

O valor esperado de uma quantidade aleatória

Paulo Cezar Pinto Carvalho

IMPA e EMAP/FGV

Um conceito simples e útil, mas que não é normalmente explorado no Ensino Básico no Brasil é o de *valor esperado de uma quantidade aleatória*. Começemos com um exemplo.

Exemplo 1. *Em um famoso jogo popular, o acertador deve apontar 1 dentre 25 resultados possíveis para um sorteio. Se acertar, ele recebe 18 vezes o valor apostado. Qual é o lucro (ou prejuízo) médio diário que deve ser esperado por alguém que aposte diariamente 1 real neste jogo por um período longo?*

A primeira observação a fazer é que não é possível obter um valor exato para o lucro médio do apostador: ele vai depender de sua “sorte”, ou seja, do experimento aleatório que consiste nos sucessivos sorteios. No entanto, é possível ter uma ideia bastante aproximada do resultado alcançado pelo apostador, como veremos a seguir.

Em cada realização do jogo, há duas possibilidades para o apostador: se ele acerta o número sorteado, o que ocorre com probabilidade $1/25$, ele tem um lucro de $18 - 1 = 17$ reais; no caso contrário, que ocorre com probabilidade $24/25$, ele tem um prejuízo de 1 real. Podemos resumir estes fatos na tabela abaixo:

Lucro (L)	Probabilidade
17	$1/25$
-1	$24/25$

Mas o que significa ter probabilidade $1/25$ de ter um lucro de 17 reais? A interpretação de probabilidade mais útil para as aplicações (a chamada interpretação “frequentista”) é a de que, se o sorteio for realizado muitas vezes, a frequência de resultados em que esse lucro é observado é aproximadamente igual a $1/25$ das realizações.

Assim, se o sorteio é feito n vezes, onde n é grande, o apostador lucra 17 reais aproximadamente $\frac{n}{25}$ vezes e tem um prejuízo de 1 real aproximadamente $\frac{24n}{25}$ vezes.

Portanto, o lucro total é aproximadamente $17 \cdot \frac{n}{25} - 1 \cdot \frac{24n}{25} = -\frac{7n}{25}$ e o lucro médio diário, obtido dividindo o lucro total pelo número n de dias de aposta, é $-\frac{7}{25} = -0,28$.

Concluimos, portanto, que o apostador contumaz terá um prejuízo médio diário aproximado de R\$ 0,28. Mas será que este resultado reflete a realidade? Um modo de nos convencermos (e a nossos alunos) da utilidade do resultado encontrado (e de suas limitações) é simularmos o processo repetido de aposta. Isto pode ser feito, em pequena escala, com materiais simples, como uma caixa da qual sortamos papéis numerados de 1 a 25. Para simulação em grande escala, é muito melhor recorremos a um computador, como mostra a planilha que pode ser obtida em <https://sites.google.com/site/papmem2015/sorteio.xlsx>. Ela mostra o resultado de simular o processo de aposta com 10 apostadores e mostra o ganho médio de cada um após 10, 100, 10000 e 100000 dias de apostas (se você obtiver esta planilha e os resultados obtidos forem diferentes dos mostrados abaixo, não estranhe: cada vez que a planilha é aberta, ela repete o experimento aleatório, fornecendo resultados diferentes a cada vez).

	Lucro médio após o número indicado de realizações				
Apostador	10	100	1000	10000	100000
1	-1	-0,46	-0,30204	-0,253	-0,27532
2	-1	-0,46	-0,22857	-0,2476	-0,26398
3	-1	-0,28	-0,22857	-0,3556	-0,29026
4	-1	0,08	-0,28367	-0,3106	-0,28522
5	-1	-0,46	-0,37551	-0,2854	-0,2809
6	-1	-0,64	-0,17347	-0,316	-0,27874
7	-1	-0,28	-0,33878	-0,2386	-0,28396
8	-1	-0,28	-0,37551	-0,2458	-0,27712
9	0,8	-0,1	-0,37551	-0,3574	-0,28
10	-1	0,08	-0,1551	-0,244	-0,29404

Após 10 realizações, há ainda muita flutuação nos resultados. Nas sequências que observamos, um dos apostadores está no lucro após 10 apostas (ele ganhou uma vez e perdeu nas outras nove). Com 100 apostas, ainda há dois apostadores que estão tendo um pequeno lucro (ambos ganharam em 6 das 100 ocasiões), mas com 1000 apostas, todos já estão tendo prejuízo. Com 10000 apostas, o prejuízo médio dos 10 apostadores está entre R\$ 0,2386 e R\$ 0,3574 e, com 100000 apostas, entre R\$ 0,26398 e R\$ 0,29404. Assim, à medida que aumenta o número de dias observados, o lucro médio observado varia, mas oscila cada vez menos em torno do valor esperado de R\$ 0,28 (este comportamento pode ser descrito com precisão em teoremas conhecidos como “Leis dos Grandes Números”).

