PL 1

Probabilidade: conceitos base

Palavras chave: probabilidade, estimação de probabilidade, experiência aleatória, espaço de amostragem, eventos, casos favoráveis, simulação, Matlab.

Responda às seguintes questões utilizando sempre o Matlab para efetuar os cálculos necessários:

1. Considere a experiência aleatória de lançar **3 vezes** uma moeda equilibrada. Pretende-se estimar por simulação a probabilidade de se obter **2 caras** no fim dos **3 lançamentos** .

   Para estimar a probabilidade por simulação, é necessário executar várias vezes a experiência aleatória de lançar 3 vezes uma moeda equilibrada e calcular a percentagem de vezes em que o resultado deu 2 caras. Em Matlab, uma forma possível de implementar este simulador é a seguinte (assumindo que a experiência é executada 10000 vezes):

```
%% Código 1

% Gerar uma matriz com 3 linhas e 10000 colunas de números aleatórios
% entre 0.0 e 1.0 (ou seja, cada coluna representa uma experiência):
  experiencias = rand(3,10000);
% Gerar uma matriz com 3 linhas e 10000 colunas com o valor 1 se o valor
% da matriz anterior for superior a 0.5 (ou seja, se saiu cara) ou com o
% valor 0 caso contrário (ou seja, saiu coroa):
  lancamentos = experiencias > 0.5; % 0.5 corresponde a 1 - prob. de caras
% Gerar um vetor linha com 10000 elementos com a soma dos valores de cada
% coluna da matriz anterior (ou seja, o número de caras de cada experiência):
  resultados= sum(lancamentos);
% Gerar um vetor linha com 10000 elementos com o valor 1 quando o valor do
% vetor anterior é 2 (ou seja, se a experiência deu 2 caras) ou 0 quando é
% diferente de 2:
  sucessos= resultados==2;
% Determinar o resultado final dividindo o número de experiências com 2
% caras pelo número total de experiências:
  probSimulacao= sum(sucessos)/10000
```

   Note-se que o código proposto está desenvolvido passo a passo para mais fácil compreensão. Muitas das operações podem ser combinadas por forma a evitar o uso de matrizes intermédias e tornar a execução do código mais eficiente. Além disso, é útil usar variáveis iniciais para os parâmetros do problema para ser mais fácil alterar o código e adaptá-lo a outros casos de interesse. Assim, uma outra forma de implementar o mesmo simulador é, por exemplo, a seguinte:

```
%% Código 1 - segunda versão

N= 1e5;  %número de experiências
p = 0.5; %probabilidade de cara
k = 2;   %número de caras
n = 3;   %número de lançamentos
lancamentos = rand(n,N) > p;
sucessos= sum(lancamentos)==k;
probSimulacao= sum(sucessos)/N
```

(a) Implemente os 2 métodos no Matlab.

(b) Determine a probabilidade de obter 2 caras em 3 lançamentos de uma moeda equilibrada por cada um dos métodos (execute várias vezes a simulação).

2. Estime a probabilidade de obter 6 caras em 15 lançamentos de uma moeda equilibrada? [1]

3. Qual é a probabilidade de obter pelo menos 6 caras em 15 lançamentos de uma moeda equilibrada?

4. Para facilitar o cálculo de outras situações similares às que tratou nos pontos anteriores, crie uma função em Matlab que permita estimar a probabilidade por simulação. A função deve ter os seguintes parâmetros de entrada: p, número de lançamentos, número de caras pretendidas e número de experiências a realizar. Deve utilizar nomes adequados para a função e para os parâmetros de entrada.

   (a) Aplique a função para voltar a estimar as probabilidades das questões anteriores assim como estimar as probabilidades para todo o espaço de amostragem[2] para os seguintes números de lançamentoss: 20, 40 e 100.

   (b) Faça um gráfico, usando a função `stem`, das probabilidades estimadas no caso de 20 lançamentos.

5. Pretende-se agora calcular de forma analítica as probabilidades estimadas nos exercícios anteriores.

   Considere para isso a seguinte expressão

$$P(k) = C_k^n p^k (1-p)^{n-k}$$

   em que p é a probabilidade de ocorrer o acontecimento em que estamos interessados (por exemplo: se o acontecimento for "sair cara" em cada lançamento, $p = 0.5$), k é o número de acontecimentos que ocorreram em n repetições da experiência aleatória.

   Em Matlab, esta expressão é determinada da seguinte forma:

```
%% Codigo 2-  cálculo analítico de probabilidade em séries experiências de Bernoulli
% Dados relativos ao problema 1
p = 0.5;
k = 2;
n = 3;
prob= nchoosek(n,k)*p^k*(1-p)^(n-k); % nchoosek(n,k)= n!/(n-k)!/k!
```

   Calcule o valor analítico para cada um dos problemas anteriores e compare os resultados obtidos com os valores estimados.

6. Considere um processo de produção fabril que produz torneiras em que a probabilidade de cada torneira ser produzida com defeito é de 30%. No processo de controlo de qualidade, é selecionada uma amostra de 5 peças.

   (a) Calcule analiticamente e por simulação a probabilidade de 3 peças da amostra serem defeituosas.

   (b) Calcule analiticamente e por simulação a probabilidade de, no máximo, 2 das peças da amostra serem defeituosas.

   (c) Baseado em simulação, construa no Matlab o histograma representativo da distribuição de probabilidades do número de peças defeituosas da amostra.

