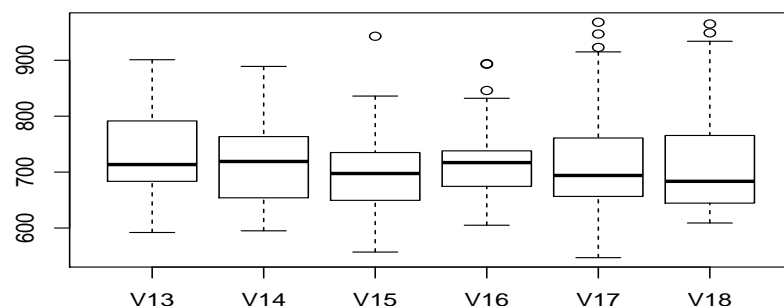


1. Considere os dados `rice` contidos na biblioteca `agricolae` do `R`, correspondentes à produção da variedade de arroz IR8 (g/m<sup>2</sup>) para estudo de uniformidade de terras. A área de cultivo tem 18×36 metros. Os dados encontram-se num *data frame* com 36 registos de observações efectuadas em 18 variáveis V1-V18. Nesta questão vamos considerar apenas seis variáveis V13-V18.

Considere o *output* que se encontra no **Anexo** no qual se estudaram as variáveis escolhidas e se apresentam alguns comandos que foram executados no `R`. Sempre que possível use os resultados para responder às seguintes questões:

- Esboce o **histograma**, indicado no *output* mas não desenhado. Que alterações devia introduzir no comando para ele ser desenhado?
- Interprete os resultados que se obtiveram com o comando `basicStats` (V18).
- Indique os cálculos efectuados para a elaboração do *boxplot* para a variável V18. Indique claramente os limites do *boxplot*.
- Compare os *boxplots* apresentados abaixo.
- Face aos resultados apresentados no *output* poder-se-á dizer que a variável V13 apresenta, em média, valores superiores a V15? Justifique convenientemente.



2. a) Pretende-se estudar o peso de coelhos criados numa propriedade. Para isso recolheu-se uma amostra de 80 coelhos naquela propriedade. Considere alguns resultados apresentados em **Anexo**.
- Dê uma estimativa pontual do peso médio dos coelhos naquela propriedade. Com 95% de confiança, qual o limite máximo do erro cometido ao estimar o verdadeiro peso médio dos coelhos pela estimativa indicada?
  - Poder-se-á dizer que mais de 20% dos coelhos têm peso superior ou igual a 5.5 kg? Justifique convenientemente.
- b) Admita que o peso de coelhos criados nessa propriedade pode ser considerado tendo uma distribuição normal com valor médio 5 kg e desvio padrão 0.8 kg.
- Qual a percentagem de coelhos que pesam mais de 7 kg?
  - Um comerciante vai comprar 5000 coelhos e pretende classificá-los de acordo com o peso da seguinte forma: 20% dos leves como pequenos, os 55% seguintes como médios, os 15% seguintes como grandes e os 10% mais pesados como extras. Quais os limites de peso para cada classificação?

3. Seja  $X$  uma v.a. que representa o peso, em kg, das uvas produzidas pelas videiras de uma dada qualidade, que se admite ter distribuição aproximadamente normal,  $X \sim \mathcal{N}(\mu, \sigma)$ , com  $\mu = 60$  kg e  $\sigma = 10$  kg.
- Suponha que pretende obter, por simulação, 50 valores daquele modelo. Quais os comandos em **R**, necessários para obter esses valores?
  - Escreva os comandos necessários para calcular, com recurso ao **R**,  $P[55 < X < 70]$ .
  - Seja  $Z$  a v.a. normal reduzida associada a  $X$ . Quais os comandos em **R** para obter os gráficos da função densidade de  $Z$  e de  $Z^2$ ?
4. Considere que se tem uma amostra aleatória de dimensão  $n$ ,  $(X_1, X_2, \dots, X_n)$ , retirada de uma população  $X$ , com função densidade assim definida, onde  $\theta > -1$  é um parâmetro desconhecido:

$$f(x; \theta) = \begin{cases} \frac{\theta + 1}{e^{\theta+1}} x^\theta & \text{se } 0 \leq x \leq e \\ 0 & \text{outros valores de } x \end{cases}$$

**Nota:** Sabe-se que  $E[X] = \frac{(\theta + 1)e}{\theta + 2}$ .

- Obtenha o estimador de  $\theta$  pelo método dos momentos.
- Obtenha o estimador de máxima verosimilhança para  $\theta$ .
- Considere os valores observados da seguinte amostra, de dimensão 30, extraída daquela população, com a qual se realizaram os cálculos apresentados:

```
> dados
[1] 2.13 2.26 1.32 1.88 1.63 2.27 2.65 2.60 1.80 2.00
[11] 1.85 1.54 0.63 2.47 1.67 2.14 1.47 2.30 2.02 2.57
[21] 1.77 2.40 1.96 2.33 2.12 1.62 0.82 2.39 2.57 2.59

> sum(dados)          > sum(log(dados))
[1] 59.77              [1] 19.39522
```

- Determine estimativas para  $\theta$ .
- Determine uma estimativa de máxima verosimilhança para  $E[X]$ .

