

# ESTADÍSTICA UOC – PRIMAVERA 2018

## EXAMEN 1

### EXERCICI 1

S'ha preguntat a les alumnes i als alumnes d'una classe pel temps (en minuts) que triguen a arribar des de la seva casa fins a la Universitat. Les respostes es recullen en aquesta taula:

TEMPS (MINUTS)	[0, 5)	[5, 10)	[10, 15)	[15, 20)	[20, 25)
N.º D'ALUMNES/AS	10	6	9	3	2

A partir de la marca de classe,  $m_i$ , de cada interval:

- Calculeu la mitjana del temps (en minuts) que triguen a arribar.
- Calculeu la seva desviació típica i la variància poblacional.
- Representeu gràficament la variable estudiada.

Observació: és necessari escriure les fòrmules i els passos intermedis a cada apartat.

*Criteris de puntuació i valoració (sobre 10): a) 3 punts; b) 4 punts; c) 3 punts.*

### SOLUCIÓ:

Trobem la marca de classe,  $m_i$ , de cada interval i fem la taula de freqüències:

INTERV	$m_i$	$n_i$	$n_i m_i$	$n_i m_i^2$
[0, 5)	2,5	10	25	62,5
[5, 10)	7,5	6	45	337,5
[10, 15)	12,5	9	112,5	1406,25
[15, 20)	17,5	3	52,5	918,75
[20, 25)	22,5	2	45	1012,5
		30	280	3737,5

a) Mitjana:

$$\bar{x} = \frac{\sum n_i m_i}{N} = \frac{280}{30} = 9,33$$

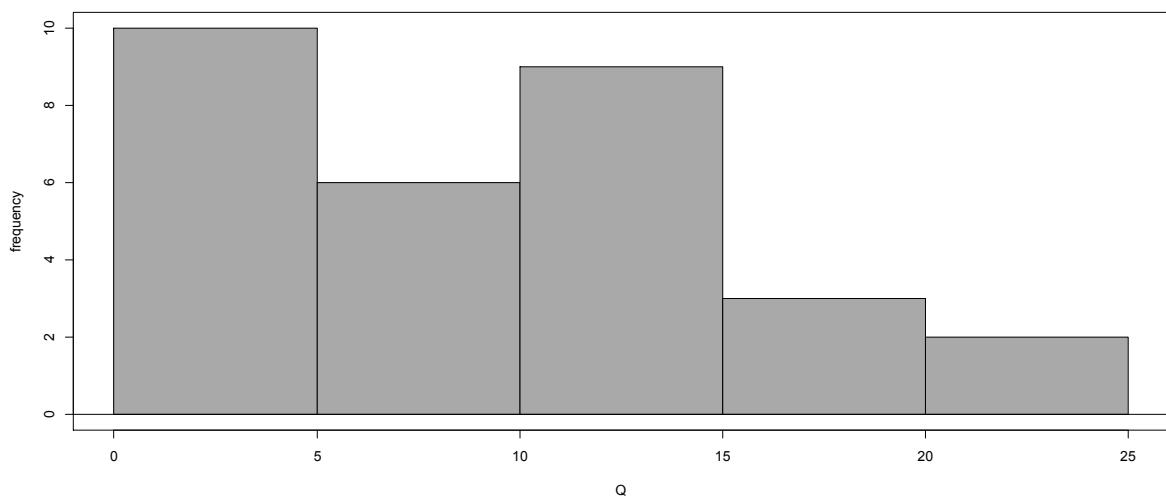
b) Desviació típica i variància poblacional:

$$S_x^2 = \frac{\sum n_i m_i^2}{N} - \bar{x}^2 = \frac{3737,5}{30} - 9,33^2 = 37,53$$

$$S_x = \sqrt{37,53} = 6,13$$

Els alumnes i les alumnes triguen, de mitjana, 9,33 minuts, amb una desviació típica de 6,13 minuts.

c)



---

## EXERCICI 2

---

En una classe de 30 alumnes hi ha 18 que han aprovat matemàtiques, 16 que han aprovat anglès i 6 que no han aprovat cap de les dues.

Triem a l'atzar un alumne d'aquesta classe:

- Quina és la probabilitat que hagi aprovat anglès?
- Quina és la probabilitat que hagi aprovat anglès i matemàtiques?
- Sabent que ha aprovat matemàtiques, quina és la probabilitat que hagi aprovat anglès?
- Són independents els successos "Aprovar matemàtiques" i "Aprovar anglès"?

Observació: us pot ajudar organitzar les dades en una taula de contingència.

