

EXAMEN 1

1.- La següent taula mostra el nombre de llibres que han comprat durant l'últim mes 50 estudiants de la UOC

NLlibres	NEstudiants
1	12
3	2
4	X
6	2
Y	4

Es demana:

- Calculeu X i Y sabent que la mitjana dels llibres comprats pels estudiants de la mostra és 5.
- Calculeu els quartils del nombre de llibres comprats

Criteris de puntuació i valoració (sobre 10): a) (6 punts) Càlcul d' X: 1 punt. Càlcul d'Y: 5 punts. b) (4 punts) 2 per la mediana i 2 pels altres quartils.

Solució:

- X=30, Y=25
- Q1=3, Q2=4, Q3=4

2.- Sabem que un cert projecte es pot desenvolupar en C+ o en Python i que els clients pot ser que quedin satisfets o que no. Sabent que

- La probabilitat de desenvolupar-ho en C+ és de 0.7
- La probabilitat de desenvolupar-ho en Python i que el client quedi satisfet és de 0.27
- La probabilitat que el client quedi satisfet és 0.48.

Es demana

- La probabilitat de desenvolupar el projecte en Python.
- La probabilitat que un client no quedi satisfet sabent que el projecte s'ha desenvolupat en C+.
- Raoneu si els successos “Desenvolupar el projecte en C+” i “Estar el client satisfet” són independents o no.

Criteris de puntuació i valoració (sobre 10): a) (2 punts) Identificar la probabilitat que es demana 1 punt; càlcul 1 punt. B) (5 punts) Identificar la probabilitat que es demana 1 punt; càlcul 4 punts. c) (3 punts) Identificar el que es demana 1 punt; càlculs 1 punt; conclusió 1 punt.

Solució:

- $P(\text{Python})=0.3$
- $P(\text{No satisfet} \mid \text{C+})=0.7$
- Com que $P(\text{C+})=0.7$, $P(\text{Satisfet})=0.48$ $P(\text{C+ i Satisfet})= 0.21$, clarament $P(\text{C+}) \cdot P(\text{Satisfet}) \neq P(\text{C+ i Satisfet})$ per la qual cosa no són independents.

3.- Considerem la variable aleatòria X que ens diu el nombre de vegades que un servidor web se col·lapsa en un dia. Ens diuen que la taula de la seva funció de massa de probabilitat és la següent:

x	0	1	2	3	4
p(X=x)	0.25	0.15	0.2	0.3	0.1

- Comproveu que és una funció de massa de probabilitat.
- Calculeu el valor esperat del nombre de vegades que el servidor se col·lapsa en un dia.
- Què val la probabilitat que el servidor se col·lapsi com a màxim una vegada en un dia?

Criteris de puntuació i valoració (sobre 10): a) (3 punts). Comprovació que és una funció de probabilitat, 3 punts b) (3 punts). Càlcul del valor esperat, 3 punts. b) (4 punts). Escriure bé la probabilitat demanada, 2 punts, càlcul de la probabilitat, 2 punts.

Solució:

- La funció p és una funció de massa ja que els seus valors estan entre 0 i 1 i la suma dels seus valors val 1:

$$0.25 + 0.15 + 0.2 + 0.30 + 0.1 = 1.$$

- L'esperança de X serà: $E(X) = 0 \cdot 0.25 + 1 \cdot 0.15 + 2 \cdot 0.2 + 3 \cdot 0.3 + 4 \cdot 0.1 = 1.85$ vegades.
- La probabilitat demanada serà: $p(X \leq 1) = p(X = 0) + p(X = 1) = 0.25 + 0.15 = 0.4$.

4.- Considerem l'experiment aleatori de llançar dos daus i sumar els resultats obtinguts. Considerem la variable aleatòria Y que consisteix en calcular les vegades en que la suma anterior ha estat 5 si hem repetit l'experiment 100 vegades.

- Digueu quina és la distribució de la variable Y .
- Calculeu la probabilitat aproximada que la suma dels daus hagi estat 5 entre 10 i 20 llançaments (inclosos). Indicació: Feu servir el Teorema del Límit Central.

