2NOM:___________________________COGNOM:____

Contesteu cada pregunta en el seu lloc. Expliciteu justifiqueu els càlculs.

Problema 1 (A)

Segons les dades d'una enquesta d'intoleràncies alimentaries, aquestes es distribueixen de la següent manera segons si la edat es superior a 45 anys o inferior o igual a 45 anys. Es van anotar les intoleràncies al gluten (G), Lactosa (L) i a Cap de les dues anteriors (C). La taula següent presenta les probabilitats corresponents:

Edat	Gluten (G)	Lactosa (L)	Cap de les anteriors (C)
Edat>45	0.1	0.05	0.3
Edat<=45	0.25	0.12	0.18

a) (0.5 punts) Les probabilitats a la taula són probabilitats condicionades o conjuntes? Raoneu la resposta.

Solució:

Es tracta de probabilitats conjuntes ja que la suma de totes les probabilitats a la taula es 1.

b) (0.5 punts) Les variables Edat i Intolerància son dependents, expliqueu com ho podem saber.

Solució:

Ho sabem perquè les distribucions condicionades són diferents. La probabilitat d'una intolerància varia d'un grup d'edat a l'altre.

c) (1 punt) Si fossin independents, com quedaria la taula anterior? Heu de procurar mantenir les mateixes proporcions pels grups d'edat, i també pels grups d'intolerància.

Solució:

Edat	Gluten (G)	Lactosa (L)	Cap de les anteriors (C)
Edat>45	0.1575	0.0765	0.216
Edat<=45	0.1925	0.0935	0.264

P(G)= 0,1+0,25 Calculem la intersecció de cada parella per posar a la taula fent les multiplicacions

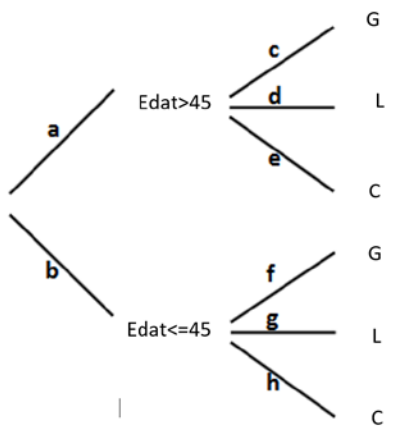
P(L)=0,05+0,12 de les 2 probabilitats que corresponen.

P(C)=0,3+0,18

P(Edat>45)=0,1+0,05+0,3

P(Edat<=45)=0.55

d) (1 punt) Calculeu les probabilitats a fins a h del següent arbre indiqueu les fórmules necessàries per fer els càlculs.



Solució:

Probabilitats marginals a i b:

P(Edat<45) = P(Edat>45,G) + P(Edat>45, L) + P(Edat>45) = 0.45 → P(Edat<=45) = 0.55

• Calcul de les probabilitats condicionades c fins a h:

P(G|Edat>45) = P(G, Edat>45)/P(Edat>45)=..

...

P(C|Edat<=45) =P(C,Edat<=45)/P(C|Edat<=45)=..

a=0.45; b=0.55; c= 0.1/0.45=0.22 ; d=0.05/0.45=0.11; e=0.3/0.45=0.67 ; f= 0.25/0.55=0.45 ; g=0.12/0.55=0.22; h=0.18/0.55=0.33.

e) (1 punts) Si una persona de qualsevol edat no té intolerància al gluten, quina és la probabilitat de que sigui més gran de 45 anys?

Solució:

P(Edat>45|L+C)=[P(Edat>45 intersecció L)+ P(Edat>45 intersecció C)]/[P(L)+P(C)]=[0.05+0.3]/[0.17+0.48]=0.35/0.65=0.538

Disposem de 2 tests diferents per detectar toleràncies alimentaries (1=Si, 0=No), però com no són del tot fiables es sol utilitzar els dos amb el mateix pacient. El resultats del 1r test es recullen a la variable aleatòria X1 i els del segon test a la variable aleatòria X2 i a continuació s'indiquen les probabilitats de les combinacions dels possibles resultats:

	TEST 2	
TEST 1	**0**	**1**
0	0.1	0.2
1	0	0.7

f) (0.5 punts) Calculeu l'esperança dels resultats obtinguts en cadascun dels tests.

