NOM:____

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

Problema 1 (B1).

El Sr. MR valora si presentar-se a unes oposicions en les que ha de defensar oralment 3 temes escollits a l'atzar de una llista de 30 possibles. Quan veu el temari, observa que el temps necessari per preparar segueix el model de Pareto: per estudiar 24 temes (24 de 30: 80%) necessitarà 6 mesos, però pels 6 restants (20%) necessitarà 2 anys addicionals

1) **(2p)** Si el proper concurs tindrà lloc d'aquí 6 mesos i MR només estudia aquest 80% dels temes, quina probabilitat hi ha que els 3 temes siguin dels que sí que ha preparat?

P (Sí el 1 ∩ Sí el 2 ∩ Sí el 3) = $P^S_1 * P^S_2 * P^S_3 = 24/30 * 23/29 * 22/28 \approx 0.4985 \approx 0.50$ (2p)

Però: si NO indica càlculs **o** $0.8^3 = 0.512 \approx 0.51$ (1p)

I si NO indica càlculs **i** $0.8^3 = 0.512 \approx 0.51$ (0p)

En realitat, les normes diuen que l'estudiant ha de defensar 2 dels 3 temes seleccionats a l'atzar.

2) **(3p)** Quina probabilitat hi ha que MR pugui seleccionar 2 temes dels que sí ha preparat?

P (nº Sí >= 2) = $P^S_1 \wedge P^S_2 \wedge P^S_3 \cup P^S_1 \wedge P^S_2 \wedge P^N_3 \cup P^S_1 \wedge P^N_2 \wedge P^S_3 \cup P^N_1 \wedge P^S_2 \wedge P^S_3 =$

$= P^S_1 * P^S_2 * P^S_3 + P^S_1 * P^S_2 * P^N_3 + P^S_1 * P^N_2 * P^S_3 + P^N_1 * P^S_2 * P^S_3 =$

=24/30*23/29*22/28 + 24/30*23/29*6/28 + 24/30*6/29*23/28 +6/30*24/29*23/28

= =24/30*23/29*22/28 + 3*((24*23*6)/(30*29*28)) ≈ 0.5045 + 0.4079 ≈ 0.9133 ≈ 0.91 (3p)

Però: si NO indica càlculs **o** ho considera com a reposició (0.8 i 0.2 constants) (2p)

I si NO indica càlculs **i** ho considera com a reposició (0.8 i 0.2 constants) (1p)

Suposem que sempre s'aprova si surten 2 o més temes dels que ha estudiat.

3) **(1p)** Quina probabilitat hi ha que MR suspengui?

P (nº Sí < 2) =1 - P (nº Sí >= 2) ≈ 1- 0.9133 ≈ 0.0867≈ 0.09 (1p)

4) **(3p)** L'any passat el cosí de MR també es va presentar a les oposicions amb la mateixa estratègia. Si sabem que va suspendre les proves (B), quina probabilitat hi ha que cap dels temes sigui dels que s'havia preparat (A)

A = "cap dels temes sigui dels que s'havia preparat"

P(A) = P $[p^N_1 \wedge p^N_2 \wedge p^N_3]$ =6/30*5/29*4/28 ≈ 0,004926108 ≈ 0,005

B= "suspendre les proves"

P (nº Sí < 2) = 1 - P (nº Sí >= 2) ≈ 1- 0.9133 ≈ 0.0867≈ 0.09

P (A | B) = P(A ∩ B) / P(B) = P(A) / P(B) ≈ 0.005 / 0.09 ≈ 0,056829127 ≈ 0,06 (3p)

5) **(1p)** Proposi una regla que minimitzi el paper de l'atzar. És a dir, alguna forma de minimitzar que un estudiant pugui aprovar simplement per que tingui la sort de que li toquin els temes que ha estudiat (no cal demostrar-ho).

Augmentar la quantitat d'informació, fent un examen de, p.e., 5 preguntes. (1p)

NOM:____

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

Problema 2 (B2).

Denotem amb X i Y el nombre de compres al mes realitzades a dos proveïdors A i B respectivament. La següent taula expressa la funció de probabilitat conjunta de X i Y:

		Proveïdor B		
		1	***2***	***3***
Proveïdor A	***1***	0	0.0	0.3
	2	0.1	0.2	0
	3	0.4	0	0

Indiqueu les funcions de probabilitat i de distribució de probabilitat de X i de Y (1 punt)

k	$p_X(k)$	$F_X(k)$	k	$p_Y(k)$	$F_Y(k)$
1	0.3	0.3	1	0.5	0.5
2	0.3	0.6	2	0.2	0.7
3	0.4	1	3	0.3	1

Quantes compres s'espera fer mensualment, en mitjana, al proveïdor A? I al proveïdor B? (1 punt)

