TarefALEA - Probabilidade

N.º 8 – Subir ou descer uma escada ao sabor do acaso

O Pedro e a Clarinha moram numa moradia com três pisos. Do rés-do chão, piso 0, acede-se ao piso -1, por uma escada com quinze degraus, e ao piso 1, por uma escadas com dezoito degraus.

Certo dia, o Pedro encontrou um dado na gaveta da sua secretária e lembrou-se de propor o seguinte jogo à Clarinha, que era uma grande fã deste tipo de jogos, em que entrava o "acaso", desde que tinha começado a dar, em Matemátrica A, o tema "Probabilidade"...

O jogo consistia no seguinte: lançavam o dado e, conforme o resultado, subiam ou desciam alguns degraus (eles estavam no rés-do-chão), de acordo com o seguinte esquema:

Se saísse 1, 2 ou 3, desciam, respetivamente 5, 3, ou 1 degraus;

Se saísse 4, 5 ou 6, subiam, respetivamente 2, 4 ou 6 degraus.

Pretendiam saber qual a probabilidade de, após 3 lançamentos do dado, estarem a caminho do piso 1 ou a caminho do piso -1!

Formalização da questão

Esta questão pode ser formalizada da seguinte forma: considere-se a variável aleatória (v.a.) X, que representa o resultado do lançamento do dado, em que se admite que:

O dado é equilibrado;

Os lançamentos são independentes uns dos outros.

Assumindo as condições anteriores, admitimos para X um modelo de probabilidade, traduzido pela seguinte função massa de probabilidade (f.m.p.):

$X = i$	-5	-3	-1	2	4	6
$P(X = i)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

Ao considerar o resultado do lançamento do dado 3 vezes, está a pensar-se na soma de 3 variáveis aleatórias $X_1 + X_2 + X_3$, em que X_1, X_2 e X_3 são variáveis aleatórias com a mesma distribuição de X, independentes, que se representam por X_1, X_2 e X_3, para significar as 3 vezes em que se lançou o dado (nesta situação, podemos considerar 3 dados diferentes, ou o mesmo dado que se lança 3 vezes).

O que se pretende é obter a $P(X_1 + X_2 + X_3 > 0)$ e a $P(X_1 + X_2 + X_3 < 0)$!

Esta questão pode ser respondida utilizando a definição **clássica** de Probabilidade, a que se dará o nome de abordagem **teórica**, ou utilizando a abordagem **empírica** ou **frequencista**, para estimar a probabilidade.

Abordagem teórica

Para calcular $P(X_1 + X_2 + X_3 > 0)$ e $P(X_1 + X_2 + X_3 < 0)$, é necessário começar por obter a função massa de probabilidade de $X_1 + X_2 + X_3$. Esta função é constituída pelos valores possíveis para a soma dos resultados dos 3 lançamentos, em número de 6^3, e respetivas probabilidades. O processo será facilitado se se começar por obter os valores possíveis para $X_1 + X_2$, em número de 6^2, utilizando-se uma tabela como a que se apresenta a seguir, com os resultados possíveis para a soma:

X_2 \ X_1	-5	-3	-1	2	4	6
-5	-10	-8	-6	-3	-1	1
-3	-8	-6	-4	-1	1	3
-1	-6	-4	-2	1	3	5
2	-3	-1	1	4	6	8
4	-1	1	3	6	8	10
6	1	3	5	8	10	12

Para obter os valores possíveis para $X_1 + X_2 + X_3$, considere-se uma tabela idêntica à anterior, mas com os valores da soma de $X_1 + X_2$ e X_3:

X_3 \ $X_1 + X_2$	-10	-8	-8	-6	-6	-6	-3	-4	-4	-3	-1	-1	-2	-1	-1	1	1	1
-5	-15	-13	-13	-11	-11	-11	-8	-9	-9	-8	-6	-6	-7	-6	-6	-4	-4	-4
-3	-13	-11	-11	-9	-9	-9	-6	-7	-7	-6	-4	-4	-5	-4	-4	-2	-2	-2
-1	-11	-9	-9	-7	-7	-7	-4	-5	-5	-4	-2	-2	-3	-2	-2	0	0	0
2	-8	-6	-6	-4	-4	-4	-1	-2	-2	-1	1	1	0	1	1	3	3	3
4	-6	-4	-4	-2	-2	-2	1	0	0	1	3	3	2	3	3	5	5	5
6	-4	-2	-2	0	0	0	3	2	2	3	5	5	4	5	5	7	7	7

(continuação da tabela anterior)