*Criteris de puntuació i valoració (sobre 10): a), b) i c) 2 punts cadascun; d) 4 punts.*

**SOLUCIÓ:**

Organitzem les dades en una taula de contingència, completant els que falten:

	APROVEN MATEMÀTIQUES	NO APROVEN MATEMÀTIQUES	
APROVEN INGLÉS	10	6	16
NO APROVEN INGLÉS	8	6	14
	18	12	30

Sigui  $M$  = "Aprova matemàtiques",  $I$  = "Aprova anglès".

a) Quina és la probabilitat que hagi aprovat anglès?

$$P[I] = \frac{16}{30} = 0,53$$

b) Quina és la probabilitat que hagi aprovat anglès i matemàtiques?

$$P[M \cap I] = \frac{10}{30} = \frac{1}{3} = 0,33$$

c) Sabent que ha aprovat matemàtiques, quina és la probabilitat que hagi aprovat anglès?

$$P[I/M] = \frac{10}{18} = \frac{5}{9} = 0,56$$

d) Són independents els successos "Aprovar matemàtiques" i "Aprovar anglès"?

$$P[M] \cdot P[I] = \frac{18}{30} \cdot \frac{16}{30} = \frac{3}{5} \cdot \frac{8}{15} = \frac{24}{75} = \frac{8}{25}$$

$$P[M \cap I] = \frac{1}{3} \neq \frac{8}{25}$$

No són independents, ja que:

$$P[M \cap I] \neq P[M] \cdot P[I]$$

### EXERCICI 3

Un perillós virus informàtic ataca una carpeta que conté 100 fitxers. El virus contamina de

manera independent a uns fitxers i a d'altres. Cada fitxer té una probabilitat de 0,045 de ser contaminat.

- a) Quina és la probabilitat que més d' 1 fitxer estiguin contaminats del virus?
- b) Quina és la probabilitat que exactament 9 fitxers estiguin contaminats del virus?
- c) Un programa antivirus comença a examinar un per un els fitxers de la carpeta. Quina és la probabilitat que el primer fitxer que trobi contaminat sigui el cinquè en ser examinat?

A cadascun dels apartats cal indicar quina és la variable aleatòria que estudiem, quina és la seva distribució i els càlculs que es fan.

*Criteris de puntuació i valoració (sobre 10): a) 1 punt per indicar la llei i 2 punts la probabilitat b) 1 punt per indicar la llei i 2 punts la probabilitat i c) 1 punt per indicar la llei i 3 punts la probabilitat.*

**SOLUCIÓ:**

a) Sigui  $X$ =”nombre de fitxers contaminats”.  $X$  segueix una distribució binomial de paràmetres:  $n=100$  i  $p=P[\text{un fitxer estigui infectat}]=0,045$ . Per tant,

$$P(X > 1) = 1 - P(X \leq 1) = 1 - (P(X = 0) + P(X = 1)) = 1 - (0,010 + 0,047) = 0,943$$

b) De nou hauríem d'aplicar la distribució binomial:

$$P[X = 9] = 0,021801$$

c) Si ara anomenem:  $Y$ =”nombre de fitxers examinats fins el primer contaminat”, resulta que  $Y$  segueix una distribució geomètrica:

$$P[Y = 5] = (1 - 0,045)^4 \cdot 0,045 = 0,037430$$

#### **EXERCICI 4**

D'una mostra aleatòria de 200 alumnes d'un grau mitjà es va observar que 110 eren dones. Determineu un interval de confiança al 95% per la proporció de dones entre els alumnes d'aquest grau. Quina ha de ser la mida de la mostra si volem tenir un interval per la proporció amb un 95% de confiança i amb un marge d'error menor que 0.05?

Valors de probabilitats que us poden ser útils; si no trobeu exactament el que necessiteu, useu el més proper

$p(X \geq x)$	$X \sim N(0,1)$	$X \sim t$ de Student amb 9 graus de llibertat	$X \sim t$ de Student amb 10 graus de llibertat
0,01	2,326	2,821	2,763
0,025	1,96	2,262	2,228
0,05	1,645	1,833	1,812
0,005	2,575	3,249	3,169

*Criteris de puntuació i valoració: Identificar l'estadístic de contrast i la seva llei 3 punts, plantejar i obtenir l'interval correcte 4 punts, plantejar correctament i obtenir la mida de la mostra 3 punts.*

**SOLUCIÓ:**

$$\text{Sigui } \hat{p} = \frac{110}{200} = 0,55;$$

$$\text{i } s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0,55 * 0,45}{200}} = 0,035.$$

L'estadístic de contrast serà

$$\frac{\hat{p}(1-p)}{\sqrt{\frac{p(1-p)}{n}}}$$

que segueix una distribució normal estàndard. Aleshores si  $Z$  és una normal estàndard imosem  $P(-z \leq Z \leq z) = 0,95$  i s'obté  $z = 1,96$ .

L'interval de confiança és

$$\begin{aligned} (\hat{p} - 1,96s_{\hat{p}}, \hat{p} + 1,96s_{\hat{p}}) &= (0,55 - (1,96 * 0,035); 0,55 + (1,96 * 0,035)) \\ &= (0,55 - 0,068; 0,55 + 0,068) = (0,481; 0,618) \end{aligned}$$

D'altra banda, hem de determinar  $n$  tal que:

$$n \geq z^2 \frac{p(1-p)}{ME^2}$$

Però com que  $p$  és desconegut agafem

$$n \geq z^2 \frac{1}{4ME^2} = 1,96^2 \frac{1}{4(0,05)^2} = 384,16$$

Per tant,  $n \geq 385$ .