E(X) = 1*0.3 + 2*0.3 + 3*0.4 = 0.3 + 0.6 + 1.2 = 2.1

E(Y) = 1*0.5 + 2*0.2 + 3*0.3 = 0.5 + 0.4 + 0.9 = 1.8

Quant val la desviació típica del nombre de compres mensual al proveïdor A? (1 punt)

$\mathrm{Sqrt}(V(X)) = \mathrm{sqrt}\ ((1-2.1)^2 0.3 + (2-2.1)^2 0.3 + (3-2.1)^2 0.4) \approx 0.83$

$\mathrm{Sqrt}(V(Y)) = \mathrm{sqrt}\ ((1-1.8)^2 0.5 + (2-1.8)^2 0.2 + (3-1.8)^2 0.3) \approx 0.87$

Quina és la probabilitat de que observem més d'una compra al mes al proveïdor A? (1 punt)

P(X>1) = P(X=2)+P(X=3) = 0.3 + 0.4 = 0.7
= 1 – P(X<=1) = 1-0.3 = 0.7

Calculeu la probabilitat que un mes el nombre de compres a cada proveïdor sigui la mateixa (1 punt)

P(X=1,Y=1) + P(X=2,Y=2) + P(X=3,Y=3) = 0+0.2+0 = 0.2

Amb quina probabilitat es realitzaran 2 compres al proveïdor B si se n'ha fet 1 al proveïdor A? I 2 compres al proveïdor A si se n'ha fet 1 al proveïdor B ? (1 punt)

$P(Y=2|X=1) = P(Y=2,X=1) / P(X=1) = 0 / 0.3 = \mathbf{0}$

$P(X=2|Y=1) = P(X=2,Y=1) / P(Y=1) = 0.1 / 0.5 = \mathbf{0.2}$

Calculeu el valor de la covariància i de la correlació de les dues variables considerades (2 punts)

$$Cov(X,Y) = (1-2.1)(1-1.8)0+(1-2.1)(2-1.8)0+(1-2.1)(3-1.8)0.3+(2-2.1)(1-1.8)0.1+(2-2.1)(2-1.8)0.2+(2-2.1)(3-1.8)0$$
$$+(3-2.1)(1-1.8)0.4+(3-2.1)(2-1.8)0+(3-2.1)(3-1.8)0 =$$
$$= -0.4 + 0.008 - 0.004 - 0.288 = \mathbf{-0.68}$$

$Corr(X,Y) = -0.68 / (0.83*0.87) = \mathbf{-0.94}$

($Sqrt(V(Y)) = sqrt ((1-1.8)^2 0.5 + (2-1.8)^2 0.2 + (3-1.8)^2 0.3) = 0.87$)

Interpreteu els resultats de la covariància i la correlació. Relacioneu-ho amb la dependència o independència de les compres a un i altre proveïdor i també amb el altres resultats (1 punt)

Covariància negativa, correlació negativa i propera a -1.
Per tant relació negativa, inversa, entre el nombre de compres en un proveïdor i l'altre.
Si es compra molt a un proveïdor no es compra molt a l'altre.

La correlació negativa indica que no són independents les compres a un proveïdor i a l'altre. La dependència és que si es compra molt a un proveïdor no es compra molt a l'altre. A més a més, la mitjana de compres al proveïdor A és més alta que al proveïdor B, tot i que la variabilitat (desviació) també és més gran.

Quant val el valor esperat del nombre de compres total (al proveïdor A més al proveïdor B)? (1 punt)

$E(X+Y) = E(X) + E(Y) = 2.1 + 1.8 = \mathbf{3.9}$

NOM:____

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

Problema 3 (B3).

Segons alguns estudiosos melòmans, les cançons dels grups punkies tenen una durada mitjana de dos minuts i mig (en segons, 150) i aquesta durada es distribueix com una campana de Gauss (o model Normal).

1. Quina és la desviació tipus de la durada d'una cançó punk, sabent que només una de cada 5 supera els 166 segons? *(1.5pt)*

T: "durada en segons d'una cançó punk" ~ N(μ=150, σ)

Sabem que $P(T > 166) = 1/5$, i d'aquí deduir σ.

