

NOM: _____ COGNOM: _____
(Contesteu cada pregunta en el seu lloc. Expliqueu i justifiqueu els càlculs)

Problema de B1

Suposem que en una ciutat concreta, l'aeroport A controla el 50% del trànsit aeri de passatgers, i que els aeroports B i C controlen el 30% i 20% restant respectivament. Les taxes de detecció d'armes (probabilitat que un passatger porti una arma) en aquests aeroports són de 0.10, 0.20 i 0.20 respectivament.

a) Dibuixa l'arbre de probabilitats que es dedueix de l'enunciat i el conjunt de resultats (1 punt)

b) Quina és la probabilitat de que un passatger, en un dels aeroports qualsevol, no porti una arma? (1 punt)

c) S'informa que s'ha detectat un passatger que porta una arma. De quin aeroport és més probable que sigui? (2 punts)

d) Si un passatger ve de l'aeroport A o B, quina és la probabilitat de que no porti cap arma? (1 punt)

e) En aquesta ciutat es proposa que independentment de l'aeroport i de si s'ha detectat arma o no, la meitat dels passatgers s'han de treure les sabates. Quina és la probabilitat que un passatger que no porta cap arma i no s'ha tret les sabates estigui a l'aeroport A? I que estigui a l'aeroport A si sabem que no porta cap arma però no sabem res de les sabates? (2 punts)

f) Suposem una altra ciutat que també té tres aeroports amb el mateix repartiment del trànsit aeri i amb les mateixes taxes de detecció d'armes en els dos primers aeroports. Quines són les taxes en el tercer aeroport si sabem que en aquesta ciutat la probabilitat de que un passatger no porti arma en un qualsevol dels aeroports és del 84% ? (2 punts)

e) En aquesta segona ciutat, quina és la probabilitat d'estar a l'aeroport B o C i portar una arma? (1 punt)

NOM: _____ COGNOM: _____
(Contesteu cada pregunta en el seu lloc. Expliqueu i justifiqueu els càlculs)

Problema de B2

Algunes assignatures de la FIB, donada la situació del COVID 19, han volgut quantificar la probabilitat de fer exercicis setmanalment en les plataformes online que posen a disposició dels estudiants. En una assignatura X la distribució de probabilitats de fer entre 0 i 4 exercicis és la següent:

$X = x$	0	1	2	3	4
$P(X = x)$	0.1	0.2	0.25	0.25	0.2

1) Quina és la probabilitat de fer algun exercici? Argumenta si es compleix la condició que espera el professorat de que quatre de cada cinc estudiants facin algun exercici (1 punt)

2) Suposem que sabem que un alumne ha fet algun exercici. Quina és la probabilitat que n'hagi fet més de dos? (1 punt)

3) Quin és el valor esperat i la desviació típica del nombre d'exercicis realitzats setmanalment? Comenteu i interpreteu els resultats (1.5 punts)

4) Per a una assignatura amb 100 matriculats i amb l'anterior distribució de probabilitats per a cada estudiant (i suposant que el nombre d'exercicis que fa un estudiant no afecta als demés), calculeu, i justifiqueu formalment, quin és el valor esperat i la desviació del nombre d'exercicis que recollirà l'assignatura setmanalment (1.5 punts)

5) Dues assignatures, X i Y, han estudiat la distribució de probabilitat conjunta dels seus estudiants comuns per fer exercicis d'una i altra setmanalment, obtenint:

	X=0	X=1	X>1
Y=0	0.05	0	0.05
Y=1	0	0.1	0.1
Y>1	0.05	0.1	0.55

Quina és la probabilitat de fer 2 o més exercicis entre les dues assignatures? (1 punt)

Indiqueu i justifiqueu com és la relació entre fer exercicis d'una assignatura i l'altra (1 punt)

6) Les dues assignatures anteriors també tenen estudiada la distribució del temps en minuts que dediquen els seus estudiants en resoldre els exercicis. Per a l'assignatura X els minuts dedicats segueixen la distribució $\exp(-k)$, i per a l'assignatura Y la distribució $1/9 k^2$ ($0 \leq k \leq 3$)

Indiqueu les funcions de probabilitat i de distribució de probabilitat en cada cas i calculeu quina és la probabilitat d'estar menys de 2 minuts per fer un exercici en cada cas? (1.5 punts)

En quin dels dos casos el valor esperat de minuts per fer un exercici és més gran? (1.5 punts)

NOM: _____ COGNOM: _____
(Contesteu cada pregunta en el seu lloc. Expliqueu i justifiqueu els càlculs)

Problema de B3

En una determinada marca de roba estan estudiant la quantitat de peces de roba de les comandes que reben online a través del nou web. En concret estan estudiant la quantitat de comandes que només compren una peça de roba i han trobat que, de mitjana, aquestes representen un 25% de les comandes.

