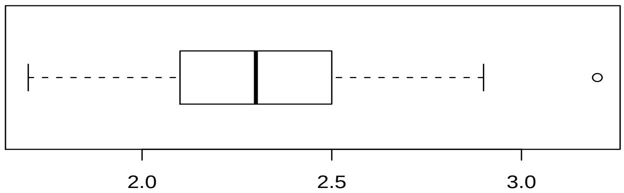
Instituto Superior de Agronomia
Soluções do 1º Teste de Estatística (2020/2021)

Nota: Não se trata de uma resolução, apenas se apresenta as soluções de cada pergunta, por vezes com um pouco de explicação para apoio

1. Seja x a variável "peso seco (g) dos mirtilos"; $n = 20$

a) Para a construção do *boxplot* são necessários os seguintes indicadores:
$\min(x_i) = 1.7$; $Q_1 = 2.1$; $Q_2 = 2.3$; $Q_3 = 2.5$ e $\max(x_i) = 3.2$
BI=$Q_1 - 1.5(Q_3 - Q_1) = 1.5$ e BS=$Q_3 + 1.5(Q_3 - Q_1) = 3.1$
Como BS> $\max(x_i)$ e olhando para os dados ordenados vemos que há 1 candidato a *outlier* à direita, que é 3.2
Os "'bigodes" esquerdo e direito são: 1.7 e 2.9, respectivamente.



b) $\overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = 2.325$ g; $s_x^2 = \dfrac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} = 0.1346$ g^2

c) A característica a que estamos a referir-no é o quantil de ordem 0.30. $Q_{0.3}^* = 2.15$

d) Sejam x' e x'' referentes aos pesos da 1ª amostra de 20 mirtilos e da 2ª amostar de 10 mirtilos, respectivamente;
$\overline{x}_{30} = \dfrac{\sum_{i=1}^{20} x_i' + \sum_{i=1}^{10} x_i''}{30} = 2.4167$ g;
$$s_x^2 = \frac{30\left(\sum_{i=1}^{20} x_i'^2 + \sum_{i=1}^{10} x_i''^2\right) - \left(\sum_{i=1}^{20} x_i' + \sum_{i=1}^{10} x_i''\right)^2}{29 * 30} = 0.1521 \text{ g}^2$$

2. $n = 25$; $y-$ diâmetro médio (em cm) $x-$altura média (em m)

a) Pelo output $r = 0.8914$ (próximo de 1, associado a uma relação linear perfeita) e a nuvem de pontos apresenta-se próxima de uma recta, logo é admissível a existência de uma relação linear entre y e x.

b) $y = 1.9288 + 0.9814\,x$.

c) O coeficiente de regressão é o declive da recta $= 0.9814$, que significa que, ao aumento de 1 metro na altura média das árvores corresponde, em média, um aumento de 0.9814 cm no diâmetro médio das árvores.

d) x em metros e $x' = 100x$ em cm; $b_1' = r\dfrac{s_y}{s_x'} = r\dfrac{s_y}{100 s_x} = \frac{1}{100} r\dfrac{s_y}{s_x} = r\dfrac{b_1}{100} = 0.009814$ cm/cm

3. Sejam os seguintes acontecimentos: TP–"teste positivo" e PI–"pessoa está infectada"

a) $P(TP) = P(PI) * P(TP|PI) + P(\overline{PI}) * P(TP|\overline{PI}) = 0.265$

b) $P(\overline{PI}|TP) = 0.679$

c) $P(errar) = P(\overline{PI} \cap TP) + P(PI \cap \overline{TP}) = 0.195$

4. As v.a. na pergunta são: $X-$ designando a mortaliddae das árvores; $Y-$ designando a espécie arbórea Vamos completar o quadro da função massa de probabilidades conjunta, colocando as probabilidades marginais

