

NOM: _____

(Contesteu cada pregunta en el seu lloc. Expliqueu i justifiqueu els càlculs.)

Problema 1 (B1-B2)

La taula adjunta mostra les proporcions (poblacionals) corresponents als estudiants que obtenen simultàniament una qualificació numèrica X a certa assignatura i una qualificació literal a una determinada competència transversal CT:

	3	5	7	9	10
A	0.006	0.012	0.067	0.04	0.005
B	0.009	0.021	0.096	0.033	0.018
C	0.072	0.144	0.169	0.088	0.044
D	0.055	0.075	0.036	0.01	0

- (2 pt) Trobeu la funció de probabilitats de la variable aleatòria X , i calculeu el valor esperat i la desviació tipus. Interpreteu el resultat.
- (2 pt) Trobeu la distribució de la variable X condicionada a haver obtingut una D a la CT. Quina és la qualificació esperada entre aquests?
- (1 pt) Quina és la proporció de B o C, en general? I entre els que han tret un 9? Expresses formalment aquestes probabilitats abans de contestar.
- (1 pt) Els estudiants poden optar a Matrícula d'Honor (MH) si obtenen un 10, o si obtenen un 9 i una A. Expliqueu la diferència existent entre trobar la probabilitat que un alumne amb MH hagi tret un 9, o que un alumne amb un 9 hagi tret MH. Calculeu-les.
- (1 pt) Si un estudiant escollit a l'atzar resulta que ha obtingut una qualificació alta a la CT (alguna entre A o B, però no sabem quina), quina és la probabilitat que la seva nota sigui superior a 5?

6. (2 pt) Són independents les qualificacions X i CT? Trobeu i representeu gràficament les funcions $P_{X|CT=B}(x_i)$ i $P_{X|CT=C}(x_i)$. Interpreteu el resultat, vinculant la resposta a la pregunta inicial.

7. (1 pt) Podem calcular la correlació entre X i CT?

8. (1 pt / •) Pensem ara en la nota X com una variable contínua amb la següent funció de distribució:

$$F_X(x) = \frac{ax + b}{x + 1}, 0 \leq x \leq 10, 0 \text{ si } x < 0, 1 \text{ si } x > 10$$

- És cert que aquesta funció podria ser una funció de distribució? Què hauria de complir?
- Determineu els valors de les constants a i b per tal que sigui funció de distribució;
- Calculeu la probabilitat que la nota estigui entre 5 i 7, i la probabilitat que sigui major que 9;
- Trobeu i dibuixeu la funció de densitat de X.
- Com es calcularia l'esperança de X? (no cal que ho resoleu). Marqueu a la gràfica el lloc aproximat on es situaria el valor esperat.

NOM: _____

(Contesteu cada pregunta en el seu lloc. Explíciteu i justifiqueu els càlculs.)

Problema 2 (B3-B4).

Suposem que fem una cerca amb Google. La probabilitat de *clicar* en un enllaç patrocinat que apareix a la primera posició de la primera pàgina al realitzar una cerca és 0.3. Una empresa de reparació d'equipaments informàtics té una plana web i s'està plantejant contractar aquest enllaç patrocinat per a que aparegui al fer una cerca amb els termes "reparació" i "ordinadors".

1. A l'àmbit d'actuació de l'empresa es realitzen 20 cerques en una hora amb la combinació d'aquests 2 termes.
 - a. Quina és la distribució de la variable aleatòria "*Nombre de clicks fets en aquesta hora a l'enllaç patrocinat*"? Quina és la seva esperança i desviació estàndard? **(1.25 punts)**
 - b. Quina és la probabilitat que el nombre de *clicks* en aquesta hora sigui igual o superior a 10? **(1.25 punts)**
2. En un dia concret, durant les 8 hores laborables que treballen els tècnics de reparació s'han fet 15 cerques per hora d'aquests termes. La probabilitat que un visitant de la plana web contracti un servei de reparació és 0.15. Considereu la variable aleatòria "*Nombre de serveis contractats en 8 hores*"
 - a. Quina és la distribució d'aquesta variable aleatòria? Quina és la seva esperança? **(1.25 punts)**
 - b. El nombre màxim de reparacions diàries que es poden fer per dia són 8. Quina és la probabilitat de no poder realitzar alguna reparació per sobrepassar aquest número de demandes en aquest dia? **(1.25 punts)**
 - c. Quin és el llinard de temps transcorregut entre dues demandes que es sobrepassarà amb una probabilitat de 0.8? **(1.25 punts)**
3. Al entrar en una plana web venint de *Google*, hi ha dos tipus d'usuari, els que surten rebotats de seguida perquè veuen que no els interessa la pàgina web (Usuaris A) i els que naveguen una estona buscant més informació (Usuaris B). El temps de la visita dels usuaris A a la pàgina web es distribueix amb una exponencial de $\lambda=2/\text{minuts}$, mentre que la distribució dels temps dels usuaris B es distribueix com una Normal amb $\mu=6$ minuts i $\sigma=2$ minuts. També sabem que hi ha un 50% d'usuaris de cada tipus.
 - a. Quina serà la durada esperada d'una visita qualsevol? **(1.25 punts)**
 - b. Quina serà la durada que sobrepassaran els usuaris B amb una probabilitat de 0.99? **(1.25 punts)**

