NOM:_____ COGNOM:_____

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

Problema 1 (B1)

En un estudi sobre els lectors de dues publicacions de distribució gratuïta entre els estudiants universitaris es troba que el 40% dels estudiants universitaris llegeixen el diari *Vida universitària,* el 35% el diari *Campus* i el 15% ambdues publicacions.

1) A partir de les dades de l'estudi, completa la següent taula de doble entrada (0,5 punts):

	Llegeix el diari *Vida Universitària*	No llegeix el diari *Vida Universitària*	
Llegeix el diari *Campus*			
No llegeix el diari *Campus*			
			100

SOLUCIÓ:

	Llegeix el diari *Vida Universitària*	No llegeix el diari *Vida Universitària*	
Llegeix el diari *Campus*	15	20	35
No llegeix el diari *Campus*	25	40	65
	40	60	100

2) Quina és la probabilitat que un estudiant universitari llegeixi algun d'aquests diaris? (1 punt)

Anomenem VU a l'esdeveniment "Llegir el diari *Vida Universitària*" i C a l'esdeveniment "Llegir el diari *Campus*". Aleshores es té que,

$P(VU \cup C) = P(VU) + P(C) - P(VU \cap C) = 40/100 + 35/100 – 15/100 = 60/100 = 0,6$

3) Quina és la probabilitat que només llegeixi un dels diaris? (1 punt)

$P(\text{només un dels dos diaris}) = P(VU \cap \neg C) + P(\neg VU \cap C) = 25/100 + 20/100 = 45/100 = 0,45$

4) Sabent que un estudiant no llegeix el diari *Campus, q*uina és la probabilitat que llegeixi el diari *Vida Universitària*? (1 punt)

$P(VU \mid \neg C) = P(VU \cap \neg C) / P(\neg C) = 25/100 : 65/100 = 25/65 = 5/13 = 0,3846$

En el diari *Campus* els articles estan escrits un 50% en OpenOffice, un 30% en Word i un 20% en Latex. Actualment estan fent un estudi de les errades tipogràfiques que es troben en els articles del seu diari segons en quina eina hagin estat escrits. A partir de l'estudi, han trobat que les errades tipogràfiques tenen un 0,1 de probabilitat d'ocórrer si s'ha escrit en OpenOffice, un 0,6 si s'ha escrit en Word i un 0,3 si s'ha escrit en Latex.

5) Quina és la probabilitat que hi hagi una errada tipogràfica en un article del diari escollit a l'atzar? (1'5 punts)

$P(\text{Errada}) = P (\text{Errada} \cap \text{OpenOffice}) + P (\text{Errada} \cap \text{Word}) + P (\text{Errada} \cap \text{Latex}) =$
$= P (\text{Errada} \mid \text{OpenOffice}) \cdot P (\text{OpenOffice}) + P (\text{Errada} \mid \text{Word}) \cdot P(\text{Word}) + P (\text{Errada} \mid \text{Latex}) \cdot P(\text{Latex}) =$
$= 0.1 \cdot 0.5 + 0.6 \cdot 0.3 + 0.3 \cdot 0.2 = 0.29$

6) Quina és la probabilitat que si hi ha una errada, l'article hagi estat escrit en Latex? (1'5 punts)

P(Latex | Errada) = P (Latex ∩ Errada) / P (Errada) = P (Errada | Latex) · P(Latex) / P(Errada) = 0.3 ·0.2 / 0.29 = 0.2069

En el diari *Vida universitària* reparteixen els diaris amb tres empreses diferents: Empresa A, Empresa B i Empresa C amb el compromís de fer el lliurament abans de les 8 del matí i amb una penalització sinó compleixen amb el termini establert. A partir de les dades, les probabilitats que hi hagi un endarreriment en el lliurament es troben en la següent taula:

	Empresa A	Empresa B	Empresa C
Lliurament amb endarreriment	0.07	0.10	0.06
Lliurament sense endarreriment	0.93	0.90	0.94

7) Les probabilitats que es troben a la taula són probabilitats condicionades o conjuntes? Raona la resposta i indica de manera explícita a quina probabilitat correspon el valor 0.07 de la taula. (1 punt)

Són probabilitats condicionades ja que la suma és 1 per cada columna. En concret el condicionament és per cada empresa, així 0.07 = P(Lliurament amb endarreriment | Empresa A).