Enfim, a análise que fizemos anteriormente, que nos levou a encontrar o valor esperado de R\$ 0,28 para o lucro médio, é útil para prever o comportamento a longo prazo do lucro médio do apostador (embora não represente o valor exato a ser observado em uma dada sequência de apostas). Isto motiva a definição de valor esperado de uma quantidade numérica aleatória (ou, mais tecnicamente, de uma *variável aleatória*).

Definição: Suponhamos que, em um experimento aleatório, uma quantidade observada X possa assumir os valores x_1, x_2, \dots, x_n , com probabilidades respectivamente iguais a p_1, p_2, \dots, p_n (com $p_1, p_2, \dots, p_n \geq 0$ e $p_1 + p_2 + \dots + p_n = 1$). O *valor esperado* ou *valor médio* dessa quantidade (ou variável) aleatória é

$$E X = p_1 x_1 + p_2 x_2 + \dots + p_n x_n = \sum_{k=1}^n p_k x_k$$

Observe que isto corresponde exatamente ao que fizemos acima para calcular o lucro médio diário esperado após um grande número de realizações: $E(\text{Lucro}) = \frac{1}{25} \cdot 17 + \frac{24}{25} \cdot (-1) = -\frac{7}{25} = -0,28$.

Note, também, que o valor esperado de uma quantidade aleatória é a média ponderada dos valores que ela pode assumir, com pesos respectivamente iguais às respectivas probabilidades (se você está sentindo falta da divisão pela soma dos pesos, lembre-se que $p_1 + p_2 + \dots + p_n = 1$).

Exemplo 2. Uma moeda honesta é lançada três vezes. Considere as seguintes variáveis aleatórias observadas neste experimento:

$X = \text{número de caras}$

$Y = \text{número de sequências de resultados iguais consecutivos}$

Calcular os valores esperados de X , Y e $X+Y$.

A tabela abaixo dá os resultados possíveis para os resultados do experimento aleatório, suas respectivas probabilidades e os valores de X , Y e $X+Y$ em cada caso.

Resultado	Probabilidade	X	Y	$X+Y$
cara, cara, cara	$1/8$	3	1	$3+1=4$
cara, cara, coroa	$1/8$	2	2	$2+2=4$
cara, coroa, cara	$1/8$	2	3	$2+3=5$
coroa, cara, cara	$1/8$	2	2	$2+2=4$
cara, coroa, coroa	$1/8$	1	2	$1+2=3$
coroa, cara, coroa	$1/8$	1	3	$1+3=4$
coroa, coroa, cara	$1/8$	1	2	$1+2=3$
coroa, coroa, coroa	$1/8$	0	1	$0+1=1$

A partir daí, podemos extrair, para cada variável aleatória, os valores que ela assume e as respectivas probabilidades. Por exemplo, para X , temos:

X	Probabilidade
0	$1/8$
1	$3/8$
2	$3/8$
3	$1/8$

(Observe que, por exemplo, $X=1$ ocorre em três resultados, com uma probabilidade total de $3/8$).

Logo, o valor esperado de X é

$$E X = \frac{1}{8} \cdot 0 + \frac{3}{8} \cdot 1 + \frac{3}{8} \cdot 2 + \frac{1}{8} \cdot 3 = \frac{12}{8} = 1,5.$$

Na verdade, não é necessário obter a tabela das probabilidades de X para calcular seu valor esperado. O cálculo pode ser feito diretamente a partir da tabela de resultados, multiplicando a probabilidade de cada resultado pelo valor de X em cada caso.

$$E X = \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 0 = \frac{12}{8} = 1,5.$$

Observe que o cálculo anterior corresponde a simplesmente agrupar os valores iguais de X em um único termo, por exemplo, o termo $\frac{3}{8} \cdot 1$ é a soma dos três termos $\frac{1}{8} \cdot 1$ correspondentes aos resultados (cara, cara, coroa), (cara, coroa, cara) e (coroa, cara, cara).

Fazendo o mesmo com cada uma das outras variáveis, obtemos:

$$E Y = \frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 1 = \frac{16}{8} = 2.$$

$$E (X+Y) = \frac{1}{8} \cdot (3+1) + \frac{1}{8} \cdot (2+2) + \frac{1}{8} \cdot (2+3) + \frac{1}{8} \cdot (2+2) + \frac{1}{8} \cdot (1+2) + \frac{1}{8} \cdot (1+3) + \frac{1}{8} \cdot (1+2) + \frac{1}{8} \cdot (0+1) =$$

$$\begin{aligned} & \left(\frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 0 \right) \\ & + \left(\frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 1 \right) = \\ & = \frac{12}{8} + \frac{16}{8} = 3,5 \end{aligned}$$

(Organizamos as contas acima de modo a deixar claro que o valor esperado de $X+Y$ é sempre igual à soma dos valores esperados de X e Y , já que a expressão que fornece $E(X+Y)$ pode ser decomposta na soma das expressões de EX e EY .)

