

Instituto Superior de Agronomia
Estatística – 2ª Chamada de Exame (2020/2021)

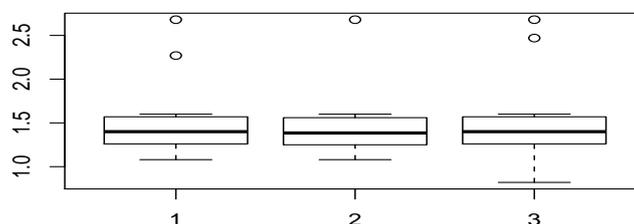
Duração: 3 horas

21 de Janeiro de 2021

1. (4 val.) Um produtor entregou um lote de abóboras manteiga na central de comercialização, tendo sido calibrado um conjunto de 50 frutos, seleccionados ao acaso. O calibrador introduziu os dados no *software* \mathbb{R} , no objecto `massa` que designa o vector com as massas (em quilograma) dos 50 frutos analisados:

```
> sort(massa)
[1] 1.08 1.20 1.21 1.22 . . . 1.58 1.59 1.60 2.27 2.68
> mean(massa)
[1] 1.4637
> coeficiente_variação #valor obtido após aplicação da fórmula de cálculo.
[1] 21.5182
> quantile(massa, type=2)
 0% 25% 50% 75% 100%
1.08 1.26 1.40 1.57 2.68
```

- a) Identifique e classifique, justificando, a variável em estudo.
b) Calcule a variância da amostra analisada.
c) Dos três *boxplots* apresentados na figura, qual o que poderá corresponder à amostra analisada? Justifique, apresentando os cálculos que necessitou de efectuar.



- d) Considere que a empresa pretende apenas comercializar frutos com massas no intervalo $[1.2, 1.6]$ kg.
- Determine uma estimativa da proporção de frutos do produtor com massa no intervalo $[1.2, 1.6]$ kg.
 - Usando uma análise estatística adequada, critique a seguinte afirmação do produtor: “Mais de 85% das abóboras têm massa no intervalo $[1.2, 1.6]$ kg. Justifique convenientemente.
2. (2.5 val.) Seja $\{(x_i, y_i), i = 1, \dots, 15\}$, uma amostra bivariada de dimensão 15. Determinou-se a recta dos mínimos quadrados entre x e y assim dada $y = 3.5 + 0.1x$.
- Alguém comentou: “O coeficiente de correlação é próximo de zero porque o declive da recta tem um valor baixo”. O que pensa deste comentário? Justifique.
 - Sendo $\sum_{i=1}^{15} x_i = 157.5$ e $\sum_{i=1}^{15} x_i^2 = 1750.75$, determine:
 - $\sum_{i=1}^{15} y_i$;
 - $cov(x, y)$.
 - Suponha que foi efectuada a transformação afim, $x' = kx$, k real não nulo. Mostre que b'_1 , o declive da recta associada aos pares $\{(x'_i, y_i)\}$, se exprime em função do declive da recta dada como $b'_1 = 0.1/k$.

3. (3.5 val.) Sabe-se que um povoamento florestal misto é caracterizado pela seguinte proporção de espécies: 20% de pinheiro manso, 25% de carvalho cerquinho e 55% de pinheiro bravo. Nesta região o ataque da lagarta processionária (*Thaumetopoea pityocampa*) verifica-se em 40% dos pinheiros mansos e 80% dos pinheiros bravos, não havendo ataque nos carvalhos cerquinhos.

- a) Escolhida uma árvore ao acaso neste povoamento, qual a probabilidade de:
 - i) ser pinheiro bravo e estar atacada com processionária?
 - ii) estar atacada com processionária?
- b) Nas árvores daquele povoamento não atacadas por processionária, qual a proporção de árvores de cada espécie?
- c) Escolhidas aleatoriamente 100 árvores neste povoamento:
 - i) Qual é o número de árvores atacadas que se espera encontrar?
 - ii) Qual a probabilidade, aproximada, de nessas 100 árvores encontrar 40 ou mais atacadas? Justifique convenientemente.

4. (4 val.) Considere a variável aleatória contínua X cuja função densidade é assim definida:

$$f_X(x) = \begin{cases} 4x(1-x^2) & \text{se } 0 \leq x \leq 1, \\ 0 & \text{outros valores de } x. \end{cases}$$

- a) Determine a função distribuição cumulativa de X .
- b) Calcule o valor esperado de X . Será que a mediana de X é superior ao valor esperado?
- c) Considere o par aleatório contínuo (X, Y) no qual X é a v.a. com densidade dada acima. A função densidade conjunta do par é assim definida:

$$f(x, y) = \begin{cases} 8xy & \text{se } 0 \leq x \leq y \leq 1, \\ 0 & \text{outros valores de } (x, y). \end{cases}$$

- i) Esboce a região do plano onde $f(x, y)$ toma valores não nulos.
- ii) Serão X e Y v.a. independentes? Justifique.
- iii) Determine $P[Y > 1/2 | X = 1/4]$.