8) Indica si les variables "Empresa de repartiment" i "Lliurament" són o no independents. Raona adequadament la resposta justificant-la a partir de les dades donades. (1 punt)

No són independents perquè, per exemple:
P(Lliurament amb endarreriment | Empresa A) = 0.07 ≠ 0.10 = P(Lliurament amb endarreriment | Empresa B)

És a dir, la probabilitat de "Lliurament amb endarreriment" és diferent si se sap l'empresa de repartiment i per tant, no hi ha independència dels esdeveniment.

9) Se sap que l'empresa A gestiona el 50% dels lliurament. Si es vol aconseguir que els lliuraments amb endarreriment siguin del 8%, quin tant per cent de lliurament ha de gestionar l'empresa B? (1'5 punts)

Anomenem p a la P(Empresa B). Com P(Empresa A) + P(Empresa B) + P(Empresa C) = 1, P(Empresa C) = 0.5-p

P(Lliurament amb endarreriment) =
P(Lliurament amb endarreriment ∩ Empresa A) +
P(Lliurament amb endarreriment ∩ Empresa B) +
P(Lliurament amb endarreriment ∩ Empresa C) =
P(Lliurament amb endarreriment | Empresa A) · P(Empresa A) +
P(Lliurament amb endarreriment | Empresa B) · P(Empresa B) +
P(Lliurament amb endarreriment | Empresa C) · P(Empresa C) = 0.07 ·0.5 + 0.1· p + 0.06· (0.5-p) = 0.08

Per tant, 0.035 + 0.1p + 0.03 -0.06p = 0.04p + 0.065 = 0.08 i p=0.375

L'empresa B ha de gestionar el 37,5% dels repartiments.

PROBLEMA 2 (B2). SOLUCIÓ

Per les festes de Sant Josep, a l'escola han instal·lat una diana per a que els nens practiquin tir amb arc. Es sap que la meitat dels tirs no impacten a la diana. De l'altra meitat, un 50% valen 10 punts, un 30% 20 punts, el 12% obtenen 30 punts i la resta assoleixen el màxim, 100 punts. Per als nens grans, els tirs fora de diana compten 0 punts, per als nens petits els deixen repetir fins a impactar a dins.

1. Especifiqueu les funcions de probabilitat de la puntuació obtinguda per nens grans i per nens petits.

   G: variable aleatòria "puntuació un tir de nen gran"

   T: variable aleatòria "puntuació un tir de nen petit"

g	P(G=g)	
0	0,5	"la meitat dels tirs no impacten ..." (0 punts)
10	0,25	"De l'altra meitat, un 50% valen 10 punts"
20	0,15	"...un 30% 20 punts"
30	0,06	"...el 12% obtenen 30 punts"
100	0,04	"...la resta assoleixen el màxim, 100 punts"

t	P(T=t)
10	0,5
20	0,3
30	0,12
100	0,08

2. Compareu les puntuacions esperades i les variàncies. En quin grup la variable aleatòria "puntuació" és més dispersa? Per què?

   $E(G) = 11,3$

   $E(T) = 22,6$

   $E(G^2) = 539$; $V(G) = 539 - 11,3^2 = 411,31$

   $E(T^2) = 1078$; $V(T) = 1078 - 22,6^2 = 567,24$

   Veiem que la variància més gran és la de la puntuació dels nens petits, encara que el rang de punts és més estret (de 10 a 100, en lloc del rang dels grans, de 0 a 100). La raó és que el valor 0 concentra la meitat dels valors dels grans.

3. Els petits fan una competició, consistent en dos tirs vàlids (a la diana). La taula del costat mostra com es distribueixen conjuntament el primer i el segon tir. Si un nen ha obtingut la primera vegada 30 punts, com es distribueix la probabilitat del seu següent tir?

		2n tir			
		10	20	30	100
1er tir	10	0,4	0,06	0,04	
	20	0,05	0,2	0,05	
	30		0,02	0,08	0,02
	100			0,03	0,05

   Tenim T1 (primer tir, files) i T2 (segon tir, columnes).

   Es demana $P(T2=t \mid T1=30) = = \frac{P(T1=30 \cap T2=t)}{P(T1=30)}$

   Fixem-nos en la tercera fila: [0 0,02 0,08 0,02]. Veiem que suma 0,12, que ja havíem obtingut a l'apartat 1 per la variable T. De fet, la distribució de la variable T1 coincideix amb la de T. Per tant, el denominador val 0,12.

t	P(T2=t | T1=30)
10	0
20	0,02/0,12 = 1/6
30	0,08/0,12 = 2/3
100	0,02/0,12 = 1/6

4. Digueu si us sembla que el segon tir en general millora o empitjora el resultat previ. Justifiqueu-ho

t	10	20	30	100
P(T2=t)	0,45	0,28	0,2	0,07

   (sumant-hi les columnes)

   Per respondre a la qüestió "en general", utilitzarem el valor esperat de T2. Amb els valors anteriors:

   $E(T2) = 10 \cdot 0,45 + 20 \cdot 0,28 + 30 \cdot 0,20 + 100 \cdot 0,07 = 23,1 > 22,6 = E(T1)$

   Per tant, en mitjana el segon tir és lleugerament millor (0,5 punts més) que el primer.