Solució:

$$\mathbb{E}[X_1] = 0 \cdot 0.1 + 0 \cdot 0.2 + 1 \cdot 0 + 1 \cdot 0.7 = 0.7$$
$$\mathbb{E}[X_2] = 0 \cdot 0.1 + 1 \cdot 0.2 + 0 \cdot 0 + 1 \cdot 0.7 = 0.9$$

g) (1 punt) Amb els dos tests es construeix una nova escala igual a 4 TEST1 + 6 TEST2. Trobeu quina és la variància d'aquesta escala.

Solució:

$Var(Y)=E(Y^2)-E(Y)^2$

$Var(4 X_1 + 6X_2)=E(4^2 X_1^2+6^2 X_2^2+48X_1X_2)- E(4 X_1+6 X_2)^2$

$= 4^2E(X_1^2)+6^2 E(X_2^2)+48E(X_1X_2)-(4 E(X_1)+6E(X_2))^2$

$=4^2(0,7-0,7^2)+6^2 (0,9-0,9^2)+48(0,7-0,7*0,9) \quad =9.96$

h) (1.5 punts) Calculeu la covariància dels resultats obtinguts en els dos test.

Solució:

$$\begin{aligned}\mathrm{Covar}(X_1, X_2) =&(0.1)(0-0.7)(0-0.9)\\ &+(0.2)(0-0.7)(1-0.9)\\ &+(0)(1-0.7)(0-0.9)\\ &+(0.7)(1-0.7)(1-0.9)\\ &=0.07\end{aligned}$$

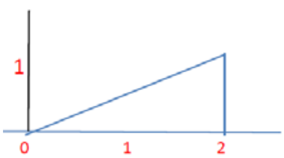
Es pot trobar la covariància d'una altra manera: $cov(X_1,X_2)=E(X_1X_2)-E(X_1)*E(X_2)=0.07$; serà igualment correcte. A més, amb la COV es pot resoldre més fàcilment l'apartat de la variància de la nova escala (16 * 0.7 * 0.3 + 36 * 0.9 * 0.1 + 2*4*6*0.07 = 9.96)

El temps entre les realitzacions dels 2 tests (TD) va d'entre 0 i 2 hores i segueix la funció lineal creixent: $f_{TD}(t)$ = t/2.

i) (0.5 punts) Representeu amb un gràfic esquemàtic la funció de densitat

Solució:

$f_{TD}(t) = \frac{1}{2}t \quad (0 \le t \le 2)$



j) (1 punts) Trobeu la funció de distribució de probabilitat de la variable.

Solució:

$$F_{TD}(t) = \tfrac{1}{4}t^2 \quad (0 \le t \le 2)$$

$$\int_0^t \tfrac{1}{2}x\,dx = \tfrac{1}{4}x^2 \,|_0^t = \tfrac{1}{4}t^2$$

k) (0.5 punts) Calculeu la probabilitat que el temps entre 2 tests sigui inferior a 1 hora.

Solució:

$P(TD < 1) = F_{TD}(1) = \frac{1}{4}1^2 = 0.25$

l) (1 punt) Trobeu la mitjana del temps (el valor esperat) i la mediana del temps (la meitat dels pacients triguen aquest temps entre test i test), *en minuts*.

Solució:

$E(x)=\int_0^2 x\, f(x)dx=\int_0^2 x\frac{x}{2}\, dx$=1,3333hores->80 minuts

$Med=\int_0^X f(x)dx$=0,5 ->(1/4) t^2 avaluat a x i 0 → (1/4) x^2 =0 →x=sqrt(2) hores → 84,85 minuts

Problema 2 (BB)

Un departament de la universitat CPU té un parell d'impressores, PH i Conan, que no sempre funcionen bé. En particular, de la impressora PH s'ha observat al llarg del temps que un de cada 12 documents enviats a imprimir no s'imprimeix bé, o fins i tot no s'imprimeix.