Valors de les taules de la $N(0,1)$ (utilitzeu els que considereu necessaris, si necessiteu un valor que no és exacte, utilitzar el valor més proper)

x	p(X >= x)
2.85	0.0022
2.83	0.0023
2.8	0.0026
2.5	0.0062
2	0.0228
0.6	0.2743
0.65	0.2578
0.5	0.3085

0.55	0.2912
0.4	0.3446
0.45	0.3264
0.35	0.3632
0.3	0.3821

Criteris de puntuació i valoració (sobre 10): a) (4 punts). Donar la distribució, 1.5 punts. Donar els paràmetres del que depèn, 2.5 punts. b) (6 punts). Escriure bé la probabilitat demanada, 3 punts, càlcul de la probabilitat, 3 punts.

Solució:

a) Els casos en que la suma dels dos daus és 5 són els següents: $\{(1,4), (4,1), (2,3), (3,2)\}$. Per tant, la probabilitat d'obtenir un 5 val $p = \frac{4}{36} = \frac{1}{9}$.

La distribució de Y serà binomial de paràmetres $n = 100$ i $p = \frac{1}{9}$

b) Segons el Teorema del Límit Central, la variable Y es pot aproximar a una normal de paràmetre $\mu = n \cdot p = \frac{100}{9} \approx 11.1111$ i $\sigma = \sqrt{n \cdot p \cdot (1 - p)} =$

$\sqrt{\frac{800}{81}} \approx 3.1427$. La probabilitat demanada serà:

$$\begin{aligned}
 p(10 \leq Y \leq 20) &\approx p\left(\frac{10 - \frac{100}{9}}{\sqrt{\frac{800}{81}}} \leq Z \leq \frac{20 - \frac{100}{9}}{\sqrt{\frac{800}{81}}}\right) \\
 &= p(-0.35 \leq Z \leq 2.83) = p(Z \leq 2.83) - p(Z \leq -0.35) \\
 &= 0.9977 - 0.3618 = 0.6358,
 \end{aligned}$$

on Z representa la normal $N(0,1)$.

5.- Un proveïdor d'Internet publicita que la seva velocitat de baixada és de 256 kbps. Mesurem, en horaris diferents, la velocitat, en kbps, obtinguda per 10 usuaris d'aquest proveïdor. Obtenim el següent resultat en R:

One Sample t-test

data: adsl\$vel

t = -1.2162, df = 9, p-value = 0.1274

alternative hypothesis: true mean is less than 256

95 percent confidence interval:

-Inf 265.0181

sample estimates:

mean of x

Suposant que la velocitat segueix una distribució normal, contrasteu l'afirmació de l'empresa, al nivell de significació del 5%. Feu-ho amb el p-valor i amb l'interval de confiança. Expliciteu les hipòtesis del contrast. És coneguda la desviació típica poblacional? Per què?

Criteris de puntuació i valoració (sobre 10): Donar bé les hipòtesis, 2 punts. Conclusió correcta amb el p-valor, 3 punts. Conclusió correcta amb l'interval de confiança, 3 punts. Raonar correctament si és coneguda la desviació típica poblacional, 2 punts.

Solució:

$$H_0 : \mu = 256$$

$$H_1 : \mu < 256$$

Com el p-valor (0.1274) és més gran que el nivell de significació 0.05 acceptem que $\mu = 256$.

IC al 95% para μ és $(-\infty ; 265.0181)$ i com 256 està dins de l'interval acceptem també la hipòtesi nul·la. Això està d'acord amb la conclusió que hem tret amb el p-valor. No és coneguda la desviació típica poblacional, ja que estem fent un contrast amb la t-student.

6.- Els preus de les TV LED, d'una coneguda marca són els següents:

Mida (polzades)	15	17	19	24
Preu (euros)	251	301	357	556

- Calculeu la recta de regressió per explicar el preu a partir de la mida.
- Es tracta d'un bon model?

Criteris de puntuació i valoració (sobre 10): a) 6 punts, b) 4 punts.

Solució:

a) Les mitjanes de les variables són: $\bar{x} = 18.75$, $\bar{y} = 366.25$. Les variàncies i la covariància valen: $s_x^2 = 14.92$, $s_y^2 = 17876.92$, $s_{x,y} = 513.42$.