De modo mais geral, temos o seguinte teorema:

Linearidade do valor esperado: Se X e Y são variáveis aleatórias relativas ao mesmo experimento aleatório, temos $E(aX + bY) = aEX + bEY$.

Mas cuidado: não vale, em geral, a propriedade análoga para variáveis definidas por outras operações. Por exemplo, na mesma situação, temos

$$E(X^2) = \frac{1}{8} \cdot 3^2 + \frac{1}{8} \cdot 2^2 + \frac{1}{8} \cdot 2^2 + \frac{1}{8} \cdot 2^2 + \frac{1}{8} \cdot 1^2 + \frac{1}{8} \cdot 1^2 + \frac{1}{8} \cdot 1^2 + \frac{1}{8} \cdot 0^2 = \frac{24}{8} = 3,$$

enquanto $(EX)^2 = 1,5^2 = 2,25$; portanto, $E(X^2) \neq (EX)^2$ (na verdade, a única situação em que os dois valores coincidem é aquela em que X tem um valor constante, não importando o resultado do experimento aleatório).

Exemplo 3: Um dado honesto é lançado 10 vezes. Qual é o número esperado de lançamentos em que o resultado obtido é 6?

É possível obter a tabela de probabilidades relativas à variável aleatória X , definida como o número de lançamentos em que o resultado é 6. Os valores que X pode assumir são 0, 1, 2, ..., 10 e podemos calcular, para cada k neste conjunto de valores, a probabilidade de que X seja igual a k . A probabilidade de que sejam registrados k resultados iguais a 6 consecutivos e, a seguir, $10 - k$ resultados diferentes de 6 é $\frac{1 \cdot 1 \cdot \dots \cdot 1 \cdot 5 \cdot 5 \cdot \dots \cdot 5}{6^k} = \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{10-k}$. Mas os k resultados iguais a 6 não precisam ocorrer todos juntos no início. Os locais em que eles ocorrem podem ser escolhidos de C_{10}^k modos. Logo, temos $p_k = P(X = k) = C_{10}^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{10-k}$, para $k = 0, 1, 2, \dots, 10$. Daí, podemos calcular o valor esperado de X :

$$EX = \sum_{k=0}^{10} p_k k = \sum_{k=0}^{10} C_{10}^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{10-k} \cdot k$$

Se fizermos as contas (trabalhosas) expressas pelo somatório, vamos encontrar $EX = 10/6 = 5/3$.

Mas poderíamos chegar a esse resultado de forma bem mais rápida, usando a linearidade do valor esperado. Para $k = 1, \dots, 10$, vamos chamar de X_k a variável aleatória que descreve se ocorreu ou não um resultado igual a 6 no k -ésimo lançamento. Isto é, vamos considerar que $X_k = 1$ se saiu 6 nesse lançamento e $X_k = 0$ caso contrário. O número total de lançamento em que saiu 6 é simplesmente a soma $X = X_1 + X_2 + \dots + X_{10}$. Pela propriedade aditiva do valor esperado, $EX = EX_1 + EX_2 + \dots + EX_{10}$. Mas, para cada k ,

$$EX_k = P(X_k = 0) \cdot 0 + P(X_k = 1) \cdot 1 = \frac{5}{6} \cdot 1 = \frac{5}{6}.$$

$$\text{Portanto, } EX = \frac{5}{6} + \frac{5}{6} + \dots + \frac{5}{6} = \frac{10}{6} = \frac{5}{3}.$$

Terminamos com uma situação envolvendo variáveis aleatórias que podem assumir uma infinidade de valores. Se uma variável aleatória assume uma quantidade infinita mas enumerável de valores (isto é, que possam ser escritos em uma lista, estabelecendo uma correspondência biunívoca com os números naturais), a definição de valor esperado continua a mesma: $EX = \sum p_k x_k$, mas o somatório agora tem um número infinito de parcelas e pode ou não convergir para um número real.