(a) (0,6 punts)
Si un professor envia 10 documents per setmana a imprimir a la impressora PH, quina és la probabilitat que tots s'imprimeixin bé. Indiqueu quina distribució feu servir per a aquest càlcul.

Solució:

$$X \sim \mathcal{B}(10, 11/12) \Longrightarrow \mathrm{P}(X = 10) = 0.419.$$

(b) (0,6 punts)
Quin és el valor esperat d'aquesta variable? Doneu una interpretació del valor obtingut.

Solució:

$$\mathrm{E}(X) = 10 \cdot 11/12 = 9.17.$$

La mitjana de documents que s'imprimeixen bé (entre cada 10 documents enviats) al llarg del temps s'acosta cada cop més a 9.17.

(c) (0,6 punts)
Quina és la distribució de la variable 'Nombre d'enviaments de documents a la impressora fins s'hagin imprès bé els deu documents? Quin és el seu valor esperat?

Solució:

$$Y \sim \mathcal{BN}(10, 11/12) \Longrightarrow \mathrm{E}(Y) = \frac{10}{11/12} = 10.9.$$

De la impressora Conan se sap per experiència que la mitjana de incidències que requereixen l'ajuda de l'informàtic en una setmana de 5 dies i 10 hores laborables diàries (de 8 a 18h) és 4. Se sap també que els problemes d'aquesta impressora són independents dels problemes de la impressora PH.

(d) (0,6 punts)
Quin és el valor esperat de nombre d'incidències en un dia? Quina distribució feu servir?

Solució:
Sigui Z_1 la variable 'Nombre de incidències en 1 dia'.

$$Z_1 \sim \mathcal{P}(4/5 = 0.8) \Longrightarrow \mathrm{E}(Z_1) = 0.8.$$

(e) (0,8 punts)
Si entre dilluns i dimecres hi hagut 4 incidències, quina és la probabilitat que el nombre total d'incidències setmanals superi 6?

Solució:
Siguin Z_2, Z_3 i Z_5 les variables 'Nombres d'incidències en 2, 3 i 5 dies', respectivament.

$$\mathrm{P}(Z_5 > 6 | Z_3 = 4) = \mathrm{P}(Z_2 > 2) = 1 - \mathrm{P}(Z_2 \leq 2) = 0.217,$$

on $Z_2 \sim \mathcal{P}(1.6)$.

(f) (1 punt)
Un matí qualsevol, entre les 8 i 14h, s'envien 6 documents a imprimir a la impressora PH. Calculeu la probabilitat que aquest dia l'informàtic del departament passi un matí (de 8 a 14h) tranquil, és a dir sense cap incidència de cap de les dues impressores.

Solució:
Siguin X_6 i Z_M les variables 'Nombre de documents (entre 6) que imprimeixi bé PH' i 'Nombre d'incidències de Conan al llarg d'un matí', respectivament. Llavors, utilitzant independència entre les dues impressores:

$$\begin{aligned}\mathrm{P}(\text{Matí tranquil per al informàtic}) &= \mathrm{P}(X_6 = 6) \cdot \mathrm{P}(Z_M = 0) \\ &= \left(\frac{11}{12}\right)^6 \cdot e^{-0.8 \cdot 6/10} = 0.593 \cdot 0.619 = 0.367.\end{aligned}$$

(g) (0,8 punts)
Quina és la distribució de la variable 'Temps entre dues incidències de la impressora Conan' i quin és el seu valor esperat (en hores laborables)?

Solució:
Sigui T la variable 'Temps entre dues incidències'.

$$T \sim Exp(0.8/10 \text{ (per hora)}) \Longrightarrow \mathrm{E}(T) = 10/0.8 = 12.5 \text{ hores}.$$

(h) (1 punt)
Quina és la probabilitat que durant almenys 5 de les 6 hores del matí no hi hagi cap incidència de la impressora Conan?