X_3 \ $X_1 + X_2$	1	1	1	3	3	4	3	3	5	6	6	5	8	8	8	10	10	12
-5	-4	-4	-4	-2	-2	-1	-2	-2	0	1	1	0	3	3	3	5	5	7
-3	-2	-2	-2	0	0	1	0	0	2	3	3	2	5	5	5	7	7	9
-1	0	0	0	2	2	3	2	2	4	5	5	4	7	7	7	9	9	11
2	3	3	3	5	5	6	5	5	7	8	8	7	10	10	10	12	12	14
4	5	5	5	7	7	8	7	7	9	10	10	9	12	12	12	14	14	16
6	7	7	7	9	9	10	9	9	11	12	12	11	14	14	14	16	16	18

Uma vez que os 216 valores da tabela anterior são igualmente prováveis, obtém-se a seguinte função massa de probabilidade para a v.a. $X_1 + X_2 + X_3$:

Função massa de probabilidade da v.a. $X_1 + X_2 + X_3$

$X_1 + X_2 + X_3 = i$	-15	-13	-11	-9	-8	-7	-6	-5	-4	-3
$P(X_1 + X_2 + X_3 = i)$	1/216	3/216	6/216	7/216	3/216	6/216	9/216	3/216	18/216	1/216
	0,005	0,014	0,028	0,032	0,014	0,028	0,042	0,014	0,083	0,005

Função massa de probabilidade da v.a. $X_1 + X_2 + X_3$ (continuação)

$X_1 + X_2 + X_3 = i$	-2	-1	0	1	2	3	4	5	6
$P(X_1 + X_2 + X_3 = i)$	21/216	3/216	18/216	9/216	9/216	18/216	3/216	21/216	1/216
	0,097	0,014	0,083	0,042	0,042	0,083	0,014	0,097	0,005

Função massa de probabilidade da v.a. $X_1 + X_2 + X_3$ (continuação)

$X_1 + X_2 + X_3 = i$	7	8	9	10	11	12	14	16	18
$P(X_1 + X_2 + X_3 = i)$	18/216	3/216	9/216	6/216	3/216	7/216	6/216	3/216	1/216
	0,083	0,014	0,042	0,028	0,014	0,032	0,028	0,014	0,005

Uma vez obtido o modelo de probabilidade para a v.a. $X_1 + X_2 + X_3$, podem-se calcular as probabilidades pretendidas, ou seja:

$$P(\text{Após 3 lançamentos do dado estarem a caminho do piso 1}) \\ = P(X_1 + X_2 + X_3 > 0) = \frac{117}{216} = 0{,}542$$

$$P(\text{Após 3 lançamentos do dado estarem a caminho do piso -1}) \\ = P(X_1 + X_2 + X_3 < 0) = \frac{81}{216} = 0{,}375$$

Conclui-se que, após 3 lançamentos do dado, a probabilidade de estarem a subir é aproximadamente 54%, enquanto a de estarem a descer é, aproximadamente, 38%.

Outros resultados interessantes...

Da tabela da f.m.p., podem-se obter outras probabilidades interessantes, como, por exemplo, as probabilidades de, após 3 lançamentos, não saírem do piso 0, chegarem ao piso -1 ou chegarem ao piso 1, respetivamente iguais a 8,3%, 0,5% e 0,5%.

Abordagem empírica

Para simular o processo, pode-se utilizar a seguinte metodologia, considerando uma folha de cálculo, como o Excel:

1. Numa coluna da folha do Excel, gera-se um número aleatório[1] entre 1 e 6 para simular o resultado do lançamento do dado, e na coluna seguinte atribui-se o valor que corresponde ao resultado observado, de acordo com as regras enunciadas na introdução da tarefa;
2. Repete-se o processo anterior 3 vezes, para simular o lançamentpo do dado 3 vezes;
3. Considera-se uma coluna com a soma dos valores obtidos nos 3 lançamentos;
4. Repetem-se os passos anteriores "muitas" vezes;
5. Calculam-se as frequências relativas dos valores observados para a soma dos resultados nos 3 lançamentos e constrói-se o diagrama de barras, que é a imagem estatística da função massa de probabilidade de $X_1 + X_2 + X_3$.

Apresenta-se de seguida uma parte da folha de Excel, com as fórmulas correspondentes aos passos 1, 2 e 3:

	A	B	C	D	E	F	G
1	Resultado 1º lanç		Res ulta		Resultado 3º lanç		Resultado 3 lan
2	=RANDBETWEEN(1;6)	=IF(A2=1;-1;IF(A2=2;-3;IF(A2=3;-5;IF(A2=4;2;IF(A2=5;4;6)))))	=R	=	=RANDBETWEEN(1;6)	=IF(E2=1;-1;IF(E2=2;-3;IF(E2=3;-5;IF(E2=4;2;IF(E2=5;4;6)))))	=B2+D2+F2
3	=RANDBETWEEN(1;6)	=IF(A3=1;-1;IF(A3=2;-3;IF(A3=3;-5;IF(A3=4;2;IF(A3=5;4;6)))))	=R	=	=RANDBETWEEN(1;6)	=IF(E3=1;-1;IF(E3=2;-3;IF(E3=3;-5;IF(E3=4;2;IF(E3=5;4;6)))))	=B3+D3+F3
4	=RANDBETWEEN(1;6)	=IF(A4=1;-1;IF(A4=2;-3;IF(A4=3;-5;IF(A4=4;2;IF(A4=5;4;6)))))	=R	=	=RANDBETWEEN(1;6)	=IF(E4=1;-1;IF(E4=2;-3;IF(E4=3;-5;IF(E4=4;2;IF(E4=5;4;6)))))	=B4+D4+F4
5	=RANDBETWEEN(1;6)	=IF(A5=1;-1;IF(A5=2;-3;IF(A5=3;-5;IF(A5=4;2;IF(A5=5;4;6)))))	=R	=	=RANDBETWEEN(1;6)	=IF(E5=1;-1;IF(E5=2;-3;IF(E5=3;-5;IF(E5=4;2;IF(E5=5;4;6)))))	=B5+D5+F5
6	=RANDBETWEEN(1;6)	=IF(A6=1;-1;IF(A6=2;-3;IF(A6=3;-5;IF(A6=4;2;IF(A6=5;4;6)))))	=R	=	=RANDBETWEEN(1;6)	=IF(E6=1;-1;IF(E6=2;-3;IF(E6=3;-5;IF(E6=4;2;IF(E6=5;4;6)))))	=B6+D6+F6
7	=RANDBETWEEN(1;6)	=IF(A7=1;-1;IF(A7=2;-3;IF(A7=3;-5;IF(A7=4;2;IF(A7=5;4;6)))))	=R	=	=RANDBETWEEN(1;6)	=IF(E7=1;-1;IF(E7=2;-3;IF(E7=3;-5;IF(E7=4;2;IF(E7=5;4;6)))))	=B7+D7+F7
8	=RANDBETWEEN(1;6)	=IF(A8=1;-1;IF(A8=2;-3;IF(A8=3;-5;IF(A8=4;2;IF(A8=5;4;6)))))	=R	=	=RANDBETWEEN(1;6)	=IF(E8=1;-1;IF(E8=2;-3;IF(E8=3;-5;IF(E8=4;2;IF(E8=5;4;6)))))	=B8+D8+F8
9	=RANDBETWEEN(1;6)	=IF(A9=1;-1;IF(A9=2;-3;IF(A9=3;-5;IF(A9=4;2;IF(A9=5;4;6)))))	=R	=	=RANDBETWEEN(1;6)	=IF(E9=1;-1;IF(E9=2;-3;IF(E9=3;-5;IF(E9=4;2;IF(E9=5;4;6)))))	=B9+D9+F9
10	=RANDBETWEEN(1;6)	=IF(A10=1;-1;IF(A10=2;-3;IF(A10=3;-5;IF(A10=4;2;IF(A10=5;4;	=R	=	=RANDBETWEEN(1;6)	=IF(E10=1;-1;IF(E10=2;-3;IF(E10=3;-5;IF(E10=4;2;IF(E10=5;4	=B10+D10+F1

...

	A	B	C	D	E	F	G
1		Resultado 1.º lanç		Resultado 2.º lanç		Resultado 3.º lanç	Soma resultados 3 lançamentos
2	3	-5	2	-3	1	-1	-9
3	1	-1	6	6	6	6	11
4	6	6	5	4	1	-1	9
5	3	-5	6	6	1	-1	0
6	2	-3	1	-1	2	-3	-7
7	4	2	4	2	4	2	6
8	5	4	6	6	6	6	16
9	6	6	4	2	5	4	12
10	1	-1	6	6	1	-1	4

...

[1] Mais corretamente, os números são pseudo-aleatórios, já que são gerados por um mecanismo determinista, comportando-se, no entanto, como se fossem aleatórios.

	A	B	C	D	E	F	G
4992	5	4	1	-1	4	2	5
4993	3	-5	1	-1	1	-1	-7
4994	1	-1	4	2	2	-3	-2
4995	2	-3	6	6	2	-3	0
4996	6	6	3	-5	3	-5	-4
4997	1	-1	2	-3	5	4	0
4998	2	-3	2	-3	4	2	-4
4999	3	-5	4	2	1	-1	-4
5000	3	-5	6	6	1	-1	0
5001	1	-1	1	-1	2	-3	-5

Foram feitas 5 000 repetições, tendo-se obtido a seguinte tabela de frequências para os valores obtidos como soma dos resultados nos 3 lançamentos do dado:

Tabela frequências da soma dos resultados nos 3 lançamentos

Valor	-15	-13	-11	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
Freq. rel.	0,004	0,013	0,026	0,031	0,017	0,027	0,04	0,013	0,085	0,006	0,095	0,014	0,087

Tabela frequências relativas da soma dos resultados nos 3 lançamentos (continuação)

1	2	3	4	5	6	7	8	9	10	11	12	14	16	18
0,042	0,039	0,088	0,013	0,102	0,006	0,082	0,012	0,043	0,025	0,014	0,034	0,026	0,012	0,005

A soma das frequências relativas dos valores menores que 0 e maiores que 0 são, respetivamente, 0,3708 e 0,5424, donde se conclui que as estimativas obtidas para as probabilidades de, ao fim de 3 lançamentos do dado, estarem a descer ou a subir, são respetivamente, 37% e 54%.