## EXERCICI 5

Volem estudiar el temps que triguen els estudiants de la Universitat A (variable UniA) a arribar a la universitat des de casa seva i el que triguen els de la Universitat B (variable UniB). Amb R s'han obtingut els següents resultats (en els quals s'ha perdut el p-valor):

```
## 
## Two Sample t-test
##
## data: UniA and UniB
## t = 0.51002, df = 298, p-value = XXXXX
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.9229512 1.5686886
## sample estimates:
## mean of x mean of y
## 100.3616 100.0388
```

També sabem que

```
pt(c(0.50, 0.51, 0.52), 298)
```

```
## [1] 0.6912780 0.6947855 0.6982752
```

i es demana:

a) Indiqueu quin contrast s'ha usat i quines són les seves hipòtesis.

- b) Trobeu el p-valor.
- c) Expliqueu clarament quina conclusió s'obté partint del p-valor i relacioneu-la amb l'interval de confiança.
- Criteris de puntuació i valoració (sobre 10): Indicar contrast i hipòtesis 4 punts. p-valor 3 punts. Conclusió 3 punts.*

### SOLUCIÓ:

- a) És un contrast de diferència de mitjanes amb variàncies iguals i suposant normalitat. És bilateral  $\mu_0 = \mu_1, \mu_0 \neq \mu_1$  amb un nivell de significació del 10%.
- b) El p-valor és  $(1-0,6947849)*2= 0,6104302$
- c) Com el p-valor és major que el nivell de significació no podem dir que les mitjanes siguin diferents. Es confirma aquesta conclusió ja que l'interval de confiança conté el valor 0.

### EXERCICI 6

S'està realitzant un estudi sobre el cost d'imprimir 100 prospectes segons el seu nombre de pàgines i s'han obtingut les següents dades:

Nombre de pàgines	50	60	70	75
Cost	700	780	1015	1068

- a) Calculeu la recta de regressió del cost en funció del nombre de pàgines, indicant les fórmules i els càlculs realitzats.
- b) Interpreteu els coeficients obtinguts.
- c) Quin seria el cost per imprimir 100 prospectes de 65 pàgines?
- Criteris de puntuació i valoració (sobre 10): a) Fórmules i càlculs 5 punts. b) 3 punts. c) 2 punts.*

### SOLUCIÓ:

- a) Tenim que  $\bar{pag} = 63,75$ ;  $\bar{cost} = 890,75$ ;  $s_{pag} = 11,0875$ ;  $s_{cost} = 178,4253$ ;  $s_{pag-cost} = 1936,25$  i per tant  $\hat{\beta}_1 = \frac{1936,25}{11,0875^2} = 15,7526$  i  $\hat{\beta}_0 = 890,75 - 15,7526 * 63,75 = -113,47825$ . La recta de regressió és  $cost = -113,47825 + 15,7526 * pag$

- b)  $\hat{\beta}_1 = 15,7526$  indica que per cada pàgina addicional el cost de 100 prospectes augmenta en 15,7526 euros.  $\hat{\beta}_0 = -113,47825$  indica que el cost d'imprimir 100 prospectes de 0 pàgines seria -113,47825 que en aquest cas no té sentit econòmic.
- c) Serà  $cost = -113,47825 + 15,7526 * 65 = 910,4407$

# EXAMEN 2

---

## EXERCICI 1

---

Tirem successivament una moneda i anotem el nombre de llançaments que necessitem fins a obtenir per primera vegada cara. Realitzem l'experiment 100 vegades, amb els següents resultats:

LLANÇAMENT EN EL QUE SURT CARA	1	2	3	4	5	6
N.º DE VEGADES QUE HA OCORREGUT	48	25	16	4	5	2

- a) Calculeu la mitjana.
- b) Calculeu la seva desviació típica i la variància mostral.
- c) Calculeu la mediana i els quartils.
- d) Feu el diagrama de caixes de la variable.

Observació: és necessari escriure les fòrmules i els passos intermedis a cada apartat.

*Criteris de puntuació i valoració (sobre 10): a) 2 punts; b) 3 punts; c) 3 punts; d) 2 punts.*

## SOLUCIÓ:

- a) Mitjana

$$\bar{x} = \frac{\sum n_i x_i}{N} = 1,99$$

- b) La variància i la desviació típica mostral són:

$$S_x^2 = \frac{\sum n_i x_i^2}{N-1} - \bar{x}^2 = 1,585758$$
$$S_x = \sqrt{1,585758} = 1,259269$$

- c) Fem la taula de freqüències acumulades:

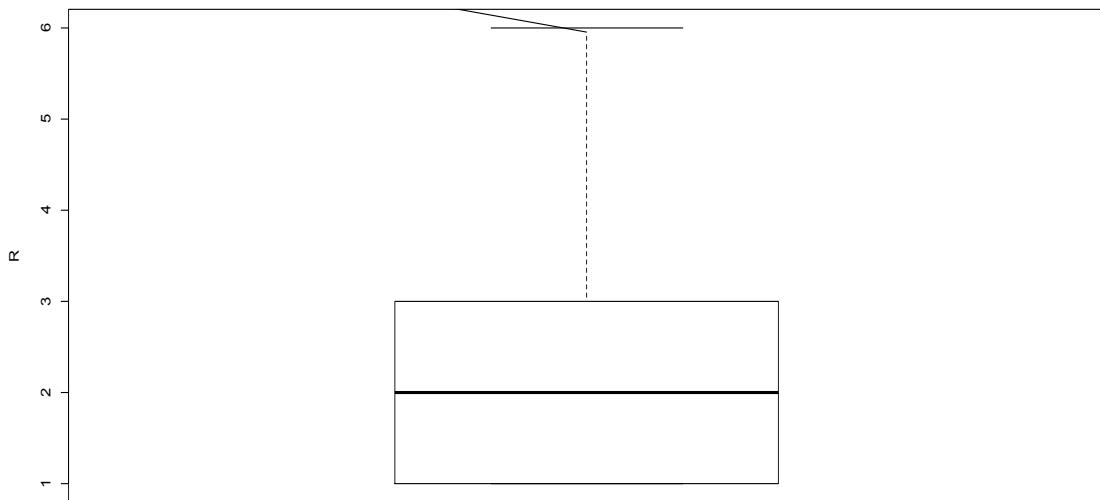
$x_i$	$n_i$	$F_i$	en %
1	48	48	48
2	25	73	73
3	16	89	89
4	4	93	93
5	5	98	98
6	2	100	100

$Me = 2$  perquè per a  $x_i = 2$ , la  $F_i$  supera el 50 %.

$Q_1 = 1$  perquè per a  $x_i = 1$ , la  $F_i$  supera el 25 %.

$Q_3 = 3$  perquè per a  $x_i = 3$ , la  $F_i$  supera el 75 %.

d) El Boxplot ( Min.: 1.00, Max.: 6.00 )



## EXERCICI 2

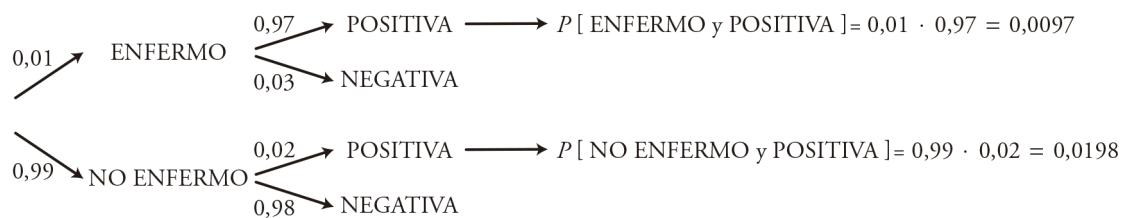
L'1 % de la població d'un determinat lloc pateix una malaltia. Per detectar-la es realitza una prova de diagnòstic. Aquesta prova dona positiva en el 97 % dels pacients que pateixen la malaltia; en el 98 % dels individus que no la pateixen dona negativa. Triem un individu a l'atzar.

- Representeu l'arbre de probabilitats
- Quina és la probabilitat que la prova doni positiva i pateixi la malaltia?
- Si sabem que ha donat positiva, quina és la probabilitat que pateixi la malaltia?
- Els successos “donar positiva” i “patir malaltia”, són successos independents? Per què?

Criteris de puntuació i valoració (sobre 10): a) 4 punts; b), c) i d) 2 punts cadascun.

## SOLUCIÓ:

a)



b)  $P(\text{MALALT I POSITIVA}) = 0,0097$

c)

$$P[\text{ENFERMO} / \text{POSITIVA}] = \frac{P[\text{ENFERMO y POSITIVA}]}{P[\text{POSITIVA}]} = \frac{0,0097}{0,0097 + 0,0198} = \frac{0,0097}{0,0295} = 0,33$$

d) No són independents, ja que:

$$P(\text{MALALT I POSITIVA}) \neq P(\text{MALALT}) \cdot P(\text{POSITIVA})$$

$$P(\text{MALALT I POSITIVA}) = 0,0097$$

$$P(\text{MALALT}) \cdot P(\text{POSITIVA}) = 0,01 \cdot 0,0295 = 0,000295$$

---

### EXERCICI 3

---

El 20% del clients que arriben a una benzinera entren a la botiga de la benzinera a comprar alguna cosa. Suposem que cada client actua de manera independent de la resta.

- a) Diem X al nombre de clients que arriben a la benzinera fins que n'arriba un que no entra a la botiga a comprar res. Quina llei segueix la variable X? Calculeu  $P(X > 2)$ ,  $P(X = 2 | X < 2)$ ,  $P(X = 10)$  i  $P(X = 0,5)$ .
- b) Considerem 6 clients agafats a l'atzar. Quin és el nombre esperat de clients que entraran a la botiga a comprar alguna cosa?
- c) Si el temps que tarda en arribar el primer client a la benzinera segueix una exponencial de paràmetre 0,1 (en minuts), quina és la probabilitat que després d'obrir s'hagi d'esperar més de 5 minuts l'arribada del primer client?

A cadascun dels apartats cal indicar quina és la variable aleatòria que estudiem, quina és la seva distribució i els càlculs que es fan.

