para os 36 pares de valores (X_1, X_2)

$$\begin{matrix} (1,1) & (2,1) & \ldots & (6,1) \\ (1,2) & (2,2) & \ldots & (6,2) \\ \cdot & \cdot & \ldots & \cdot \\ (1,6) & (2,6) & \ldots & (6,6) \end{matrix}$$

vêm os seguintes valores de ν:

$$\begin{matrix} 2 & 3 & 0 & 1 & 2 & 3 \\ 3 & 0 & 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 3 & 0 & 1 \\ 1 & 2 & 3 & 0 & 1 & 2 \\ 2 & 3 & 0 & 1 & 2 & 3 \\ 3 & 0 & 1 & 2 & 3 & 0 \end{matrix}$$

pelo que (usando de novo a fórmula de Laplace):

$$p_2(0) = p_2(2) = 9/36$$
$$p_2(1) = 8/36$$
$$p_2(3) = 10/36$$

(De novo, $\sum_{\nu} p_2(\nu) = 1$, como aliás deveria ser). Agora, para resolver o problema para qualquer valor natural de n teremos de estabelecer uma fórmula de recorrência que, depois de devidamente resolvida, nos dará a expressão analítica desejada de $p_n(\nu)$. De facto,

$$p_n(\nu) = P(X \equiv \nu \bmod 4) =$$
$$= P\left(\sum_{\rho=1}^{n-1} X_\rho \equiv 0 \bmod 4\right) \cdot P(X_n \equiv \nu \bmod 4) +$$
$$+ P\left(\sum_{\rho=1}^{n-1} X_\rho \equiv 1 \bmod 4\right) \cdot$$
$$\cdot P(X_n \equiv \nu - 1 \bmod 4) + P\left(\sum_{\rho=1}^{n-1} X_\rho \equiv 2 \bmod 4\right) \cdot$$
$$\cdot P(X_n \equiv \nu - 2 \bmod 4) + P\left(\sum_{\rho=1}^{n-1} X_\rho \equiv 3 \bmod 4\right) \cdot$$
$$\cdot P(X_n \equiv \nu - 3 \bmod 4) = p_{n-1}(0)\pi_\nu +$$
$$+ p_{n-1}(1)\pi_{\nu-1} + p_{n-1}(2)\pi_{\nu-2} +$$
$$+ p_{n-1}(3)\pi_{\nu-3},$$

onde $\pi_{\nu-j} = p_1(\nu - j)$, ou, no caso de $\nu - j$ ser negativo se toma $\pi_{\nu-j} = p_1(\nu - j + 4)$ uma vez que $\nu - j + 4$ é agora o número positivo mais pequeno que com $\nu - j$ é congruente mod 4. Temos portanto:

$$(3) \qquad p_n(\nu) = \sum_{j=0}^{3} \pi_{\nu-j}\, p_{n-1}(j)$$
$$(\nu, j = 0, 1, 2, 3).$$

Quer dizer, se designarmos por p_n o vector do espaço euclideano a 4 dimensões cujas componentes são $p_n(\nu)$ $(\nu = 0, 1, 2, 3)$, i. e.

$$p_n = \begin{bmatrix} p_n(0) \\ p_n(1) \\ p_n(2) \\ p_n(3) \end{bmatrix}$$

virá então

$$(4) \qquad p_n = A\, p_{n-1},$$

onde A é uma matriz quadrada de ordem 4, cujos elementos $a_{\nu j}$ são precisamente $\pi_{\nu-j} = P(X_\rho \equiv \nu - j \bmod 4) = p_1(\nu - j \bmod 4)$. Mas estes valores já são nossos conhecidos do caso $n = 1$, nomeadamente

$$p_1(0) = p_1(3) = 1/6$$
$$p_1(1) = p_1(2) = 1/3 \, .$$

Logo, teremos:

$$A = \begin{bmatrix} 1/6 & 1/6 & 1/3 & 1/3 \\ 1/3 & 1/6 & 1/6 & 1/3 \\ 1/3 & 1/3 & 1/6 & 1/6 \\ 1/6 & 1/3 & 1/3 & 1/6 \end{bmatrix}.$$

OBSERVAÇÃO: A primeira coluna da matriz A é precisamente o vector p_1 e as outras obtêm-se dela substituindo sucessivamente cada elemento pelo seu antecedente e o primeiro pelo último. A matriz A é uma *matriz de* MARKOV, i. e. a soma dos elementos de uma linha é a unidade, e, mais do que isso, é uma *matriz duplamente estocástica*, assim chamada pelo facto de a soma dos elementos de qualquer linha ou coluna dar sempre a unidade.