$Z = (T - \mu) / \sigma$

$P(Z > (166 - 150)/ \sigma) = 1/5$, o $P(Z < (166 - 150)/ \sigma) = 0.80$

A les taules, el quantil 0.80 és aprox. 0.84: llavors, $(166 - 150)/ \sigma = 0.84 \rightarrow \underline{\sigma = 19.05}$ segons

2. El temps típic d'una cançó pop és de tres minuts i mig. Quina és la probabilitat que prenent una cançó punk a l'atzar tingui almenys una durada com aquesta? (si no heu trobat una solució a l'apartat 1., utilitzeu variància=400) *(1pt)*

3 minuts i mig = 210 segons. Trobarem $P(T > 210) = P(Z > (210 - 150)/ 19.05) = 1 - F(3.15)$

$= 1 - 0.9992 = \underline{0.0008}$

Amb la dada de la variància igual a 400 (desviació tipus 20), el resultat seria: $P(Z > (210 - 150)/ 20) = 1 - F(3) = 0.0013$

3. Hem agafat un disc d'un grup punk molt conegut, al qual comptem 12 cançons. Si admetem que aquestes cançons tenen durades independents i que segueixen el model anteriorment descrit, calculeu un interval per a la durada total del disc, per tal que puguem afirmar: "amb probabilitat 90%, el disc dura entre ... i ... minuts en total". *(1.5pt)*

La durada del disc és la suma dels temps de 12 cançons independents. Com que la suma de Normals també és Normal, hem de trobar primer el valor esperat i la variància de la suma:

D: "durada del disc de 12 cançons"; $E(D) = 12 \times 150 = 1800$ segons (30 minuts)

$V(D) = 12 \times 19.05^2 = 4355\ s^2$ ($\underline{\sigma_D = 65.99\ s}$)

Després, calcularem a i b tal que $P(a < D < b) = 0.90$. L'interval més senzill és un centrat al valor esperat. Llavors, busquem c tal que $P(1800-c < D < 1800+c) = 0.90$ o, equivalentment:

$P(D < 1800+c) = 0.95$. Si estandarditzem: $P(Z < (1800+c-1800)/65.99) = 0.95$, c/65.99 és igual al quantil 0.95 = 1.645 (taules) → $c = 1.645 \times 65.99 = 108.55 \rightarrow a = 1800-c = 1691.45$, $b = 1800+c = 1908.55$. (1691.45, 1908.55) seg., o <u>(28' 11.4", 31' 48.6")</u>

$V(D) = 12 \times 400 = 4800\ s^2$ ($\sigma_D = 69.28$ s); interval (28' 6", 31' 54")

4. Quin model de probabilitat és l'adequat per a la variable "nombre de cançons d'un disc punk que no arriben als dos minuts de durada"? Assumiu que els discos són de 12 títols. *(1pt)*

Nombre de cançons que no arriben als dos minuts (S), segueix el model <u>Binomial</u>, amb <u>n=12</u> i <u>p</u> = P(T<120):

$p = P(Z< (120-150)/19.05) = F(-1.575) = 1-F(1.575)$.

$F(1.57) = 0.9418$; $F(1.58) = 0.9429$; llavors $\underline{p} \approx 1-0.9423 = \underline{0.0577}$

$p = F(-1.5) = 0.0668$

5. Valor esperat i variància de la variable anterior. *(1pt)*

$E(S) = np = 12 \times 0.0577 = 0.6924$ *0.8017*

$V(S) = np(1-p) = 0.65245$ *0.7481*

6. Probabilitat que un disc punk tingui més de 2 cançons de menys de dos minuts de durada. *(1.5pt)*

$P(S > 2) = 1-P(S\leq 2) = 1-p_S(0) - p_S(1) - p_S(2) = 1-(1-p)^{12} - 12p(1-p)^{11} - 12\cdot 11/2\ p^2(1-p)^{10} =$

$= 1 - 0.4900837 - 0.3601125 - 0.1212795 = \underline{0.02852}$ *0.04159*

7. Hem posat el canal de spotify de música punk (una inacabable seqüència de temes independents). Volem trobar tres cançons de durada superior a 3 minuts. Descriviu el model de probabilitat per al nombre de cançons que hem d'escoltar per a seleccionar aquestes tres cançons. Quin serà el nombre mitjà de cançons a escoltar, i la desviació tipus? *(1.5pt)*

La variable en aquest cas (Q) segueix un model Binomial Negativa (r=3, p=P(T>180))

És a dir: p = 1 – P(Z<30/19.05) = 1 – F(1.575) = 0.0577 (n'hi ha tantes cançons que no arriben als 2 minuts com les que passen dels 3 minuts).

$\underline{E(Q) = r/p = 51.99}$ *44.91*

$\underline{\sigma_Q = \sqrt{(1-p)r/p^2} = 29.14}$ *25.05*

8. Justifica si es podria o no es podria emprar una distribució Normal amb el valor esperat i desviació tipus trobats a l'apartat anterior com una aproximació del model autèntic d'aquesta variable aleatòria. *(1pt)*

La Binomial Negativa es pot veure com un cas on es sumen r variables aleatòries independents geomètriques de paràmetre p. Si r és un nombre gran, el TCL es podria aplicar, però com que en aquest punt r val 3 (petit), i la distribució geomètrica és extremadament asimètrica, aquesta variable no es pot aproximar per una Normal.

Al costat podeu veure l'aspecte real de la distribució de la variable Q. És molt clar que és lluny de semblar una Normal.