

**Questão 1.** A partir do pacote `dslabs`, leia o conjunto de dados `heights`:

```
library(dslabs)
data(heights)
```

Este conjunto de dados apresenta informações das alturas (em polegadas) de homens e mulheres.

- (a) Crie uma nova coluna, chamada `alt.cm`, no data frame `heights` para representar a altura de cada indivíduo em centímetros (1 polegada = 2.54 cm).
- (b) Qual a proporção de homens neste conjunto de dados?
- (c) Qual a maior altura em centímetros? Esta altura é de um homem ou de uma mulher?
- (d) Qual a menor altura em centímetros? Esta altura é de um homem ou de uma mulher?
- (e) Calcule a média e o desvio padrão da variável `alt.cm`.
- (f) Divida o data frame por sexo, isto é, crie a partir de `heights` dois novos data frames. Um data frame conterá apenas as mulheres e o outro apenas os homens. Em seguida, calcule a média e o desvio padrão da altura em centímetros para cada sexo. Compare os resultados com aqueles em (e).
- (g) Apresente na mesma janela os boxplots de `alt.cm` por sexo. Interprete os resultados.

**Questão 2.** Vovó Juju e seu neto Jorel estão apostando num jogo de cara ou coroa. Uma moeda é lançada: se sair cara, Vovó Juju vence e recebe 1 real de Jorel; se sair coroa, Jorel vence e Vovó Juju paga 1 real para seu neto. No início do jogo, Vovó Juju tem 18 reais e Jorel tem 7 reais. O jogo termina quando algum dos jogadores perde todo o seu dinheiro. Estime, via Monte Carlo, a probabilidade de Vovó Juju terminar o jogo sem dinheiro.

**Questão 3.** Considere uma variável discreta  $X$  que possui a seguinte lei de probabilidade:

$$P(X = i) = \frac{1}{6}, \quad i = 1, 2, 3, 4, 5, 6.$$

- (a) Crie uma função chamada `gerador` cuja entrada seja  $n$  e cuja saída seja um vetor com  $n$  valores gerados de  $X$  a partir da distribuição uniforme (`runif`) em  $[0, 1]$ .
- (b) A variável  $X$  dessa questão pode ser interpretada como o valor da face superior de um dado após seu lançamento. Considere o seguinte experimento: um dado será lançado; se sair 6 consideremos isso como um "fracasso"; caso contrário consideremos o resultado como "sucesso". O dado será lançado até que se obtenha o segundo fracasso. O número  $N$  de sucessos que foi obtido antes do segundo fracasso é registrado. Exemplo: suponha que a sequência de valores obtidos foi 1,1,6,3,1,2,4,6; então, nesse caso,  $N = 7$ . Utilize a função criada em (a) com  $n = 1$  para estimar, via Monte Carlo, o valor de  $E[N]$ , isto é, a quantidade média de sucessos obtidos antes do segundo fracasso.

**Pontuação:** Q1: 10 pontos, Q2: 10 pontos, Q3: 15 pontos.