Ora, de (4) concluímos que

$$p_n = A^2 p_{n-2} = \cdots = A^{n-1} p_1 = A^n p_0 , \tag{6}$$

onde

$$p_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} .$$

E agora, de (3) e de (4) conclui-se imediatamente que

$$A = \frac{1}{6} I + \frac{1}{3} J^1 + \frac{1}{3} J^2 + \frac{1}{6} J^3 = \\ = \sum_{\nu=0}^{3} p_1(\nu) J^\nu ,$$

onde I é a matriz identidade de ordem 4 e

$$J = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

é uma matriz cíclica de ordem 4 ($J^4 = I$). (A nossa terminologia segue aqui a de ZURMÜHL [4, pg. 270], diferindo ligeiramente da usada por VICENTE GONÇALVES [3, p. 228].) Portanto (cf. [4, pg. 270, Teorema 6]), os valores próprios de J são as raízes de índice 4 da unidade, i. e.

$$\omega_\nu = e^{2\pi i \nu/4} = \cos \frac{\nu\pi}{2} + i \operatorname{sen} \frac{\nu\pi}{2} ,$$

onde i é a unidade imaginária ($i^2 = -1$) e $\nu = 0, 1, 2, 3$. Além disso os vectores próprios de J, y_ν, sendo por definição tais que

$$J y_\nu = \omega_\nu y_\nu ,$$

terão de ser os vectores

$$y_\nu = \begin{bmatrix} \omega_{3\nu} \\ \omega_{2\nu} \\ \omega_\nu \\ \omega_0 \end{bmatrix} (\nu = 0, 1, 2, 3) .$$

Os valores próprios de A são então

$$\lambda_{k+1} = \sum_{\nu=0}^{3} p_1(\nu) \omega_{k\nu} = \\ = \frac{1}{6} + \frac{1}{3} e^{2\pi i k/4} + \frac{1}{3} e^{4\pi i k/4} + \frac{1}{6} e^{6\pi i k/4}$$

($k = 0, 1, 2, 3$), e assim, em virtude de $e^{i\pi/2} = \cos \pi/2 + i \operatorname{sen} \pi/2 = i$,

$$\lambda_{k+1} = \frac{1}{6} + \frac{1}{3} i^k + \frac{1}{3} i^{2k} + \frac{1}{6} i^{3k}$$

ou, mais explìcitamente,

$$\begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= -\frac{1}{6}(1 - i) \\ \lambda_3 &= 0 \\ \lambda_4 &= -\frac{1}{6}(1 + i) = \bar{\lambda}_2 , \end{aligned} \tag{7}$$

onde a barra denota o complexo conjugado de λ_2. De modo semelhante obtêm-se os vectores próprios de A que correspondem a estes valores próprios. São eles:

$$\begin{aligned} x_1 &= (1, 1, 1, 1)' \\ x_2 &= (1, -i, -1, i)' \\ x_3 &= (1, i, -1, -i)' \\ x_4 &= (1, i, -1, -i)' , \end{aligned} \tag{8}$$

onde a pelica denota a operação de transposição vectorial. Ora, como

$$\sum_{l=1}^{4} x'_l = (4,0,0,0) = 4p'_0, \text{ vem } p_0 = \frac{1}{4}\sum x_l .$$

E de (6) concluímos que

$$p_n = A^{n-1}(A p_0) = A^{n-1}(A \frac{1}{4}\sum_l x_l) =$$
$$= \frac{1}{4} A^{n-1}(\sum_l A x_l) = \frac{1}{4} A^{n-1}(\sum_l \lambda_l x_l) =$$
$$= \frac{1}{4} A^{n-2}(\sum_l \lambda_l A x_l) = \frac{1}{4} A^{n-2}\sum_l \lambda_l^2 x_l =$$
$$= \cdots = \frac{1}{4}\sum_l \lambda_l^n x_l .$$

Daqui tiramos a seguinte expressão do vector p_n, em função dos valores e vectores próprios (conhecidos) da matriz A:

$$(9) \qquad p_n = \frac{1}{4} x_1 + \frac{1}{4}\lambda_2^n x_2 + \frac{1}{4}\lambda_4^n x_4 ,$$

pois que $\lambda_1 = 1$ e $\lambda_3 = 0$. Mas como

$$1 - i = \sqrt{2}\, e^{-i\frac{\pi}{4}} =$$
$$= \sqrt{2}\left(\cos\left(-\frac{\pi}{4}\right) + i \operatorname{sen}\left(-\frac{\pi}{4}\right)\right)$$

$$1 + i = \sqrt{2}\, e^{i\frac{\pi}{4}} =$$
$$= \sqrt{2}\left(\cos\frac{\pi}{4} + i \operatorname{sen}\frac{\pi}{4}\right),$$

sai

$$\lambda_2^n = \left[-\frac{1}{6}(1-i)\right]^n = (-1)^n\left(\frac{\sqrt{2}}{6}\right)^n \cdot$$
$$\cdot\left(\cos\left(-\frac{n\pi}{4}\right) + i \operatorname{sen}\left(-\frac{n\pi}{4}\right)\right)$$

$$\lambda_4^n = \left[-\frac{1}{6}(1+i)\right]^n = (-1)^n\left(\frac{\sqrt{2}}{6}\right)^n \cdot$$
$$\cdot\left(\cos\frac{n\pi}{4} + i \operatorname{sen}\frac{n\pi}{4}\right).$$