*Criteris de puntuació i valoració: a) Trobar la llei de X 1 punt i cada probabilitat 1 punt, b) plantejar correctament l'esperança 1 punt i calcular-la 1 punt, c) plantejar correctament la probabilitat 1 punt i calcular-la 2 punts.*

### SOLUCIÓ:

- a) Sigui X="número de clients que arriben fins que n'arriba un que no entra a la botiga" es una geomètrica de paràmetre 0.8. Ens demanen

$$P(X > 2) = 1 - P(X \leq 2) = 1 - P(X = 1) - P(X = 2) = 1 - 0,8 - 0,2(0,8) = 0,04$$

$$P(X = 2 | X < 2) = \frac{P(X = 2)}{P(X < 2)} = \frac{0,2(0,8)}{0,8 + 0,2(0,8)} = \frac{0,2}{1,2} = 0,166$$

$$P(X = 0,5) = 0$$

$$P(X = 10) = 0,2^9 0,8 = 0,00000041.$$

- b) Sigui Y="número de clients entre 6 escollits a l'atzar que entren a la botiga a comprar" és una binomial de paràmetres 6, 0.2. Ens demanen

$$E(Y) = 6 * 0,2 = 1,2$$

- c) Sigui Z una variable exponencial de paràmetre 10. Aleshores hem de calcular:

$$P(Z > 5) = \int_5^\infty 0,1 \exp(-0,1x) dx = -\exp(-0,1x)]_5^\infty = \exp(-0,5) = 0,606$$

---

### EXERCICI 4

---

Tenim una nevera per conservar uns medicaments a zero graus. Prenem la temperatura durant quatre dies i obtenim les següents temperatures: 0,1; 0,4; -0,4 i 0,3.  
Suposem normalitat de les dades.

- Determineu un interval de confiança al 90% per la temperatura. A quina conclusió podem arribar?
- Si fem un interval al 95%, obtindrem un interval més gran o més petit?

Valors de probabilitats que us poden ser útils; si no trobeu exactament el que necessiteu, useu el més proper.

$p(X \geq x)$	$X \sim N(0,1)$	$X \sim t$ de Student amb 4 graus de llibertat	$X \sim t$ de Student amb 3 graus de llibertat
0,01	2,326	3,764	4,540
0,025	1,96	2,776	3,182
0,05	1,645	2,131	2,350
0,005	2,575	4,604	5,840

Criteris de puntuació i valoració: identificar l'estadístic de contrast i la seva llei 3 punts, plantejar i obtenir l'interval correcte 3 punts, arribar a conclusions correctes 2 punts, raonar correctament sobre la mida de l'interval 2 punts.

### SOLUCIÓ:

a) Sabem que la variable aleatòria

$$\frac{\bar{X} - \mu}{\sqrt{\frac{s}{n}}}$$

segueix una distribució t Student amb  $n-1=3$  graus de llibertat.

$$P(-t \leq t_3 \leq t) = 0.90, \text{ s'obté } t = 2.35.$$

L'interval de confiança és així

$$\left( \bar{X} - t \sqrt{\frac{s}{n}}, \bar{X} + t \sqrt{\frac{s}{n}} \right) = \left( 0,1 - 2,35 \sqrt{\frac{0,35}{4}}, 0,1 + 2,35 \sqrt{\frac{0,35}{4}} \right) \\ = (0,1 - 0,695; 0,1 + 0,695) = (-0,595; 0,795).$$

No hi ha evidències per dir que la nevera no manté la temperatura de zero graus.

b) Si agafem un interval al 95%, obtindrem un interval més gran.

### EXERCICI 5

S'ha preguntat a 4 estudiants d'una classe pel temps en minuts que invertien en arribar des de casa seva fins a la Universitat, abans d'instal·lar aparcaments de bicicletes i després d'instal·lar-los. S'han obtingut els següents resultats

Abans	50	40	23	15
Després	50	30	22	16

Es desitja saber si la instal·lació d'aquests aparcaments fa reduir els temps de desplaçament amb un nivell de significació de  $\alpha=0,2$ . Plantegeu el corresponent contrast i expliqueu clarament a quina conclusió s'arriba.

Valors de probabilitats que us poden ser útils; si no trobeu exactament el que necessiteu, useu el més proper.

$p(X \geq x)$	$X \sim N(0,1)$	$X \sim t$ de Student amb 3 graus de llibertat	$X \sim t$ de Student amb 4 graus de llibertat
0,2	0,84	0,97	0,94
0,1	1,28	1,63	1,53
0,4	0,25	0,27	0,27
0,05	1,64	2,53	2,13

*Criteris de puntuació i valoració (sobre 10): Plantejar el contrast 4 punts. Fórmules i càlculs 3 punts.  
Conclusió 3 punts.*

### SOLUCIÓ:

S'ha de fer un test t de dades aparellades; hem de contrastar  $\mu_d = 0, \mu_d > 0$ ; l'estadístic de contrast és  $t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} = \frac{2.5}{5.66228/2} = 0,9869$  amb tres graus de llibertat. El valor més proper de la taula és 0,97 i per tant el p-valor una mica més petit que 0,2 pel que acceptarem que s'ha reduït el temps de desplaçament.