Solució:
Càlcul de la probabilitat que no hi hagi cap incidència de Conan durant una hora:

$$\mathrm{P}(T > 1) = e^{-0.8/10} = 0.923.$$

Nota 1: També es pot fer el càlcul amb la distribució Poisson amb paràmetre $0.8/10 = 0.008$.
Sigui ara H la variable 'Hores al matí sense cap incidència'. Per tant:

$$H \sim \mathcal{B}(6, 0.923) \Longrightarrow \mathrm{P}(H \geq 5) = \mathrm{P}(H = 5) + \mathrm{P}(H = 6) = 0.928.$$

Nota 2: Per a la següent solució es dona **0,7 punts**:

$$\mathrm{P}(T \geq 5) = e^{-0.8/10 \cdot 5} = 0.67.$$

Si la impressora PH no dona problemes, el temps (en segons) d'imprimir una prova de seguiment de l'assignatura Ep per a 20 estudiants segueix una distribució normal amb paràmetres $\mu = 30$ i $\sigma = 2$.

(i) (0,6 punts)
Calculeu la probabilitat que el temps d'impressió superi 35 segons? Feu servir per a aquest càlcul un dels valors de les funcions `*norm` a la taula al final d'aquest full.

Solució:
Sigui S la variable 'Temps (en segons) per imprimir una prova de seguiment d'Ep per a 20 estudiants'.

$$S \sim \mathcal{N}(30, 2) \Longrightarrow \mathrm{P}(S > 35) = 1 - \mathrm{P}(S \leq 35) = 1 - \mathrm{P}(Z \leq 2.5) = 0.006,$$

on Z segueix una distribució normal estandaritzada.

(j) (1,2 punts)
Assumint independència entre la impressió d'una prova de seguiment i la següent, calculeu la probabilitat que la diferència de temps entre una impressió i la següent sigui de 5 segons o més.

Solució:
Sigui D la variable 'Diferència de temps d'impressió de dues proves de seguiment d'Ep per a 20 estudiants'.

$$D \sim \mathcal{N}(0, \sqrt{4+4} = 2.83) \Longrightarrow \mathrm{P}(|D|) > 5) = 2 \cdot \mathrm{P}(D < -5) = 2 \cdot \mathrm{P}(Z < -5/2.83 = -1.77) = 0.076,$$

on Z segueix una distribució normal estandaritzada.

La quantitat de fulls que gasta un professor qualsevol del departament setmanalment amb les impressions a la impressora Conan es pot modelar amb una distribució uniforme amb paràmetres 0 i 100 (fulls).

(k) (0,6 punts)
Quins són el valor esperat, la mediana i la desviació de la variable 'Nombre de fulls setmanals impresses (per professor/a)?

Solució:

$$U \sim \mathcal{U}_{[0,\,100]} \Longrightarrow \mathrm{E}(U) = \mathrm{Med}(U) = 50,\ \mathrm{SD}(U) = \sqrt{100^2/12} = 28.87.$$

(l) (1,6 punts)
Amb quants fulls s'ha de carregar la impressora Conan a l'inici de la setmana per tal que la probabilitat que hi hagi fulls suficients per als 20 professors i professores del departament per a tota la setmana amb una probabilitat del 95%? Feu servir per a aquest càlcul un dels valors de les funcions `*norm` a la taula al final d'aquest full.

Solució:
Sigui S_{20} la suma de tots els fulls gastats setmanalment pels 20 professors i professores. Llavors,

$$S_{20} \sim \mathcal{N}(20 \cdot 50 = 1000, \sqrt{20} \cdot 28.87 = 129.11)$$

$$s_{20;0.95} = 1000 + 129.11 \cdot 1.645 = 1212.386,$$

on 1.645 és el quantil del 95% de la distribució normal estandaritzada. Per tant, s'hauria de carregar la impressora amb 1213 fulls a l'inici de la setmana per tal de garantir que la probabilitat que no es quedi sense paper durant la setmana sigui 0.95.