De modo que

$$\lambda_2^n x_2 + \lambda_4^n x_4 = \left(-\frac{1}{6}\right)^n [(1-i)^n x_2 +$$
$$+ (1+i)^n x_4] = \left(-\frac{\sqrt{2}}{6}\right)^n \begin{bmatrix} 2\cos\left(-\frac{n\pi}{4}\right) \\ 2\operatorname{sen}\left(-\frac{n\pi}{4}\right) \\ -2\cos(-n\pi/4) \\ -2\operatorname{sen}(-n\pi/4) \end{bmatrix}.$$

E, finalmente, a solução do problema de MARKOV é:

$$p_n = \begin{bmatrix} p_n(0) \\ p_n(1) \\ p_n(2) \\ p_n(3) \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} +$$

$$(10) \qquad + \frac{1}{2}\left(-\frac{\sqrt{2}}{6}\right)^n \begin{bmatrix} \cos\left(-\frac{n\pi}{4}\right) \\ \operatorname{sen}\left(-\frac{n\pi}{4}\right) \\ -\cos\left(-\frac{n\pi}{4}\right) \\ -\operatorname{sen}\left(-\frac{n\pi}{4}\right) \end{bmatrix}.$$

Para $n = 1$ e $n = 2$ obtemos imediatamente os valores a que tínhamos chegado em (1) e (2). Para $n = 3$ obtemos, a partir de

$$\cos\left(-\frac{3\pi}{4}\right) = -\frac{\sqrt{2}}{2}$$

e de

$$\left(-\frac{\sqrt{2}}{6}\right)^3 = -\frac{\sqrt{8}}{216},$$

os valores

$$p_3(0)=\frac{55}{216}, p_3(1)=\frac{55}{216}, p_3(2)=$$
$$=\frac{53}{216}, p_3(3)=\frac{53}{216}, \text{ com } \sum p_3(\nu)=1.$$

Como tínhamos afirmado de início, os valores de $p_n(\nu)$ tendem alternadamente para o limite $1/4$, independente de ν. Isso é agora evidente a partir de (10), pois que por $(-\sqrt{2}/6)^n \to 0$, ao passarmos ao limite, a segunda parcela do lado direito de (10) tende para o vector zero. Este resultado é, aliás, susceptível de uma justificação heurística, pois, sendo o dado ideal (como considerámos na formulação do problema), no limite para um número infinito de lançamentos obtém-se uma simetria para as 4 diferentes congruências possíveis, pelo que se deverá ter $p_n(\nu) \to 1/4$, para qualquer ν (este raciocínio é, porém, falso no caso de o dado ser viciado, e. g. se só forem positivas as probabilidades de obtenção de faces com um número par de pontos). Teríamos chegado ainda ao mesmo resultado, observando que a matriz A define uma aplicação de contracção no sub-espaço do vulgar espaço métrico euclideano a 4 dimensões formado pelos possíveis vectores p_n (de coordenadas $0 \leq p_n(\nu) \leq 1$, com $\Sigma p_n(\nu)=1$). Pelo teorema de Banach relativo aos pontos fixos, deverá haver um e um só ponto fixo, que logo se vê ser o vector cujas 4 coordenadas são todas iguais a $1/4$.

Nota. A sucessão $\{p_n; n=0,1,2,\cdots\}$ dos vectores p_n constitui um dos exemplos mais simples de uma sucessão de distribuições de probabilidade em que cada uma delas depende apenas da precedente p_{n-1}, tal como num processo iterativo simples $\{p_n; n=0,1,\cdots\}$ é portanto um exemplo, extremamente simples, de uma cadeia de Markov.

BIBLIOGRAFIA

[1] Markov, A. A. (1912). *Wahrscheinlichkeitsrechnung*. Teubner Verlag. Leipzig.

[2] Richter, H. (1956). *Wahrscheinlichkeitstheorie*. Springer-Verlag. Berlin-Göttingen-Heidelberg.

[3] Vicente Gonçalves, J. (1950). *Curso de Álgebra Superior* 2. Lisboa.

[4] Zurmühl, R. (1963). *Matrizen* (3. Auflage). Springer-Verlag. Berlin-Göttingen-Heidelberg.

Pari-mutuel betting

by John J. Wiorkowski
Chicago

Pari-mutuel betting is a form of betting on horses in which those who bet on the winning horse share the total amount bet on all horses less a small per cent which is paid to the management. Considering that statisticians and mathematicians have always been particularly interested in gambling systems, it is surprising that pari-mutuel betting has received such sparse attention. This lack of attention is perhaps attributable to the inherent difficulty of determining the actual probability that a horse will win a race, since this probability depends on the other horses in the race, the conditions of the track, and a plethora of other factors. To circumvent this difficulty, let us suppose that the better