Es defineix la variable aleatòria X : "nombre de comandes amb una sola peça de roba comprada". Si en una hora concreta han rebut 20 comandes,

1) Quina és la probabilitat que cap de les 20 comandes hagi estat d'una sola peça de roba? (explícita la distribució que segueix la variable X) (1 punt)

2) Quina és la probabilitat que s'hagin rebut més de quatre comandes amb una sola peça de roba? (1 punt)

3) Quin és el nombre esperat de comandes fins que arribi una comanda amb una sola peça de roba? (explícita la distribució que segueix la variable que definiu) (1 punt)

A partir de l'estudi realitzat han recollit que la despesa de les comandes que reben segueixen una distribució normal amb $\mu=50\text{€}$ i $\sigma=15\text{€}$. Per incentivar les vendes online volen procurar no cobrar despeses d'enviament, però només els hi surt a compte per les comandes de més de 20€.

4) Quina és la probabilitat que una comanda sigui per un import menor de 20€? (1 punt)

De 8:00 del matí a 8:00 de la nit dels dies laborables, l'empresa compta amb un servei personalitzat d'atenció a les comandes. El nombre de trucades que reben al servei de comandes segueix una llei de Poisson amb una mitjana 3 trucades per hora.

5) Quina és la probabilitat que en una hora es rebin dues trucades? (1 punt)

6) Quina és la desviació poblacional del nombre de trucades que es reben al servei? (0,5 punts)

Considerant que la quantitat de trucades rebudes al Servei personalitzat d'atenció a les comandes segueixen una distribució de Poisson amb una mitjana de 3 trucades per hora, es defineix la variable aleatòria T : "temps transcorregut entre dues trucades consecutives al servei".

6) Quina és la probabilitat que el temps transcorregut entre dues trucades sigui més petit que 30 minuts? (Explicita la distribució de la variable aleatòria T) (1 punts)

7) Quin és el temps esperat entre dues trucades consecutives al servei? (0'5 punts)

Es vol estudiar el conjunt de les trucades que es reben al Servei d'atenció personalitzat durant les 12 hores en què està operatiu.

8) Justifica quines condicions ha de complir la situació plantejada per tal que es pugui utilitzar el Teorema Central del Límit per a realitzar l'estudi desitjat. (1 punt)

9) Determina el nombre màxim de queixes que es reben en un dia laborable de 8 del matí a 8 del vespre al Servei d'atenció personalitzat amb un error del 5%. (2 punts)

Problema de B4

1. Es vol dissenyar una nova app de pagament per mòbil que sigui un comparador de preus de productes informàtics a diferents portals (Amazon, Ali Express, eBay). Per saber quina acceptació tindria l'app al mercat s'ha fet una enquesta a 200 usuaris potencials. En aquesta enquesta es pregunta als clients quants diners estarien disposats a pagar per un producte informàtic concret (mateix per tots els enquestats). Es crea una variable de rendiment de l'app (R) que es la diferència entre el preu que pagaria el client (p_C) i el preu que oferiria l'app ($p_A = 500\text{€}$). Els creadors volen contrastar la hipòtesi de si el preu que pagarien els clients és igual al preu que oferiria l'app.