---

### EXERCICI 6

---

S'està realitzant un estudi sobre el preu unitari de compra d'uns prospectes (*pre*) en funció de quants se n'imprimeixen (*imp*). Amb R s'han obtingut els següents resultats en els quals s'ha perdut algun nombre:

```
## 
## Call:
## lm(formula = pre ~ imp)
## 
## Residuals:
##      Min    1Q   Median    3Q   Max 
## -3.7220 -1.7560  0.3621  1.5442  4.6438 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 12.074887  0.698676 17.283   <2e-16 ***
## imp          XXXXXXXXX  0.002175 -0.744    0.459  
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 2.361 on 97 degrees of freedom
## Multiple R-squared:  0.005677, Adjusted R-squared:  -0.004574 
## F-statistic: 0.5538 on 1 and 97 DF,  p-value: 0.4586
```

a) Calculeu la recta de regressió del preu unitari dels prospectes sabent que el preu esperat per a una impressió de 65 prospectes és 11,969652€; useu quatre decimals.

b) Interpreteu els coeficients obtinguts.

c) Calculeu  $R^2$  i indiqueu si és un bon ajust.

*Criteris de puntuació i valoració (sobre 10): a) Fórmules i càlculs 4 punts. b) 3 punts. c) 3 punts.*

**SOLUCIÓ:**

- a) El pendent és  $-0,001619$  ja que  $12,074887 + a \cdot 65 = 11,969652$ . Per tant la recta de regressió seria  $pre = 12,074887 - 0,001619 \cdot imp$ .
- b)  $\hat{\beta}_1 = -0,001619$  indica que per cada prospectes addicional el preu unitari disminueix en  $-0,001619$  euros.  $\hat{\beta}_0 = 12,074887$  indica que el preu de 0 prospectes seria de  $12,074887$  €.
- c)  $R^2 = 0,00567$  i, per tant, l'ajust és molt dolent.

# EXAMEN 3

---

## EXERCICI 1

---

Un periodista esportiu anota els gols marcats per un equip local en cadascuna de les 30 últimes jornades. Els resultats es resumeixen en aquesta taula:

<b>N.º DE GOLS</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
<b><math>F_{\text{ABSOLUTA}}</math></b>	<b><math>x</math></b>	<b>10</b>	<b>9</b>	<b>2</b>	<b><math>y</math></b>
<b><math>F_{\text{RELATIVA}}</math></b>	<b>0,2</b>	<b>0,33</b>	<b>0,3</b>	<b>0,07</b>	<b><math>z</math></b>

- a) Completeu la taula calculant  $x, y, z$ .
- b) Quina és la mitjana de gols marcats en cada jornada?
- c) Calculeu la seva desviació típica i variància mostral.
- d) Realitzeu el diagrama de barres.

Observació: és necessari escriure la formula i els passos intermedis a cada apartat

Criteris de puntuació i valoració (sobre 10): a) 3 punts; b) 2 punts; c) 3 punts; d) 2 punts.

### SOLUCIÓ:

- a) La freqüència relativa és el quocient entre la freqüència absoluta i el nombre total de jornades (30, en el nostre cas).

$$0,2 = \frac{x}{30} \rightarrow x = 0,2 \cdot 30 = 6 \rightarrow x = 6$$

$$6 + 10 + 9 + 2 + y = 30 \rightarrow y = 3$$

$$z = \frac{y}{30} = \frac{3}{30} = 0,1 \rightarrow z = 0,1$$

b)

N.º DE GOLS	$n_i$	$n_i x_i$
0	6	0
1	10	10
2	9	18
3	2	6
4	3	12
	30	46

$$\bar{x} = \frac{46}{30} = \frac{23}{15} \approx 1,53$$

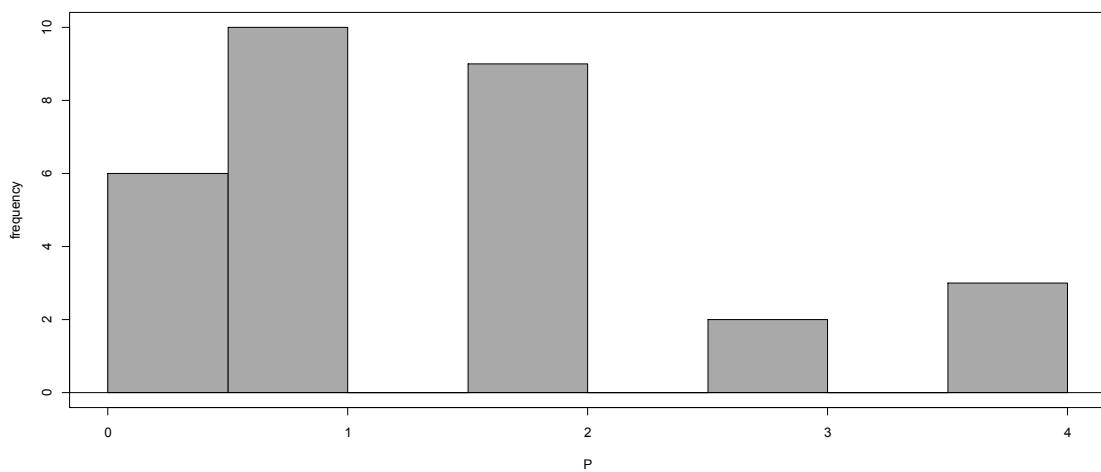
La mitjana de gols marcats en cada jornada és 1,53.