El pendent de la recta valdrà: $\hat{\beta}_1 = \frac{513.42}{14.92} = 34.42$.

L'ordenada en l'origen valdrà: $\hat{\beta}_0 = 366.25 - 34.42 \cdot 18.75 = -279.11$.

La recta de regressió serà: $\text{Preu} = -279.11 + 34.42 \text{ Mida}$

b) Per a estudiar la qualitat de l'ajust, calcularem el coeficient de correlació mostrat r :

$$r = \frac{s_{x,y}}{s_x \cdot s_y} = \frac{513.42}{\sqrt{14.92 \cdot 17876.92}} = 0.994. \text{ Per tant, es tracta d'un molt bon ajust.}$$

EXAMEN 2

1.- Hem preguntat a 10 estudiants de la UOC quants llibres havien comprat durant el darrer mes i hem obtingut les següents respostes (on X i Y representen unes respostes que hem perdut accidentalment):

0, 0, 1, 1, 3, 4, 5, 10, X, Y

Es demana:

- Calculeu X i Y sabent que la mitjana del nombre de llibres comprats és 3, la mediana 2.5 i $Y > 3$.
- Calculeu els quartils del nombre de llibres comprats.

Criteris de puntuació i valoració (sobre 10): a) (6 punts) Càlcul d' X: 3 punts. Càlcul d'Y: 3 punts. b) (4 punts) 2 per a cada quartil.

Solució:

- $X=2$, $Y=4$
- $Q1=1$, $Q2=2.5$, $Q3=4$

2.- La següent taula mostra el nombre de clients satisfets i insatisfets amb 100 projectes desenvolupats en tres llenguatges diferents

	Satisfets	No satisfets
C+	20	30
Python	10	Y
C	X	Z

Si sabem que

- La probabilitat que un client **no** estigui satisfet és 0.5
- La probabilitat que el projecte es desenvolupi en Python i el client no quedi satisfet és 0.1

Es demana:

- Trobeu els valors de X i Y i Z.
- Calculeu la probabilitat que un projecte es desenvolupi en C
- Calculeu la probabilitat que un client quedi satisfet si el projecte ha estat desenvolupat en C.

Criteris de puntuació i valoració (sobre 10): a) (3 punts) Un punt per cada incògnita. b) (3 punts) Identificar la probabilitat demanada 1 punt; càlcul 2 punt. c) (4 punts) Identificar el que es demana 2 punts; càlculs 2 punts.

Solució:

- a) $X=20$, $Y=10$, $Z=10$.
- b) $P(C)=0.3$.
- c) $P(S | C)=20/30$.

3.- En un servidor de correu arriben de mitjana 2 correus per minut. Suposem que l'arribada de correus al servidor és un procés de Poisson.

- a) Trobeu la probabilitat que arribin 4 missatges en 3 minuts.
- b) Quin és el nombre mitjà de missatges que arriben a aquest servidor en un dia? I la desviació típica del nombre de missatges que arriben en un dia?

Criteris de puntuació i valoració (sobre 10): a) (5 punts). Donar la distribució de la variable en qüestió, 1 punt, donar el paràmetre, 2 punts, càlcul de la probabilitat, 2 punts. b) (5 punts). Donar la distribució, 1 punt, donar el paràmetre, 1 punt, donar l'esperança, 1 punt, donar la desviació típica, 2 punts.

Solució:

- a) Considerem la variable X_3 que ens dona el nombre de missatges que arriben en 3 minuts. Ens demanen $p(X_3 = 4)$. Tenint en compte que la variable X_3 és de Poisson de paràmetre $2 \cdot 3 = 6$, la probabilitat serà: $p(X_3 = 4) = \frac{6^4}{4!} e^{-6} \approx 0.1339$.
- b) Considerem la variable X_{1440} que ens dona el nombre de missatges que arriben en un dia. Ens demanen $E(X_{1440})$. Tenint en compte que la variable X_{1440} és de Poisson de paràmetre $2 \cdot 1440 = 2880$, l'esperança i la variància seran: $E(X_{1440}) = 2880$, $Var(X_{1440}) = 2880$. La desviació típica serà, doncs, $sd_X = \sqrt{Var(X_{1440})} = \sqrt{2880} = 53.6656$.