5. (3.5 val.) Um negociante compra pinhas de pinheiro manso para obter pinhão negro (pinhão com casca). O preço do kg da pinha é 0.90 euros. Admita que o peso de uma pinha segue uma distribuição normal de valor médio 264 g.

- a) Sabendo que a probabilidade de uma pinha ter menos de 300 g é 0.7019, calcule o desvio padrão do peso de uma pinha.
Nota: Nas respostas seguintes utilize o valor de desvio padrão igual a 68 gramas.
- b) Qual a probabilidade de uma pinha, escolhida ao acaso, ter mais do que 400 g?
- c) As pinhas são colocadas, ao acaso, em sacas que se consideram cheias com 50 pinhas. Responda às seguintes questões, justificando convenientemente.
 - i) Caracterize a distribuição da variável aleatória que representa o peso do conteúdo de uma saca cheia.
 - ii) Qual a probabilidade de o preço de uma saca cheia ser inferior a 12.5 euros?

6. (2.5 val.) Uma fábrica localiza-se junto de dois rios, A e B. Na sua laboração é utilizado um agente químico altamente prejudicial quando libertado para o meio ambiente. Realizou-se uma acção de fiscalização em que se capturaram 12 peixes em cada um dos rios e se mediu a concentração do agente poluente em cada peixe. Os dados foram introduzidos no \mathbb{R} , nos vectores A e B e executaram-se alguns comandos, cujos resultados se encontram no Anexo. Utilize os resultados que considere adequados para responder às seguintes alíneas.

- Indique uma estimativa do valor médio e uma estimativa da variância da concentração do agente poluente nos peixes do rio A.
- Escreva a expressão dos estimadores associados às estimativas obtidas em a).
- Um ambientalista afirma que a concentração média de poluente por peixe no rio A é inferior à existente no rio B. Comente esta afirmação, indicando e justificando todos os pressupostos necessários à sua resposta.

Anexo

```
> mean(A)      >mean(B)      >var(A)      >var(B)
[1] 11.1       [1] 12          [1] 0.6345455  [1] 0.3654545

> shapiro.test(A)      > shapiro.test(B)
Shapiro-Wilk normality test      Shapiro-Wilk normality test

data:  A                      data:  B
W = 0.9376, p-value = 0.4678      W = 0.9113, p-value = 0.2219

> shapiro.test(A-B)
Shapiro-Wilk normality test

data:  A - B
W = 0.9657, p-value = 0.8606

> var.test(A,B)

      F test to compare two variances

data:  A and B
F = 1.7363, num df = 11, denom df = 11, p-value = 0.374
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.4998471 6.0314476
sample estimates:
ratio of variances
 1.736318

> t.test(A,B,var.equal=TRUE,paired=FALSE,alternative="two.sided")

      Two Sample t-test

data:  A and B
t = -3.1177, df = 22, p-value = 0.005013
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.4986756 -0.3013244
sample estimates:
mean of x mean of y
 11.1      12.0

>
> t.test(A,B,var.equal=TRUE,paired=FALSE,alternative="greater")
```

(v.s.f.f)

Two Sample t-test

```
data: A and B
t = -3.1177, df = 22, p-value = 0.9975
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -1.395697      Inf
sample estimates:
mean of x mean of y
  11.1      12.0

>
> t.test(A,B,var.equal=TRUE,paired=FALSE,alternative="less")
```

Two Sample t-test

```
data: A and B
t = -3.1177, df = 22, p-value = 0.002507
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf -0.4043031
sample estimates:
mean of x mean of y
  11.1      12.0

>
> t.test(A,B,paired=TRUE,alternative="less")
```

Paired t-test

```
data: A and B
t = -2.5379, df = 11, p-value = 0.01379
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf -0.2631377
sample estimates:
mean of the differences
      -0.9

> t.test(A,B,paired=TRUE,alternative="greater")
```

Paired t-test

```
data: A and B
t = -2.5379, df = 11, p-value = 0.9862
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -1.536862      Inf
sample estimates:
mean of the differences
      -0.9

> t.test(A,B,paired=TRUE,alternative="two.sided")
```

Paired t-test

```
data: A and B
t = -2.5379, df = 11, p-value = 0.02758
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.6805203 -0.1194797
sample estimates:
mean of the differences
      -0.9
```