	Y			
X	0	1	2	
0	0.35	A	B	0.35+A+B
1	0.05	0.25	0.1	0.4
	0.4	A+0.25	B+0.1	1

a) Então 0.35+A+B=0.6, logo A+B=0.25
$P(Y=1|X=0)=\dfrac{P(X=0,Y=1)}{P(X=0)}=\dfrac{A}{0.6}=0.25 \Longrightarrow A=0.15 \Longrightarrow B=0.1$

b) X e Y são independentes se e só se $P[X=x_i,Y=y_j]=P[X=x_i]P[Y=y_j]$, $\forall(x_i,y_j)$.
Como $P[X=0,Y=0]=0.35$ e $P[X=0]=0.6$ e $P[Y=0]=0.4$, não se verifica a igualdade para este par, portanto a espécie arbórea e a mortalidade das árvores não são independentes.

c) A distribuição de probabilidade de $Y|X=1$, é assim dada:

$$\left\{\begin{array}{ccc} 0 & 1 & 2 \\ \dfrac{0.05}{0.4} & \dfrac{0.25}{0.4} & \dfrac{0.1}{0.4} \end{array}\right.$$

5. X e Y v.a's contínuas

a) Então F_X é contínua, logo $lim_{x\to 2^-}F_X(x)=lim_{x\to 2^+}F_X(x)=F_X(2)$ e $lim_{x\to 4^-}F_X(x)=lim_{x\to 4^+}F_X(x)=F_X(4)$, donde $0=2a-1$ e $4a-1=1$, portanto $a=1/2$

b) $f_X(x)=\left\{\begin{array}{cc} F'_X(x) & \forall x \text{ onde existe derivada} \\ 0 & \text{o.v. de } x \end{array}\right. \Longrightarrow f_X(x)=\left\{\begin{array}{cc} \frac{1}{2} & 2<x<4 \\ 0 & \text{o.v. de } x \end{array}\right.$

c) $E[X]=3$ g/100 ml e a mediana, $\chi_{0.5}$, é o valor vx_o tal que $F_X(x_o)=0.5$, donde $x_0=\chi_{0.5}=$3 g/100 ml.

d) $P[2.5<X<3]=1/4$

e) X e Y são independentes se e só se $f(x,y)=f_X(x)\times f_y(y),\ \forall(x,y)\in\mathbb{R}^2$

i) então

$$f(x,y)=\left\{\begin{array}{cc} \frac{1}{8}(\frac{y}{2}-1) & 2<x<4,\ 2<\leq y\leq 6 \\ 0 & \text{o.v. de } (x,y) \end{array}\right.$$

ii) $P[Y<X]=\int\int_A f(x,y)dxdy$, com $A=\{(x,y)\in\mathbb{R}: 2<x<4;\ 2\leq y<x\}$
$P[Y<X]=\int_2^4\left(\int_2^x \frac{1}{8}(\frac{y}{2}-1)dy\right)dx=\dfrac{1}{12}$

6.

a) $\sum_{i=1}^n(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^n(x_iy_i-x_i\overline{y}-\overline{x}y_i+\overline{x}\,\overline{y})=\sum_{i=1}^n x_iy_i-\overline{y}\sum_{i=1}^n x_i-\overline{x}\sum_{i=1}^n y_i+\sum_{i=1}^n\overline{x}\,\overline{y}=\sum_{i=1}^n x_iy_i-\overline{y}n\overline{x}-\overline{x}n\overline{y}+n\overline{x}\,\overline{y}=\sum_{i=1}^n x_iy_i-n\overline{x}\,\overline{y}$

b) i) Sendo X v.a. discreta, por definição de valor esperado tem-se
$E[a-bX]=\sum_{i=1}^n[(a-bx_i)p_i]=\sum_{i=1}^n a\,p_i-\sum_{i=1}^n b\,x_i\,p_i=a\sum_{i=1}^n p_i-b\sum_{i=1}x_i\,p_i=a+bE[X]$

ii) Por definição de variância e a propriedade do valor médio demonstrada atrás tem-se
$Var[a-bX]=E\left[\{(a-bX)-(a-bE[X])\}^2\right]=E\left[\{-b(X-E[X])\}^2\right]=b^2E\left[(X-E[X])^2\right]=b^2Var[X]$.