Comparação dos resultados empíricos, com os resultados teóricos

Resumindo os resultados obtidos anteriormente, tem-se:

- $P(\text{Estar a subir após 3 lançamentos}) \approx 54\%$
  $\text{Estimativa para a probabilidade anterior} \approx 54\%$

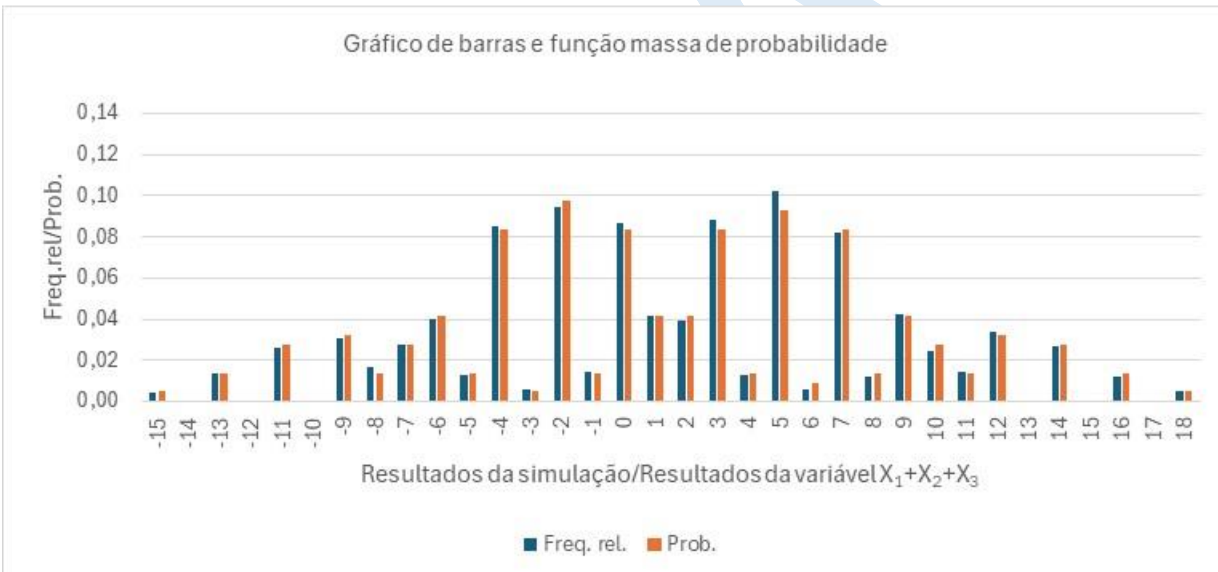
- $P(\text{Estar a descer após 3 lançamentos}) \approx 38\%$
  $\text{Estimativa para a probabilidade anterior} \approx 37\%$

pelo que se pode admitir que as estimativas obtidas são bastante razoáveis.

O diagrama de barras e a função massa de probabilidade

Pode-se dizer que o gráfico de barras é a imagem estatística da função massa de probabilidade da População de onde se recolheram os Dados com os quais se construíu o diagrama de barras, ou seja, a função massa de probabilidade é estimada pelo gráfico de barras.

No caso em estudo, a População é a variável aleatória $X_1 + X_2 + X_3$, cujas realizações foram simuladas, levando à obtenção de uma amostra de dimensão 5 000, a partir da qual se construiu o gráfico de barras. Apresenta-se a seguir um gráfico com a função massa de probabilidade e a sua estimativa, o gráfico de barras:



O gráfico anterior evidencia a simetria da função massa de probabilidade, assim como o valor de 1,5 degraus para valor médio da variável aleatória em estudo.

Nota – Embora a variável não apresente todos os valores inteiros entre -15 e 18, eles devem figurar no gráfico de barras e na função massa de probabilidade, pois só assim se obtém a informação correta, tanto para a distribuição da variável, como para a distribuição dos dados.

A grande mais-valia do diagrama de barras feito a partir de uma amostra recolhida de uma determinada população, de dados discretos, é a possibilidade de sugerir um modelo de probabilidade para a população, se este não for conhecido. No caso presente, foi fácil obter esse modelo de probabilidade, dando ocasião para verificar que o diagrama de barras seria uma boa estimativa para a função massa de probabilidade.