4.- La mitjana del temps que tarda un servidor de correu en enviar un missatge segueix una distribució normal. Hem agafat una mostra de 200 missatges sortints i hem obtingut que el servidor ha tardat 10 milisegons de mitjana en enviar-los.

- a) Suposant que la desviació típica del temps que tarda en enviar un missatge sortint val $\sigma = 2.5$ milisegons, calculau l'interval de confiança per la mitjana del temps que tarda el servidor en enviar un missatge a un 95% de confiança.
- b) Suposant que la desviació típica de la mostra val 3.5 milisegons i que no coneixem la desviació típica poblacional, calculau l'interval de confiança per la mitjana del temps que tarda el servidor en enviar un missatge a un 98% de confiança.

Valors de les taules de la $N(0,1)$ i la t de Student (utilitzeu els que considereu necessaris, si necessiteu un valor que no és exacte, utilitzar el valor més proper)

$p(X \geq x)$	x ($N(0,1)$)	x $t(199)$	x $t(200)$
0.01	2.326	2.345	2.345

0.025	1.96	1.972	1.972
0.05	1.645	1.653	1.653

Criteris de puntuació i valoració (sobre 10): a) (5 punts). Donar l'expressió de l'interval de confiança, 2.5 punts. Càlcul de l'interval de confiança, 2.5 punts. b) (5 punts). Donar l'expressió de l'interval de confiança, 2.5 punts. Càlcul de l'interval de confiança, 2.5 punts.

Solució:

a) L'interval de confiança serà:

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right) = \left(10 - 1.96 \cdot \frac{2.5}{\sqrt{200}}, 10 + 1.96 \cdot \frac{2.5}{\sqrt{200}} \right) \\ = (9.6535, 10.3465).$$

b) L'interval de confiança serà:

$$\left(\bar{x} - t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}, \bar{x} + t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}} \right) = \left(10 - 2.345 \cdot \frac{3.5}{\sqrt{200}}, 10 + 2.345 \cdot \frac{3.5}{\sqrt{200}} \right) \\ = (9.4196, 10.5804).$$

5.- Hem demanat a uns usuaris que avaluïn una determinada APP. Passats uns mesos els mateixos usuaris repeteixen el procés d'avaluació. Les puntuacions s'introdueixen en R i obtenim el següent resultat:

Paired t-test

data: V1 and V2

t = -3.3489, df = 9, p-value = 0.008539

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

95 percent confidence interval:

-0.6869539 -0.1330461

sample estimates:

mean of the differences

-0.41

Expliciteu les hipòtesis del contrast. Realitzeu el contrast d'hipòtesi amb el p-valor i amb l'interval de confiança. Traieu les conclusions pertinents. Són mostres independents o aparellades?

Criteris de puntuació i valoració (sobre 10): Donar bé les hipòtesis, 2 punts. Conclusió correcta amb el p-valor, 3 punts. Conclusió correcta amb l'interval de confiança, 3 punts. Raonar correctament si són mostres independents o aparellades, 2 punts.

Solució:

Hem de contrastar les hipòtesis: $H_0: \mu_D = 0$ vs. $\mu_D \neq 0$, on D és la diferència entre les puntuacions després de passats els mesos menys abans.

Hem de fer un contrast de diferència de mitjanes aparellades (Paired t-test) ja que son les mateixes persones les que avaluen les dues vegades.

Els resultats obtinguts ens diuen que, sobre la base de les observacions registrades, hi ha una probabilitat de 0.95 que μ_D sigui un valor de l'interval (-0.687, -0.133).

A més, amb un p-valor de 0.009 també podem afirmar que hi ha indicis suficients com per descartar la hipòtesi nul·la. Per tant, el test confirma que les dues mitjanes són diferents.

Cal destacar que aquesta conclusió és coherent amb que el valor 0 no estigui inclòs en l'interval de confiança trobat per a la diferència d'ambdues mitjanes.

6.- Si considerem 12 TV LED que hi ha en una tenda, trobem la recta de regressió $Y = -250 + 30.5X$ (on X indica la mida, en polzades, i Y preu, en euros). A més sabem que $s^2 = 82$ i $\sum (x_i - \bar{x})^2 = 39$.