c) Variància i desviació típica mostra

$$S_x^2 = \frac{\sum n_i x_i^2}{N-1} - \bar{x}^2 = 1,429885$$

$$S_x = \sqrt{1,585758} = 1,195778$$

d)




---

## EXERCICI 2

---

En una aerolínia hi ha 300 empleats: 25 pilots, 80 ajudants de pilot, i la resta, auxiliars de vol. De tots ells, solament a 15 pilots, 50 ajudants de pilot i 75 auxiliars de vol els agrada viatjar. Si triem un empleat a l'atzar, calcula les següents probabilitats:

- a)  $P$  [sigui pilot i no li agradi viatjar]
- b)  $P$  [sigui auxiliar de vol sabent que no li agrada viatjar]
- c)  $P$  [li agrada viatjar sabent que és ajudant de pilot]
- d) Els successos “ser pilot” i “no li agrada viatjar”, són independents? Raoneu la resposta.  
Observació: us pot ajudar organitzar les dades en una taula de contingència.

*Criteris de puntuació i valoració (sobre 10): a), b) i c) 2 punts cadascun; d) 4 punts.*

### SOLUCIÓ:

Organitzem les dades en una taula de contingència

	PILOTS	AJUDANTS DE PILOT	AUXILIARS DE VOL	
ELS AGRADA VIATJAR	15	50	75	140
NO ELS AGRADA VIATJAR	10	30	120	160
	25	80	195	300

a)

$$P(\text{pilot} \cap \text{no li agrada viatjar}) = \frac{10}{300} = 0,033$$

b)

$$P(\text{auxiliar de vol} | \text{no li agrada viatjar}) = \frac{120}{160} = 0,75$$

c)

$$P(\text{li agrada viatjar} | \text{ajudant de pilot}) = \frac{50}{80} = 0,625$$

d) Els successos “ser pilot” i “no li agrada viatjar”, són independents? Raoneu la resposta.

$$P(\text{pilot} \cap \text{no li agrada viatjar}) \neq P(\text{pilot}) \cdot P(\text{no li agrada viatjar})$$

$$P(\text{pilot} \cap \text{no li agrada viatjar}) = \frac{10}{300} = 0,033$$

$$P(\text{pilot}) \cdot P(\text{no li agrada viatjar}) = \frac{25}{300} \cdot \frac{160}{300} = \frac{1}{12} \cdot \frac{8}{15} = \frac{8}{180} = 0,044$$

Per tant no són independents.

### EXERCICI 3

Tirem 3 daus perfectes alhora i diem X al nombre de múltiples de 3 que obtenim.

- a) Calculeu la funció de massa de probabilitat de X.
- b) Calculeu  $P(-1 < X < 1)$  i  $P(X < 2 | X > 0,5)$ .
- c) Calculeu l’esperança de X .

Criteris de puntuació i valoració: a) 5 punts, b) cada probabilitat 1 punt, c) plantejar correctament l’esperança 1 punt i calcular-la 2 punts.

### SOLUCIÓ:

a)

Els possibles resultats són 0, 1, 2 i 3 i per tant

x	P(X=x)
0	$\left(\frac{2}{3}\right)^3 = \frac{8}{27}$
1	$3\left(\frac{2}{3}\right)^2 \frac{1}{3} = \frac{12}{27}$
2	$3 \frac{2}{3} \left(\frac{1}{3}\right)^2 = \frac{6}{27}$
3	$\left(\frac{1}{3}\right)^3 = \frac{1}{27}$

b)

$$P(-1 < X < 1) = P(X = 0) = \frac{8}{27}$$

$$P(X < 2 | X > 0.5) = \frac{P(0.5 < X < 2)}{P(X > 0.5)} = \frac{P(X = 1)}{1 - F(0.5)} = \frac{\frac{12}{27}}{1 - \frac{8}{27}} = \frac{12}{19}$$

$$\text{c) A partir d'aquí: } E(X) = 0 * \frac{8}{27} + 1 * \frac{12}{27} + 2 * \frac{6}{27} + 3 * \frac{1}{27} = \frac{27}{27} = 1.$$

#### EXERCICI 4

Estem mesurant el temps de reacció d'uns determinats components químics. Repetim l'experiment 300 vegades i mesurem 300 vegades el temps de reacció, obtenint una mitjana mostra de 100 s i una desviació estàndard mostra és de  $5\sqrt{3}$ s. Suposant que els temps segueixen una distribució normal,

- a) Calculeu un interval de confiança per a la mitjana dels temps de reacció, amb un nivell de confiança del 90%.
- b) Quina hauria de ser la mida de la mostra perquè la longitud de l'interval de confiança fos inferior a 1?