5. Descriviu la relació entre els dos tirs, i trobeu la correlació entre ambdues puntuacions (ajuda: la desviació estàndard de la suma de punts val 43.157 punts).

En primer lloc, veiem a la taula que si un nen fa un tir baix la primera vegada és molt probable que el segon també sigui baix. I si és un bon tir, també és probable que repeteixi, per tant podem dir que la relació és directa o positiva, i bastant forta, perquè no és molt probable (o impossible) obtenir dos tirs molt diferents.

Aquesta propietat es pot resumir amb un indicador, la correlació entre T1 i T2:

cor(T1, T2) = cov(T1, T2)/σ(T1)/σ(T2)

$\sigma(T1) = \sqrt{V(T1)}$ =23,817. De manera semblant a l'apartat 2, calculem la variància de T2 i d'aquí la desviació tipus:

V(T2) = 503,39; σ(T2) = 22,436

Per l'enunciat, sabem que σ(T1+T2) = 43,157, o V(T1+T2) = 1862,5

V(T1+T2) = V(T1) + V(T2) + 2 cov(T1,T2) →cov(T1, T2) = ½ (V(T1+T2) – V(T1) – V(T2))

$$\text{cor}(T1, T2) = \frac{\frac{1}{2}(1862{,}5 - 5\quad{,}24 - \quad\quad{,}39)}{23{,}817\ 22{,}436} = 0{,}741$$

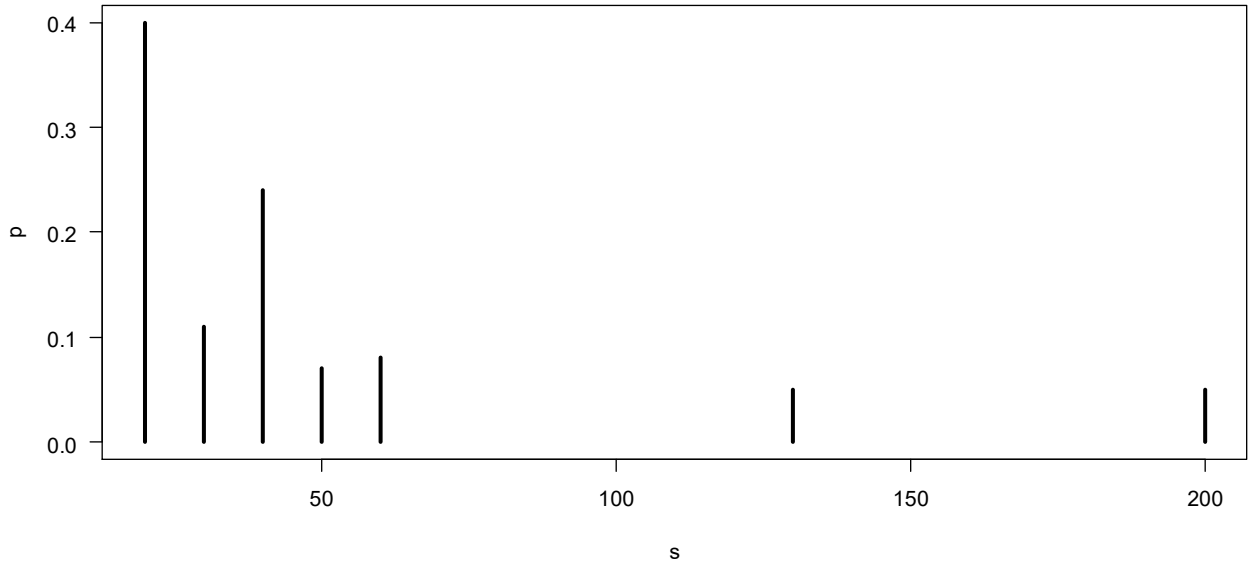
6. Representeu gràficament la funció de probabilitat de la suma de puntuacions dels dos tirs.

