

**NOM:** \_\_\_\_\_

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

**PROBLEMA 1.** Considerem el conjunt de tots els paquets de 3 bits que es poden enviar per una línia de comunicació (000,001,010,011,100,101,110,111) com a resultats d'una experiència aleatòria, i suposem que la probabilitat del 0 és el triple que la del 1.

Definim, també, dues variables aleatòries  $X$  i  $Y$ . La variable  $X$  és la suma dels 3 bits i la variable  $Y$  és el número d'alternances en la seqüència de bits de cadascun dels resultats. Per tant,  $X \in \{0,1,2,3\}$  i  $Y \in \{0,1,2\}$ .

1.- (1 punt) Indiqueu el conjunt de resultats de l'experiència aleatòria i les seves probabilitats associades justificant-les (en representació en arbre si voleu) i associeu a cada resultat un valor per les variables  $X$  i  $Y$

2.- (1 punt) Calculeu la probabilitat que la suma dels 3 bits sigui 2. I la probabilitat que la suma dels tres bits sigui 2 si sabem que el primer bit ha estat 1

3.- (2 punts) Indiqueu la taula de probabilitats conjuntes de les variables  $X$  i  $Y$ , les probabilitats marginals corresponents i la seves esperances

4.- (1 punt) Hi ha independència entre les variables  $X$  i  $Y$ ? Justifiqueu-ho i expliqueu què implica

S'ha estudiat que en l'enviament de paquets de 3 bits, es pot considerar una nova variable  $Q$  que indica en quants d'aquests bits s'ha presentat algun problema (en 0,1,2 o 3 casos), i les probabilitats de la qual són 0.4 0.3 0.2 i 0.1 respectivament

5.- (1 punt) Calculeu l'esperança i la variància de la variable  $Q$

6.- (1 punt) Calculeu la probabilitat de tenir 2 o menys problemes. I la probabilitat de tenir 2 o menys problemes si sabem que almenys 1 n'ha tingut

**Ara, enlloc de considerar els anteriors paquets de tres bits, considerem enviaments de 8 bits amb probabilitat d'enviar 0 sent el triple que la d'enviar 1.**

7.- (1 punt) Definiu la variable nombre de 0's en els 8 enviaments i indiqueu l'esperança i variància d'aquesta variable

**Finalment, estudiem el comportament més centrat en l'enviament per unitat de temps que en el nombre d'enviaments, i suposem que el nombre mitjà de bytes enviats per *unitat de temps* és de 5.**

8.- (1 punt) definiu la variable nombre de bytes enviats per *unitat de temps*, indiqueu l'esperança i variància d'aquesta variable i calculeu la probabilitat de que se n'enviïn exactament 5

9.- (1 punt) definiu la variable temps entre enviaments, indiqueu-ne l'esperança i la variància, i calculeu la probabilitat de que el temps entre enviaments sigui superior a 1 *unitat de temps*

NOM: \_\_\_\_\_

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

## PROBLEMA 2. Memòria dels ordinadors

Un servidor de la marca **Fuig Primer R100** te 12 xips de memòria de 8Gb cadascun.

- a) **(1PUNT)**. Si la probabilitat de funcionament de cada xip és 0.92, quina és la probabilitat que funcionin 9 xips?
- b) **(1PUNT)**. A continuació ens diuen que han millorat la qualitat d'aquests xips i ara la probabilitat de que funcioni cada xip és del 95%. Trobeu la probabilitat de que funcionin 10 o més xips.
- c) **(1PUNT)**. Tenim una granja de servidors que te 20 ordinadors, cadascun d'aquests servidors amb les mateixes característiques que el servidor **Fuig Primer R100**, i la probabilitat que falli un xip és del 0.05. Dieu quin serà el valor mitjà i la desviació tipus del nombre de xips que fallin a la granja de servidors.

- d) **(1PUNT)**. S'ha recollit una mostra de  $n=100$  xips, observant que la proporció de xips defectuosos és del 6%. Doneu un interval de confiança al 90% per a la proporció de xips defectuosos.
- e) **(2PUNTS)**. Ara es suposa que en una mostra de 100 xips n'hi ha 88 que funcionen. Poseu a prova si es pot admetre que la proporció de defectuosos és del 6% o no, amb un risc del 5%.
- f) **(2PUNTS)**. Hem detectat que, amb un 95% de confiança, la proporció de xips defectuosos podria estar entre 0.3 i 0.5. Calculeu quina és la grandària de la mostra que permet assumir aquests valors.
- g) **(2PUNTS)**. Tenim una mostra de 31 xips que son similars en quant a tecnologia, no obstant els preus són diferents degut a que s'han comprat a diferents proveïdors. La desviació tipus mostral d'aquests preus és 20€. Doneu un interval de confiança al 95% per a la desviació tipus poblacional d'aquests preus.