- c. Dóna la probabilitat que un usuari que sobrepassa la durada de 1.5 minuts sigui un usuari de tipus B? **(1.25 punts)**

4. Finalment, l'empresa opta per no contractar un enllaç patrocinat de *Google* i decideix contractar-lo dins del cercador *Bing*. En un mes, la descriptiva de la durada en minuts de les visites provinents de Bing és la següent:

```
> length(durada)
[1] 120
> summary(durada)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.0020  0.3655  2.8520  3.4370  6.4540  9.4800
> sd(durada)
[1] 3.146428
```

Fes la prova d'hipòtesi de si la durada mitjana de les visites provinents de *Bing* és de 2.5 minuts o no ho és amb un $\alpha=0.1$.

- a. Dóna l'expressió de l'estadístic, la seva distribució sota H_0 i les premisses adients **(2.5 punts)**

- b. Dóna el valor de l'estadístic i del punt crític. **(2.5 punts)**

- c. Dóna el p-valor i la conclusió de la prova d'hipòtesi **(2.5 punts)**

- d. Dóna el IC90% per la durada mitjana de les visites provinents de *Bing* i fes la seva interpretació **(2.5 punts)**

NOM: _____

(Contesteu cada pregunta en el seu lloc. Expliqueu i justifiqueu els càlculs.)

Problema 3 (B5-B6).

Es pren una mostra de 15 estudiants de l'assignatura de PE a la FIB i s'analitza si el nombre d'hores (**H**) que han estudiat per a un examen afecta a les puntuacions que han obtingut (**P1**). A més, tenim les puntuacions d'un altre grup (**P2**).

Les dades són: **Hores (H)**: 2, 3, 3, 4, 4, 5, 5, 6, 6, 6, 7, 7, 7, 8, 8

Puntuació (P1): 5, 5, 6, 5, 7, 7, 8, 6, 9, 8, 7, 9, 10, 8, 9

Puntuació (P2): 4.4, 4.5, 5.3, 3.7, 7.3, 6.7, 6.8, 5.8, 8.2, 8, 7, 7.8, 8, 8.7, 8

i els estadístics obtinguts a partir d'aquestes tres variables són:

Mitjana de H: $\text{mean}(H)=5.4$

Desviació tipus de H: $s_H=1.88$

Mitjana de P1: $\text{mean}(P1)=7.27$

Desviació tipus de P1: $s_{P1}=1.62$

Mitjana de P2: $\text{mean}(P2)=6.68$

Desviació tipus de P2: $s_{P2}=1.58$

Correlació entre H i P1: $\text{cor}(H,P1)=0.78$

1.- (2 punts) A partir de les desviacions tipus d'aquestes dades mostrals per a P1 i P2, creieu que les desviacions tipus poblacionals d'aquests grups són similars? Plantegeu i resolueu la prova d'hipòtesi que ens permeti rebutjar o no aquesta afirmació.

2.- A partir de les puntuacions P1 i P2 d'aquestes dades mostrals es creu que les mitjanes d'aquestes mostres podrien ser similars. Plantegeu i resolueu la prova d'hipòtesi que permeti rebutjar o no aquesta hipòtesi. Les passes a seguir són:

a) (0.5 punt) Definiu la hipòtesi nul·la i l'alternativa per al test relacionat amb les mitjanes de les puntuacions P_1 i P_2

b) (1 punt) Calculeu el valor d'aquesta prova d'hipòtesi al nivell de significació $\alpha=0.05$

c) (0.5 punt) Quin és el valor, en termes de valors absoluts, més petit de l'estadístic t pel qual la hipòtesi nul·la pot ser rebutjada?

d) (0.75 punt) Indiqueu si les següents afirmacions són vertaderes o falses i justifiqueu la vostra resposta:

i. Si es rebutja la hipòtesi nul·la H_0 al nivell de 0.05, també podem rebutjar al nivell 0.1

ii. El risc d'error de tipus I és la probabilitat de rebutjar la hipòtesi nul·la quan H_1 és certa

iii. Si el p-valor és igual a 0.15, podem rebutjar la hipòtesi nul·la al nivell del 10%

3.- (0.5 punt) Expliqueu com s'haurien de recollir les dades per a que fossin un conjunt de dades aparellades i com s'haurien de recollir si es vol treballar amb dos conjunts de dades independents

4.-a) (2 punts) Plantegeu un model lineal i estimeu la recta de regressió que permet estimar la puntuació de **P1** a partir de les hores d'estudi **H**

b) (0.5 punt) Un alumne determinat ha estudiat 6.5 hores i ha tret un 7.5: calculeu el seu valor residual

c) (0.5 punt) Calculeu un estimador no esbiaixat per a la variància dels errors d'aquest model

d) (0.75 punt) Calculeu i interpreteu el coeficient de determinació (R^2)

e) (1 punt) Obtingueu un interval de confiança, a un nivell de confiança del 95%, pel valor mitjà de la puntuació, sabent que les hores d'estudi han estat 5.5