<code>mean(R, na.rm=TRUE)</code>	<code>sd(R, na.rm=TRUE)</code>
5.71	43.57

- a) Planteja el contrast d'hipòtesis i digues explícitament si el contrast és unilateral o bilateral. (1 punt)
- b) Digues i) quin estadístic empraries; ii) fes el seu càlcul, iii) la distribució que segueix sota la hipòtesi nul·la i les premisses necessàries; i iv) digues quin o quins són el punt/s crític/s amb un nivell de significació $\alpha=5\%$. (1 punt)
- c) Conclou sobre la prova d'hipòtesi. (1 punt)
- d) Els creadors de l'aplicació, no saben massa d'estadística i no tenen clar, en el cas que la variable R no complís les premisses amb quina de les següents variables podrien dur a terme el contrast d'hipòtesi:

$$R' = |p_C - p_A| \quad R'' = \log(p_C) - \log(p_A) \quad R''' = \log(p_C - p_A)$$

Quina/es de les 3 variables et semblarien adients per dur a terme el contrast d'hipòtesis? Raona la resposta (1 punt)

2. Un altre pregunta important de l'enquesta és si estarien interessats en adquirir aquesta app en el futur. L'enquesta s'ha realitzat emprant quotes per gènere de tal manera que s'han obtingut els resultats per 100 dones i 100 homes. 32 dones i 18 homes han manifestat el seu interès en l'aplicació. Es decidirà treure l'app al mercat si el percentatge de gent interessada (dones o homes) és superior al 20% amb un nivell de significació $\alpha=5\%$.

a) Planteja el contrast de hipòtesis i digues explícitament si és unilateral o bilateral (1 punt)

b) Digues quin estadístic emprar, les premisses i la distribució que seguirà sota la hipòtesi nul·la (1 punt)

c) Calcula l'estadístic i el punt crític (1 punt)

d) Conclou sobre la prova d'hipòtesi. Trauran l'app al mercat? (1 punt)

3. La pregunta de l'apartat 2 sobre si els usuaris estarien interessats en comprar l'app és clau i els creadors voldrien estimar la proporció poblacional de gent interessada en l'app amb un error de ± 0.05 i una confiança del 90%. Digues quina hauria de ser la grandària mostral de l'enquesta per assolir aquest objectiu (2 punts)

A tots els exercicis: justifiqueu formalment les passes de les respostes

Problema B5

Un grup d'estudiants de PE es planteja estudiar les diferències de valoracions de jocs de mòbil gratuïts i de pagament. Amb aquest objectiu recullen les valoracions (sobre una escala de 0 a 100) d'una pàgina especialitzada en Internet de 31 jocs gratuïts i de pagament, respectivament. La taula a continuació mostra els indicadors numèrics més rellevants dels valors en ambdues mostres.

	<i>n</i>	Mitjana	Mediana	Desv. est.	Mínim	Màxim
<i>Jocs de pagament</i>	31	78,8	79,5	7,1	66,6	95,5
<i>Jocs gratuïts</i>	31	73,7	72,6	8,7	55,8	93,2

Definim com a variable V la valoració dels jocs i utilitzem com a subíndexs p i g per jocs de pagament i gratuïts, respectivament.

(a) (0.5 punts)

Es tracta de dues mostres independents o aparellades? Raoneu la resposta.

(b) (0.5 punts)

Només en base als valors descriptius, creieu que la variable V segueix una distribució normal en les dues mostres?

(c) (2.5 punts)

Abans d'estudiar si les valoracions mitjanes són diferents d'un tipus a un altre dels jocs de mòbil, comprovem si es compleix la igualtat de variàncies o no.

- Plantegeu la hipòtesi nul·la i l'alternativa:
- Calculeu el valor de l'estadístic de contrast:
- Resolució del contrast utilitzant un nivell de significació $\alpha = 0.1$:
- Interpretació pràctica:

(d) (3 punts)

Contrasteu ara la hipòtesi que les valoracions mitjanes de jocs de mòbil de pagament i gratuïts són iguals o no.

- Plategeu la hipòtesi nul·la i l'alternativa:
- Calculeu el valor de l'estadístic de contrast:
- Resolució del contrast utilitzant un nivell de significació $\alpha = 0.05$:
- Interpretació pràctica:

(e) (2.5 punts)

Hem fet el càlcul del interval de confiança del 95% de la diferència de les valoracions mitjanes amb R i hem obtingut el següent resultat:

$$CI(\mu_p - \mu_g; 0.95) = [1.12, 9.17].$$

Utilitzeu aquest resultat per calcular l'interval de confiança del 90% de la diferència de les valoracions mitjanes.