Valors de probabilitats que us poden ser útils; si no trobeu exactament el que necessiteu, useu el més proper.

p(X>= x)	X~ N(0,1)	X~ t de Student amb 299 graus de llibertat	X~ t de Student amb 300 graus de llibertat
0,01	2,326	2,339	2,339
0,025	1,96	1,967	1,968
0,05	1,645	1,65	1,65
0,005	2,575	2,592	2,592

Criteris de puntuació i valoració (sobre 10): S'han d'indicar les fórmules i els càlculs realitzats, així com els raonaments. a) 6 punts: plantejament 2 punts, càlculs 4 punts; b) 4 punts (plantejament 2 punts, càlculs 2 punts).

#### SOLUCIÓ:

a) És un interval de confiança de la mitjana d'una distribució normal amb la desviació estàndard desconeguda:  $\left(\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}\right) = \left(100 \pm t_{0.05, 299} \frac{5\sqrt{3}}{\sqrt{300}}\right) = \left(100 \pm 1.65 \frac{5\sqrt{3}}{\sqrt{300}}\right) = (99.175, 100.825)$

b) Si la longitud és 1, aleshores el marge d'error és 0.5 i tenim que

$$n \geq \left(t_{\frac{\alpha}{2}, n-1}\right)^2 \left(\frac{s}{0.5}\right)^2 = 1.65^2 \times (10\sqrt{3})^2 = 816,75$$

i per tant  $n = 817$

## EXERCICI 5

S'ha fet un estudi sobre com es desplacen els estudiants de la universitat A i s'ha obtingut que d'una mostra aleatòria de 200 estudiants, 40 es desplacen amb bicicleta. En canvi, a la universitat B, d'una mostra de 300 estudiants, 63 es desplacen amb bicicleta. Es desitja saber si la proporció d'ús de la bicicleta pels estudiants de la universitat A és diferent que la de la B, amb un nivell de significació de  $\alpha=0,2$ . Plantegeu el contrast, trobeu el p-valor i el valor crític del contrast i raoneu les conclusions obtingudes.

Valors de probabilitats que us poden ser útils; si no trobeu exactament el que necessiteu, useu el més proper

$p(X \geq x)$	$X \sim N(0,1)$	$X \sim t$ de Student amb 500 graus de llibertat
0,2	0,84	0,82
0,1	1,28	1,283
0,4	0,25	0,253
0,05	1,64	1,647

Criteris de puntuació i valoració (sobre 10): Plantejar el contrast 2 punts. Fórmules i càlculs 4 punts. Valor crític 2 punts. Conclusió 2 punts.

### SOLUCIÓ:

S'ha de contrastar  $p_0 = p_1, p_0 \neq p_1$  i l'estadístic de contrast és  $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.2 - 0.21}{\sqrt{0.206(0.794)\left(\frac{1}{200} + \frac{1}{300}\right)}} = -0.270861177$ . El valor crític serà aproximadament 1.28 pel que no podem dir que les proporcions siguin diferents.

## EXERCICI 6

S'està realitzant un estudi sobre el preu de lloguer (*prea* per metre quadrat) dels locals comercials d'una certa zona en funció del preu (*pred* també per metre quadrat) abans de fer reformes. Amb R s'han obtingut els següents resultats en els quals s'ha perdut algun nombre:

```
## Call:
## lm(formula = pred ~ prea)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -2.25685 -1.06082 -0.06082  0.96369  2.08621 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) XXXXXXXX  0.67661   1.032    0.305    
## prea        0.95099   0.05633  16.883   <2e-16 ***  
## ---
```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.351 on 97 degrees of freedom
## Multiple R-squared:  0.7461, Adjusted R-squared:  0.7435
## F-statistic:  285 on 1 and 97 DF,  p-value: < 2.2e-16

```

- a) Calculeu la recta de regressió del preu després de la reforma en funció del preu anterior sabent que el preu esperat per a un local pel qual es pagava 13 € per metre quadrat abans de la reforma és, després de la reforma, de 13,0608 € per metre quadrat.
- b) Interpreteu els coeficients obtinguts.
- c) Podem assegurar que el pendent de la recta de regressió és diferent de zero, amb un nivell de significació del 90%?

*Criteris de puntuació i valoració (sobre 10): a) Fórmules i càlculs 4 punts. b) 3 punts. c) 3 punts.*

### SOLUCIÓ:

- a) L'ordenada a l'origen és 0.69793 ja que  $b+0.95099*13=13.0608$ . Per tant la recta de regressió és  $\text{pred}=0.69793 +0.95099*\text{prea}$ .
- b)  $\hat{\beta}_1 = 0.95099$  indica que per cada euro de més en el preu per metre quadrat abans de les reformes, el preu després de les reformes augmenta en 0.95099.  $\hat{\beta}_0 = 0.69793$  indica que el preu d'un local que abans de la reforma era de 0 € seria després de la reforma de 0.69793 €/m<sup>2</sup>.
- c) Sí perquè el p-valor és pràcticament 0.