- Determineu un interval de confiança, al 95%, per al pendent de la recta.
- Feu el contrast d'hipòtesis sobre el pendent de la recta, amb un nivell de confiança del 95%, a partir de l'interval calculat anteriorment.

p(X>= x)	X~ N(0,1)	X~ t de Student amb 10 graus de llibertat	X~ t de Student amb 11 graus de llibertat
0.01	2.326	2.764	2.718
0.025	1.96	2.228	2.201
0.05	1.645	1.812	1.796
0.005	2.575	3.169	3.106
0.1	1.28	1.372	1.363

Criteris de puntuació i valoració (sobre 10): a) 6 punts, b) 4 punts.

Solució:

a) L'interval de confiança ve donat per:

$$(\hat{\beta}_1 - t_{\alpha/2, n-2} s_{\hat{\beta}_1}, \hat{\beta}_1 + t_{\alpha/2, n-2} s_{\hat{\beta}_1}) . \quad \text{Calculem cada element de l'interval: } \hat{\beta}_1 = 30.5,$$

$$t_{0.025, n-2} = t_{0.025, 10} = 2.23, \quad s_{\hat{\beta}_1} = \sqrt{\frac{s^2}{\sum (x_i - \bar{x})^2}} = \sqrt{\frac{82}{39}} = 1.450022. \quad \text{L'interval de}$$

confiança per al pendent serà:
 $(30.5 - 2.23 \times 1.45, 30.5 + 2.23 \times 1.45) = (27.26645, 33.73355)$.

b) La hipòtesi nul·la, H_0 , és pendent = 0, com el zero no està dins de l'interval, acceptem que el pendent de la recta és significatiu (no és zero).

EXAMEN 3

1.- La següent taula mostra el nombre de llibres que han comprat durant l'últim mes 50 estudiants de la UOC

NLlibres	Nestudiants
0	20
2	5
3	10
4	10
5	5

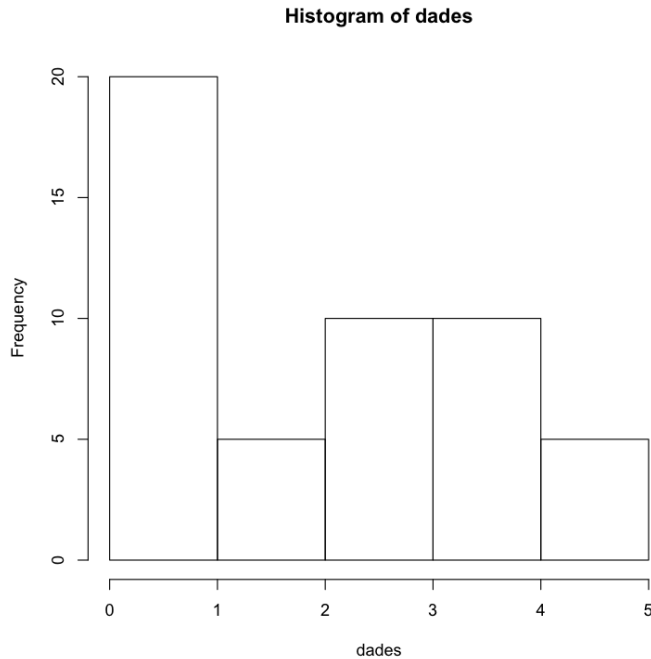
Es demana:

- Calculeu la mitjana i la mediana del nombre de llibres comprats.
- Calculeu la variància poblacional del nombre de llibres comprats.
- Representeu gràficament l'histograma de la distribució dels llibres comprats per aquests estudiants i comenteu el resultat.

Criteris de puntuació i valoració (sobre 10): a) (5 punts) Càlcul de la mitjana: 3 punts. Càlcul de la mediana 2 punts. b) (2 punts). c) (3 punts) 2 pel gràfic i 1 pel comentari

Solució:

- Mitjana=2.1. Mediana=2.5
- Variància=3.49
- Podria ser



La immensa majoria no compra llibres; entre els que compren la distribució és simètrica entre 2 i 5 llibres.