Primer observem com es distribueix a la taula conjunta la suma de puntuacions (només si probabilitat > 0):

		2n tir			
		10	**20**	**30**	**100**
1er tir	**10**	20	30	40	
	20	30	40	50	
	30		50	60	130
	100			130	200

Llavors, agrupem les probabilitats de la taula per valors de la suma de puntuacions (S = T1+T2):

s	P(S=s)
20	0,4
30	0,06+0,05 = 0,11
40	0,04 + 0,2 = 0,24
50	0,05 + 0,02 = 0,07
60	0,08
130	0,02 + 0,03 = 0,05
200	0,05



L'aeroport de Barcelona ha fet un estudi per veure quants viatgers, just abans d'embarcar, donen positiu en una prova de COVID-19 que es realitza de manera individualitzada. Considerant que hi ha 200 passatgers en un avió i que la probabilitat que un passatger doni positiu és del 1%, es plantegen les següents preguntes:

Nota: comenta tots els resultats obtinguts.

1. Quin és el model de probabilitat que ens permet estudiar el nombre de viatgers embarcant a un avió concret detectats per COVID? (1 punt)
   N: "Nombre de viatgers afectats per COVID en un embarcament "~Bin(n=200,p=0,01)
   Una premissa indispensable és que els viatgers són independents uns d'altres. Això vol dir que si al mateix vol hi va un grup (gran) de convivents, el model no és adequat.
2. Quina és la probabilitat que en un vol es detectin 3 o més positius? (1,5 punts)
   P(N>=3)=1-P(N<=2)=1- 0,676=0,324
3. En mitjana, quants positius esperem detectar en cada embarcament? Quant val la desviació tipus? (1 punt)
   E(N)=np=2 (passatgers positius per embarcament)
   Desv(N)=sqrt(np(1-p))=1,407
4. L'aeroport gestiona 400 vols diaris i cadascun d'ells té uns 200 passatgers. Quina és la probabilitat que un dia es detectin més de 500 positius a l'aeroport? (1,5 punts)
   Sabent que hi ha 400 vols diaris farem servir l'aproximació Normal
   N2:"Nombre de passatgers de l'aeroport al dia que donen positiu"~N(400*200*0,01=800;sqrt(400*200*0,01*0,99)=28.14)
   P(N2>500)=1-P(N2<500)=1-P(Z<(500-800)/28.14)=1-P(Z<-10,66)=1-0=1
   Amb tota seguretat, donades les condicions, es detectaran més de 500 positius en un dia a l'aeroport.

La realització de les proves requereix esforços organitzatius addicionals. Se sap que el temps que es triga en realitzar una prova segueix una distribució Normal de valor esperat igual a 15 minuts i desviació tipus de 5 minuts, incloent el temps d'espera, realització i obtenció dels resultats de la prova.

5. Quina probabilitat hi ha que una persona trigui més de 30 minuts a completar el procés? (1 punt)
   Donades les condicions de l'enunciat:
   T: "Temps de realització de la prova"~N(15;5)
   P(T>30)=1-P(T<30)=1-P(Z<(30-15)/5)=1-P(Z<3)=1- 0,9987 = 0,0013

6. En els espais habilitats per fer la prova arriben en mitjana 5 passatgers cada hora. Volem calcular la capacitat que haurien de tenir els espais per tal que la probabilitat de saturar-se en una hora no arribi a 5% (1 punt):
   M: "Nombre de passatgers que arriben a l'espai habilitat"~Poiss(λ=5)
   Es demana k tal que P(M > k) < 0.05. Per taules, F(9) = 0.968. 9 places.

7. Les autoritats de l'aeroport estan interessades en saber la probabilitat que entre dues arribades de passatgers al mateix espai passi més de 45 minuts. Trobeu aquesta probabilitat i interpreteu (1,5 punts):
   T: "Temps entre les arribades dels passatgers a l'espai per fer-se la prova"~Exp(λ=5)
   L'alumne hauria de comentar que són 5 passatgers/hora

   Es demana P(T>0.75) = 1-P(T<=0.75) = 1-(1-exp(-0.75·5)) = exp(-3.75) = 0.0235
   Si en mitjana hi ha un passatger cada 12 minuts, és bastant rar (2.3%) que passin 45 sense arribades.

8. Hem de donar una recomanació als passatgers perquè arribin amb temps suficient: concretament, voldríem indicar un temps que seria suficient per al 99% dels viatgers per a tramitar la prova a l'aeroport. Quants minuts abans haurien d'estar a l'aeroport respecte l'escenari pre-pandèmia? Com escriuríeu la recomanació (incorporant el resultat trobat)? (1,5 punts)
   $15 + z_{0.99}\ 5 = 15 + 2.33 * 5 = 26.65$
   La recomanació seria quelcom com "Reservar 26.65 minuts per la prova a l'aeroport és suficient temps per al 99% dels viatgers"