5. Indique, justificando se as seguintes afirmações são **Verdadeiras** ou **Falsas**. **Corrija as falsas**.

- Seja  $(X_1, \dots, X_n)$  uma amostra aleatória de tamanho  $n$ , proveniente de uma população com valor médio  $\mu$  e variância  $\sigma^2 < +\infty$ . Então  $E[X_i X_j] = \mu^2$  quaisquer que sejam  $i \neq j$ .
- $\mu$ ,  $\bar{X}$  e  $\bar{x}$  são formas diferentes de designar a mesma quantidade: a média da população em estudo.
- Seja  $X \sim \text{Normal}(0, 1)$  então  $5X \sim \text{Normal}(5, 5)$ .
- Seja  $(X_1, \dots, X_n)$  uma amostra aleatória de tamanho  $n$  e  $X_{(n)} = \max(X_i)$ . Tem-se  $P[X_{(n)} \leq x] = nP[X_1 \leq x]$ , qualquer que seja  $x$ .

## ANEXO I

##### Pergunta 1 #####

```
> library(agricolae)
> data(rice)
> rice
> attach(rice)
```

```
> hist(V15,breaks=c(500,560,620,680,740,800,950),plot=F) #HISTOGRAMA
```

\$breaks

```
[1] 500 560 620 680 740 800 950
```

\$counts

```
[1] 1 4 10 13 5 3
```

\$density

```
[1] 0.0004629630 0.0018518519 0.0046296296 0.0060185185 0.0023148148
[6] 0.0005555556
```

\$mids

```
[1] 530 590 650 710 770 875
```

```
> library(fBasics)
```

```
> basicStats(V18)
```

	V18
nobs	36.000000
NAs	0.000000
Minimum	609.000000
Maximum	965.000000
1. Quartile	644.750000
3. Quartile	764.750000
Mean	724.277778
Median	683.500000
Sum	26074.000000
SE Mean	17.855665
LCL Mean	688.028850
UCL Mean	760.526705
Variance	11477.692063
Stdev	107.133991
Skewness	1.030894

```
> boxplot(V13,V14,V15,V16,V17,V18,names=c("V13","V14","V15","V16","V17","V18"))
```

```
> shapiro.test(V13)
```

Shapiro-Wilk normality test

data: V13

W = 0.9685, p-value = 0.3865

```
>shapiro.test(V15)
```

Shapiro-Wilk normality test

data: V15

W = 0.9568, p-value = 0.1705

```
> var.test(V13,V15)
```

F test to compare two variances

data: V13 and V15

F = 1.0022, num df = 35, denom df = 35, p-value = 0.9949

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.5110389 1.9653902

sample estimates:

ratio of variances

1.002193

```
> t.test(V13,V15,var.equal=TRUE)
```

#### Two Sample t-test

```
data: V13 and V15
```

```
t = 1.8916, df = 70, p-value = 0.06269
```

```
alternative hypothesis: true difference in means is not equal to 0
```

```
95 percent confidence interval:
```

```
-1.932271 72.987826
```

```
sample estimates:
```

```
mean of x mean of y
```

```
733.5833 698.0556
```

```
> t.test(V13,V15,var.equal=TRUE,alternative="greater")
```

#### Two Sample t-test

```
data: V13 and V15
```

```
t = 1.8916, df = 70, p-value = 0.03134
```

```
alternative hypothesis: true difference in means is greater than 0
```

```
95 percent confidence interval:
```

```
4.219346 Inf
```

```
sample estimates:
```

```
mean of x mean of y
```

```
733.5833 698.0556
```

```
> t.test(V13,V15,paired=TRUE,alternative="greater")
```

#### Paired t-test

```
data: V13 and V15
```

```
t = 2.7517, df = 35, p-value = 0.004664
```

```
alternative hypothesis: true difference in means is greater than 0
```

```
95 percent confidence interval:
```

```
13.71336 Inf
```

```
sample estimates:
```

```
mean of the differences
```

```
35.52778
```

```
> detach(rice)
```

```
##### Pergunta 2 #####
```

```
> mean(peso)
```

```
[1] 4.96125
```

```
> var(peso)
```

```
[1] 0.8973402
```

```
> sort(peso)
```

```
[1] 1.9 3.1 3.2 3.3 3.5 3.6 3.7 3.7 3.8 4.0 4.0 4.1 4.1 4.2 4.2  
[16] 4.2 4.3 4.3 4.4 4.4 4.4 4.4 4.4 4.5 4.5 4.5 4.6 4.6 4.7 4.7  
[31] 4.7 4.8 4.8 4.8 4.8 4.8 4.8 4.9 4.9 4.9 4.9 5.0 5.0 5.1 5.1  
[46] 5.1 5.2 5.2 5.3 5.3 5.3 5.3 5.3 5.4 5.4 5.4 5.4 5.4 5.4  
[61] 5.5 5.5 5.5 5.6 5.6 5.7 5.7 5.7 5.9 6.0 6.0 6.1 6.5 6.5 6.5  
[76] 6.7 6.8 6.8 7.0 7.0
```

```
>
```