2.- En un grup de 200 projectes, 130 han estat desenvolupats en C+ i 100 han estat ben valorats pels clients. També sabem que els successos “desenvolupar en C+” i “estar ben valorats pels clients” són independents.

Es demana:

- Si escollim un projecte a l'atzar, quina és la probabilitat que **no** hagi estat desenvolupat en C+
- Si escollim un projecte a l'atzar, quina és la probabilitat que hagi estat desenvolupat en C+ i que estigui ben valorat pel client.
- Si escollim un projecte a l'atzar, quina és la probabilitat que no hagi estat desenvolupat en C+ i que no estigui ben valorat pel client.

criteris de puntuació i valoració (sobre 10): a) (1 punt) Un punt per identificar la probabilitat que es demana i 1 punt pel càlcul. b) (4 punts) Identificar la probabilitat demanada 2 punts; càlcul 2 punts. c) (4 punts) Identificar el que es demana 2 punt; càlculs 2 punts.

Solució:

- 0.35
- 0.325
- 35/200

3.- El temps que tarda un servidor de correu a processar si un missatge és SPAM o no segueix una distribució exponencial de mitjana 0.1 segons.

- Calculeu el valor del paràmetre λ de la distribució.

b) Calculeu la probabilitat que el servidor tardi més de mig segon en processar si un missatge és SPAM.

c) Calculeu la variància del temps que tarda el servidor a processar si un missatge és SPAM.

Criteris de puntuació i valoració (sobre 10): a) (3 punts) Càlcul del valor del paràmetre, 3 punts. b) (4 punts). Escriure bé la probabilitat demanada, 2 punts. Càlcul de la probabilitat, 2 punts. c) (3 punts). Càlcul de la variància, 3 punts.

Solució:

a) Sigui T la variable que dona el temps que tarda el servidor en processar si un missatge és SPAM. Com $T = \exp(\lambda)$ i $E(T) = \frac{1}{\lambda} = 0.1$, tenim que el valor del paràmetre λ serà: $\lambda = \frac{1}{0.1} = 10$.

b) Ens demanen: $p(X \geq 0.5) = 1 - F_X(0.5) = e^{-10 \cdot 0.5} = 0.006738$.

c) La variància de la distribució serà: $Var(X) = \frac{1}{\lambda^2} = 0.01$.

4.- El temps d'espera en segons, des que arranca un ordinador fins poder començar a treballar amb ell, és una variable aleatòria uniforme entre els 30 segons i els 3 minuts. S'agafa una mostra de 100 temps d'espera. Sigui X la variable aleatòria que ens dona la mitjana dels temps d'espera de la mostra.

a) Digueu quina és la distribució aproximada de X .

b) Calculeu la probabilitat aproximada que X agafi valors entre 100 i 150 segons.

Valors de les taules de la $N(0,1)$ (utilitzeu els que considereu necessaris, si necessiteu un valor que no és exacte, utilitzar el valor més proper)

x	p(X >= x)
0	0.5
1.15	0.1251
1.25	0.1056
1.35	0.0885
1.45	0.0735
1.55	0.0606
2	0.0228
3	0.0013
4	0
5	0

Criteris de puntuació i valoració (sobre 10): a) (5 punts). Donar la distribució, 2 punts. Donar els paràmetres del que depèn, 3 punts. b) (5 punts). Escriure bé la probabilitat demanada, 2.5 punts, càlcul de la probabilitat, 2.5 punts.

Solució:

- a) Aplicant el Teorema del Límit Central, la distribució de la variable X serà normal de mitjana $\mu = \frac{30+180}{2} = 105$, i variància $\sigma_S^2 = \frac{(180-30)^2}{12 \cdot 100} = 18.75$.
- b) Sigui $X \approx N(\mu = 105, \sigma_S^2 = 18.75)$. La probabilitat demanada serà:
 $p(100 \leq X \leq 150) \approx p\left(\frac{100-105}{\sqrt{18.75}} \leq Z \leq \frac{150-105}{\sqrt{18.75}}\right) = p(-1.15 \leq Z \leq 10.39) = p(Z \leq 10.39) - p(Z \leq -1.15) = 1 - 0.1251 = 0.8749$, on Z representa la normal $N(0,1)$.