(f) (1 punt)

Expliqueu com es pot utilitzar $CI(\mu_p - \mu_g; 0.95)$ per contrastar la hipòtesi

$$H_0: \mu_p = \mu_g \quad \text{vs.} \quad H_1: \mu_p \neq \mu_g.$$

Problema de B6

Volem comparar una possible relació lineal entre les notes d'una assignatura A1 amb unes altres dues A2 i A3.

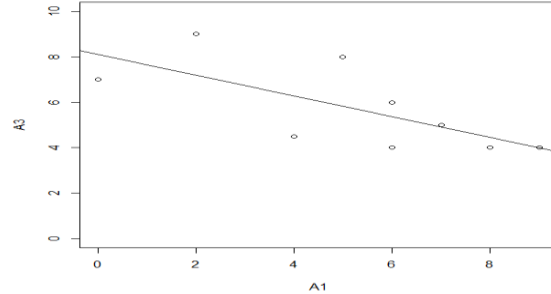
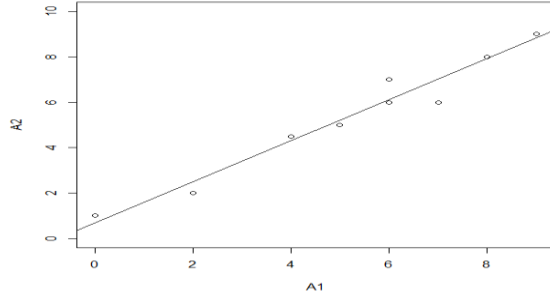
Disposem de les següents dades i resultats en 10 dels seus estudiants:

Notes A1: 0 2 4 5 6 6 7 8 9 9 $\sum A1_i = 56$ $\sum A1_i^2 = 392$

Notes A2: 1 2 4.5 5 6 7 6 8 9 9 $\sum A2_i = 57.5$ $\sum A2_i^2 = 397.25$ $\sum A1_i A2_i = 393$

Notes A3: 7 9 4.5 8 4 6 5 4 4 4 $\sum A3_i = 55.5$ $\sum A3_i^2 = 339.25$ $\sum A1_i A3_i = 275$

mean(A1)=5.6 sd(A1)=2.95 mean(A2)=5.75 sd(A2)=2.72 mean(A3)=5.55 sd(A3)=1.86 cov(A1,A2)=7.89 cov(A1,A3)=-3.98



```
lm(formula = A2 ~ A1)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	_____	0.3813	1.78	0.113
A1	_____	0.0609	14.87	0.0000004

Residual standard error: 0.5393 on 8 degrees of freedom

Multiple R-squared: 0.9651, Adjusted R-squared: 0.9607

F-statistic: 221.1 on 1 and 8 DF, p-value: 4.124e-07

1) Per comparar les rectes ajustades en els dos models indiqueu:

a) els valors dels coeficients ordenada a l'origen i del pendent de la recta ajustada en cada model (2 punts)

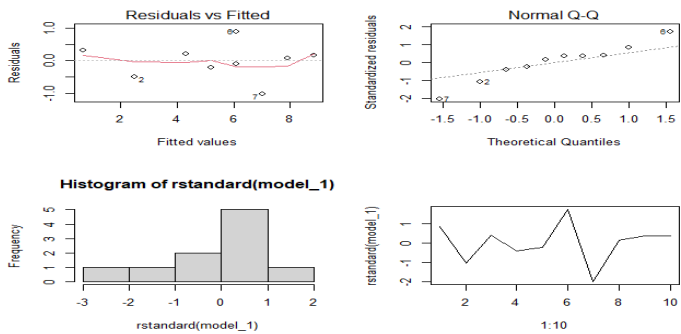
b) les equacions de les rectes estimades amb la seva interpretació (2 punts)

2) Pel cas del model entre A1 i A2, calculeu un interval de confiança amb risc 5% per al pendent de la recta, i interpreteu-lo (2 punts)

3) Indiqueu i compareu pels dos models els valors del coeficient de correlació i el de determinació (2 punts)

4) Enuncieu les premisses de la regressió lineal i compareu-les en els dos models (indiqueu els gràfics d'on es dedueixen) (2 punts)

Model lineal entre A1 i A2



Model lineal entre A1 i A3