[1] Adapte o Código 1 e o Código 2 de forma apropriada para resolver as questões de 2 a 4

[2] O espaço de amostragem, S, é o conjunto de todos os resultados possíveis para a experiência aleatória. Por exemplo, no problema 1 $S = \{0, 1, 2, 3\}$.

Para revisão, responda, sozinho, às seguintes questões através de simulação e aplicando análise combinatória:

R1 Quantas sequências diferentes de 10 bits há? E de n bits?

R2 Quantas sequências diferentes de 10 símbolos do alfabeto (A,C,G,T) há? E de n símbolos do mesmo alfabeto?

R3 Um teste tem n perguntas com respostas possíveis Verdadeiro ou Falso. Forneça uma expressão para calcular o número de maneiras diferentes de responder ao teste. Qual a probabilidade de acertar todas as respostas, escolhendo-as à sorte com igual probabilidade?

R4 Quantas chaves distintas pode ter o Totoloto antigo (5 números em 49)? E o Euromilhões (5 números em 50 e duas estrelas em 11)?

R5 Considere um baralho com 20 cartas. Dessas cartas, 10 são vermelhas e numeradas de 1 a 10. As restantes 10 são pretas e também numeradas de 1 a 10.

(a) De quantas maneiras diferentes se podem dispor as 20 cartas numa fila?

(b) Calcule a probabilidade de se obter uma sequência constituída alternadamente por cartas pretas e vermelhas.

R6 Lançam-se dois dados e toma-se nota da soma de pontos obtida.

(a) Indique o espaço de amostragem (conjunto de valores possíveis) da soma.

(b) Calcule a probabilidade de se obter a soma 9.

R7 Um conjunto de 50 peças contém 8 peças defeituosas. Escolhem-se aleatoriamente 10 peças. Qual a probabilidade de encontrar 3 defeituosas?

R8 Quantas *passwords* diferentes se podem obter nas seguintes situações:

(a) comprimento 5 e cada posição contendo um dígito entre 0 e 9;

(b) comprimento 5 e cada posição contendo uma letra mínuscula sem acentos.

(c) Qual a probabilidade de acertar em cada um dos dois casos anteriores escolhendo uma password aleatoriamente ?

(d) Qual a alteração ao valor destas probabilidades de fizermos 3 tentativas completamente independentes?

R1 — Para cada bit temos 2 algarismos possíveis, logo uma sequência com 10 bits tem 2^{10} possibilidades diferentes e uma sequência com n bits tem 2^n possibilidades.

R2 — Para cada sequência temos 4 símbolos diferentes, logo uma sequência com 10 símbolos tem 4^{10} possibilidades diferentes e uma sequência com n símbolos tem 4^n possibilidades diferentes

R3 — Para cada pergunta temos 2 possibilidades de resposta, logo para um teste com n perguntas podemos respondê-lo de 2^n maneiras diferentes. Para acertar a todas as perguntas, temos apenas 1 maneira possível, logo a probabilidade de acertar todas as perguntas é de $\frac{1}{2^n}$.

R4 — No totoloto escolhemos 5 números dos 49, logo temos $C^{49}_5 = 1906884$ chaves distintas.
No euromilhões escolhemos 5 números dos 50 e 2 estrelas das 11, logo temos $C^{50}_5 \times C^{11}_2 = 2118760 \times 55 = 116531800$ chaves distintas.

R5 — **a)** As 20 cartas podem permutar entre si, logo existem $20!$ maneiras diferentes.

b) Como casos favoráveis a essa situação temos $2 \times 10! \times 10!$, logo a probabilidade pedida é de $\frac{2 \times 10! \times 10!}{20!}$

R6 — **a)** $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

b) Existem 4 situações em que a soma dos dados é 9. A primeira situação é quando o 1º dado é 3 e o 2º dado é 6, a segunda situação é quando o 1º dado é 6 e o 2º dado é 6, a terceira situação é quando o 1º dado é 4 e o 2º dado é 5 e a terceira situação é quando o 1º dado é 5 e o 2º dado é 4.
Num total de situações possíveis temos $6 \times 6 = 36$, logo a probabilidade pedida é de $\frac{4}{36} = \frac{1}{9}$.

R7 — $P(\text{"defeito"}) = \frac{8}{50} = \frac{4}{25}$

Usando a lei binomial temos que a probabilidade pedida é $C^{10}_3 \times P(\text{"defeito"})^3 \times (1 - P(\text{"defeito"}))^{10-3} = C^{10}_3 \left(\frac{4}{25}\right)^3 \left(\frac{21}{25}\right)^7 = 0{,}145043$

R8 — **a)** Cada posição tem 10 opções de algarismo, logo temos 10^5 situações diferentes

b) Cada posição tem 26 opções de letras, logo temos 26^5 situações diferentes

c) Cada password apenas tem uma maneira de estar certa, logo no caso da alínea a) a probabilidade é de $\frac{1}{10^5}$ e no caso da alínea b) é $\frac{1}{26^5}$.

d) Utilizando a lei binomial, para a alínea a) temos que a probabilidade é $C^3_1 \left(\frac{1}{10^5}\right)^1 \left(1 - \frac{1}{10^5}\right)^2$ e para a alínea b) é $C^3_1 \left(\frac{1}{26^5}\right)^1 \left(1 - \frac{1}{26^5}\right)^2$.