5.- Durant 100 dies, hem estudiat quants dies dos servidors han estat atacats. El primer servidor ha estat atacat en 54 dies i el segon, en 48 dies. Contrasteu a un nivell de significació del 0.05 si la proporció de dies en què han sofert atacs ha estat la mateixa pels dos servidors. Heu d'especificar el tipus de contrast realitzat, les hipòtesis nul·la i alternativa, l'estadístic de contrast i els valors crítics.

$Z \sim N(0,1)$ $p(Z \geq z)$	z
0.01	2.326
0.025	1.96
0.05	1.645
0.005	2.575
0.1	1.28

Criteris de puntuació i valoració (sobre 10): Hipòtesi del contrast, 2 punts. Estadístic, 2 punts. Valors crítics, 2 punts. Conclusió, 4 punts.

Solució

Es tracta d'un contrast de proporcions entre dues mostres independents:

Plantegem el contrast

$$\left. \begin{array}{l} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{array} \right\} \text{ on } p_1 \text{ seria la proporció del nombre de dies que ha estat atacat el primer servidor i } p_2, \text{ la proporció del nombre de dies que ha estat atacat el segon servidor.}$$

Nivell de significació: $\alpha = 0.05$

Calculem $\hat{p}_1 = \frac{54}{100} = 0.54$; $\hat{p}_2 = \frac{48}{100} = 0.48$;

$$\hat{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} = \frac{54 + 48}{200} = 0.51;$$

$$s_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = 0.071.$$

Estadístic de contrast: $z = \frac{\hat{p}_1 - \hat{p}_2}{s_{\hat{p}}} = 0.849$

Valor crític, mirant la taula de la normal: 1.96. Com que $1.96 > 0.849$, acceptem la hipòtesi nul·la.

Conclusió: no rebutgem la hipòtesi nul·la pel que podem dir que la proporció de dies que han sofert atacs els dos servidors és la mateixa.

6.- Ens donen l'output de R següent on “preu” representa el preu, en euros, d'una mostra d'impressores, “pag” representa el nombre de pàgines per minut que la impressora és capaç d'imprimir.

Call:

```
lm(formula = Preu ~ pag, data = Impressores)
```

Residuals:

Min	1Q	Median	3Q	Max
-32.000	-18.625	-2.375	19.125	37.000

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	328.875	26.181	12.562	1.90e-07 ***
pag	20.188	2.903	6.954	3.93e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 25.08 on 10 degrees of freedom
Multiple R-squared: 0.8286, Adjusted R-squared: 0.8115
F-statistic: 48.36 on 1 and 10 DF, p-value: 3.928e-05

Volem ajustar la variable Y, preu de la impressora, a la variable X, nombre de pàgines per minut.

- Trobeu la recta de regressió del preu de la impressora en funció del nombre de pàgines per minut i interpreteu els resultats.
- Feu el contrast d'hipòtesi sobre el pendent de la recta, amb un nivell de confiança del 99.99%.
- Calculeu el coeficient de determinació i el coeficient de regressió lineal. És un bon model?

Criteris de puntuació i valoració (sobre 10): Tots els apartats valen igual.

Solució:

- La recta de regressió del preu (Y) en funció del nombre de pàgines per minut (X) és:

$$Y = 329 + 20.2 X.$$

Interpretació dels coeficients:

- Pendent de la recta (20.2): és l'augment de preu per cada unitat de X que augmenta —és a dir, per cada pàgina de més per minut.
- Terme independent (329): és més difícil d'interpretar. Representaria el preu de la impressora si aquesta no imprimís cap pàgina.

b) El contrast d'hipòtesi sobre el pendent de la recta es veu en la línia següent de resultats:

	Estimate	Std. Error	t value	Pr(> t)
pag	20.188	2.903	6.954	3.93e-05 ***

En tenir un p-valor (p-value: 3.928e-05) més petit que 0.0001, rebutgem la hipòtesi nul·la H_0 (pendent = 0) i acceptem que el pendent de la recta és significatiu (no és zero).

c) El coeficient de determinació val: $R\text{-Sq} = 0.829$ i el coeficient de correlació lineal serà $r = \sqrt{0.829} = 0.91$.

És un bon model en estar r molt prop d'1.