Exemplo 4. *Um apostador faz seguidas apostas em um jogo em que ele tem probabilidade $\frac{1}{2}$ de ganhar. Se ganhar, ele recebe o dobro do valor apostado. Ele adota a seguinte estratégia: começa apostando 1 real; se ganhar, se retira do jogo com um lucro de 1 real. Se perder, joga 2 reais; se ganhar, se retira do jogo com um lucro de 1 real ($4 - 2 - 1$). Se perder, joga 4 reais, e assim por diante, sempre dobrando a aposta em caso de derrota, assegurando que, quando finalmente ganhar, terá um lucro de 1 real.*

- a) *Qual é o valor esperado do número N de apostas até ganhar?*
- b) *Qual é o valor esperado do investimento total X que terá que ser feito para ganhar?*

As duas variáveis aleatórias N e X podem ambas assumir valores em um conjunto infinito. O valor de N pode ser qualquer número natural. Já X pode ser expressa em termos de N . Se forem N rodadas até ganhar, o investimento será $X = 1 + 2 + \dots + 2^{N-1} = 2^N - 1$.

A tabela de probabilidades de N é dada por $P(N=k) = \left(\frac{1}{2}\right)^k$, para $k = 1, 2, \dots$. De fato, para que haja k rodadas de apostas, o jogador deve ter perdido as $k - 1$ primeiras e ganhado a k -ésima. Como cada uma destas k ocorrências tem probabilidade $\frac{1}{2}$ e elas são independentes entre si, a probabilidade de que a vitória ocorra na k -ésima aposta é $\left(\frac{1}{2}\right)^k$. Logo:

$$EN = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{16} \cdot 4 + \dots$$

Uma forma de calcular a soma infinita acima consiste em decompor cada parcela da forma $\left(\frac{1}{2}\right)^k k$ em k parcelas iguais a $\left(\frac{1}{2}\right)^k$, conforme abaixo:

$$EN = \begin{array}{cccc} \frac{1}{2} + & \frac{1}{4} + & \frac{1}{8} + & \frac{1}{16} + & \dots \\ & \frac{1}{4} + & \frac{1}{8} + & \frac{1}{16} + & \dots \\ & & \frac{1}{8} + & \frac{1}{16} + & \dots \end{array}$$

Em cada linha, temos a soma de uma progressão geométrica de razão $\frac{1}{2}$. A soma da primeira linha é $\frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1$; da segunda é $\frac{\frac{1}{4}}{1 - \frac{1}{2}} = \frac{1}{2}$; da terceira, $\frac{\frac{1}{8}}{1 - \frac{1}{2}} = \frac{1}{4}$, e assim por diante, formando uma nova progressão geométrica de razão $1/2$. Logo

$$EN = 1 + \frac{1}{2} + \frac{1}{4} + \dots = 2.$$

(Um cálculo parecido com este aparece em [1]; juntamente com a propriedade da linearidade do valor esperado, ela é usada lá para obter o número esperado de figurinhas a serem compradas para preencher um álbum).

Portanto, em média são jogadas duas rodadas até que o apostador se retire do jogo com seu lucro de 1 real. Neste momento, poderíamos ser tentados a achar que, como o valor a ser investido X é igual a $2^N - 1$, onde N é o número de rodadas, o investimento esperado até a vitória é igual a $2^2 - 1 = 3$ reais. Na verdade, este é um exemplo que mostra, de modo radical, que, em geral, **não vale** $E(f(X)) = f(EX)$! O investimento X assume os valores $2^1-1, 2^2-1, 2^3-1, \dots$ com probabilidades respectivamente iguais a $1/2, 1/4, 1/8, \dots$. Logo:

$$EX = \frac{1}{2}(2 - 1) + \frac{1}{4}(4 - 1) + \frac{1}{8}(8 - 1) + \dots = \left(1 - \frac{1}{2}\right) + \left(1 - \frac{1}{4}\right) + \left(1 - \frac{1}{8}\right) + \dots = +\infty$$

(Note que a soma das parcelas iguais a 1 cresce sem limite, enquanto as demais parcelas somam 1.)

Portanto, o investimento esperado de nosso jogador é infinito, o que torna a estratégia descrita acima extremamente arriscada! A aparente contradição que surge ao estudar esta situação, em que quase sempre o valor investido é modesto, mas seu valor esperado é infinito, é conhecida como o “Paradoxo de São Petersburgo” e tem sido objeto de estudo e discussão desde 1738, quando foi descrito por Daniel Bernoulli .

Esperamos ter convencido o leitor que a noção de valor esperado de uma quantidade aleatória que assume um número finito de valores (ou, em alguns casos, mesmo infinito, desde que enumerável) é acessível a um aluno de Ensino Médio e pode ajudar a tornar mais palpáveis as aplicações de Probabilidade a problemas reais. Para saber mais sobre o assunto, [2] é uma leitura indicada.

Referências

[1] Paulo Cezar Pinto Carvalho. *Quantas figurinhas comprar para completar o álbum da copa?* **Revista do Professor de Matemática 73**, 2010.

[2] Sheldon Ross. **Probabilidade: um curso moderno, com aplicações**. Bookman, 2010.