NOM: \_\_\_\_\_

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

**PROBLEMA 3.** [14pts = 10 de nota] Els responsables de la xarxa social “tonti” volen provar dos tipus de perfil amb els seus usuaris. A data 1 de maig, a uns els assignen el perfil *normal*, i a uns altres el perfil *cool*. Escullen aleatòriament uns usuaris de cada tipus i analitzen les següents dades a dos moments diferents:

	perfil <i>normal</i>		perfil <i>cool</i>		
	(n=250)		(n=250)		
	$\bar{x}$	(s)	$\bar{x}$	(s)	
#F-mg	85.4	(17.3)	87.8	(19.0)	Nombre de <i>followers</i> a 1 de maig
#F-jn	91.2	(18.7)	97.5	(21.2)	Nombre de <i>followers</i> a 1 de juny
#com	127	(33.5)	146	(41.4)	Nombre de comentaris enviats al mes de maig

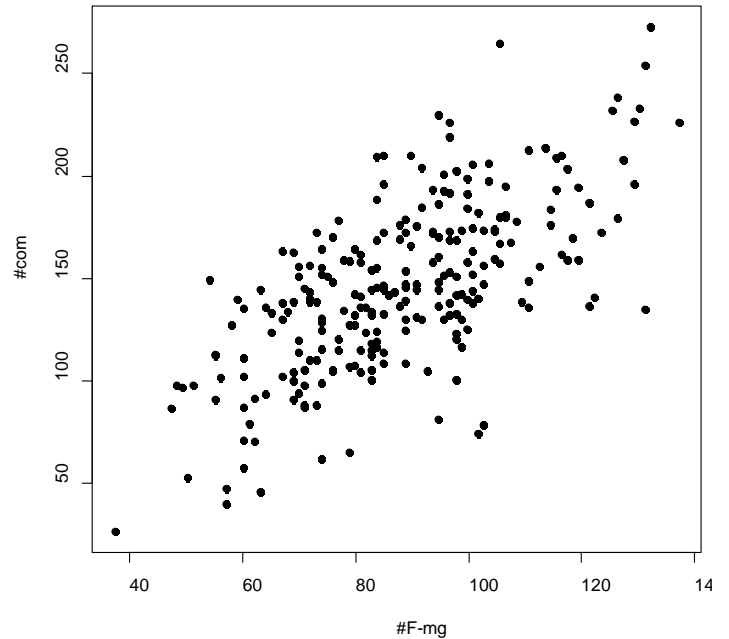
- A. Quina d'aquestes afirmacions té més sentit? No facis cap PH, només comenta. [1pt]
- L'assignació del perfil als usuaris a l'inici no s'ha fet bé: els *cools* tenen més *followers*
  - La diferència de *followers* a l'inici es pot explicar només per l'atzar
- B. Escull i comenta dos de les afirmacions. Subratlla els aspectes que no són certs, fent la correcció corresponent:
- Hauríem de saber la variació abans-després de *followers* per a cada usuari, i analitzar el canvi entre els dos grups, *normal* i *cool*, que són dues mostres independents. [qualssevol dues són 1 pt. cadascuna]
  - Hauríem de saber les diferències de *followers* entre usuari *normal* i *cool*, que són mostres aparellades, i veure si aquesta mostra prové d'una mitjana 0.
  - Si les mitjanes de #F-mg fossin les mateixes, comparariem les mitjanes de #F-jn que són dues mostres independents, però de variàncies poblacionals diferents. Però potser els *cools* tenen més *followers* perquè al principi ja tenien més.
  - Podem comparar els dos perfils utilitzant simplement la informació de la fila de juny, i assumint que poblacionalment els dos grups al maig eren idèntics.

- C. Ens fan notar que el nombre d'usuaris que, al juny, tenen més de 200 *followers* és més o menys un 5%. Què representa aquesta informació? Pot comprometre l'anàlisi estadístic que es vol fer? [1pt]
- D. Per a la diferència de *followers* entre juny i maig hem observat desviacions estàndards de 7.8 i 10.6 als grups *normal* i *cool* respectivament. Plantegeu i resoleu la PH corresponent per contestar la qüestió de si el perfil *cool* fa aconseguir més *followers* que el *normal*, treballant amb aquest variable *diferència*. [3pts]
- E. Obteniu una mesura de l'efecte del perfil, mitjançant un interval de confiança al 95%, i feu una interpretació global dels resultats. [1pt]

NOM: \_\_\_\_\_

**PROBLEMA 3 (CONT)**

F. La figura adjunta mostra la relació entre el nombre de comentaris i *followers* a l'inici en el grup *cool*. Tenim la dada de que la correlació mostrada a aquestes variables és 0.67214.



a. Estimeu quina és la recta de regressió. [1pt]

b. Interpreteu els coeficients que heu trobat. [1pt]

c. Un model lineal sembla un model apropiat? Repasseu breument les premisses del model. [1pt]

G. Què val la desviació residual estimada per al model? I el coeficient de determinació? Estan relacionats aquests dos indicadors? És a dir, si un puja l'altre ha de pujar o baixar? Si l'objectiu del model és fer previsions del nombre de comentaris a un mes, segons el nombre de *followers* d'un usuari, com s'interpreten els dos indicadors? [1pt]

H. Estimeu el pendent amb un interval de confiança del 95% [2pts]

(podeu utilitzar l'altra cara del full)