Problema 1: Solució

Una empresa d'informàtica rep una peça clau dels seus ordinadors de tres proveïdors diferents: A, B i C. Malauradament cap dels tres proveïdors produeix un 100% de les peces sense defecte. Les respectives probabilitats es troben a la taula següent, on D indica que la peça té un defecte i $\neg D$ que no en té cap.

	A	**B**	**C**
D	0,03	0,05	0,02
$\neg D$	0,97	0,95	0,98

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

Solució:
Com la suma de cada columna fa 1, es tracta de les probabilitats condicionades de la variable 'Defecte' donat cada un dels tres proveïdors. En cas de probabilitats conjuntes la suma de totes les probabilitats a la taula seria 1.

(b) **(0.5 punts)**
Són independents les variables 'Proveïdor' i 'Defecte'? Raoneu la resposta.

Solució:
Les dues variables no són independents, ja que les distribucions condicionades a la taula són diferents. La probabilitat d'un defecte varia d'un proveïdor a un altre.

(c) **(1.5 punts)**
Se sap que la meitat de les peces provenen del proveïdor A, i que B proveeix el doble de peces que C. Per tant, quina és la probabilitat que una peça qualsevol tingui un defecte?

Solució:
Tenim $\mathrm{P(A)} = 0.5$, $\mathrm{P(B)} = \frac{1}{3}$ i $\mathrm{P(C)} = \frac{1}{6}$ i per tant

$$\mathrm{P}(D) = \mathrm{P}(D|\mathrm{A})\mathrm{P(A)} + \mathrm{P}(D|\mathrm{B})\mathrm{P(B)} + \mathrm{P}(D|\mathrm{C})\mathrm{P(C)} = 0.03 \cdot 0.5 + 0.05 \cdot \tfrac{1}{3} + 0.02 \cdot \tfrac{1}{6} = 0.035.$$

(d) **(0.5 punts)**
Quina és la probabilitat que una peça qualsevol tingui un defecte i sigui del proveïdor A?

Solució:

$$\mathrm{P}(\mathrm{A} \cap D) = \mathrm{P}(D|\mathrm{A})\mathrm{P(A)} = 0.03 \cdot \tfrac{1}{2} = 0.015.$$

(e) **(1.5 punts)**
Si trobem una peça amb defecte, quina és la probabilitat que sigui del proveïdor B?

Solució:

$$\mathrm{P}(\mathrm{B}|D) = \frac{\mathrm{P}(D|\mathrm{B})\mathrm{P(B)}}{\mathrm{P}(D)} = \frac{0.05 \cdot \frac{1}{3}}{0.035} = 0.476.$$

(f) **(1.75 punts)**
Donada una peça que no és del proveïdor B, quina és la probabilitat que tingui un defecte?

Solució:

$$\mathrm{P}(D|\neg\mathrm{B}) = \mathrm{P}(D|\mathrm{A} \cup \mathrm{C}) = \frac{\mathrm{P}\big(D \cap (\mathrm{A} \cup \mathrm{C})\big)}{\mathrm{P}(\mathrm{A} \cup \mathrm{C})} = \frac{\mathrm{P}(D|\mathrm{A})\mathrm{P(A)} + \mathrm{P}(D|\mathrm{C})\mathrm{P(C)}}{0.5 + \frac{1}{6}} = 0.0275.$$

(g) **(1.75 punt)**
Ens porten dues peces de forma independent. Quina és la probabilitat que (exactament) una tingui un defecte?

Solució:
Hi ha dues possibilitats: que la primera peça tingui un defecte i no la segona o a l'inrevés. Per tant la probabilitat és:

$$\mathrm{P}(\text{Una de dues peces té un defecte}) = 2 \cdot 0.035 \cdot (1 - 0.035) = 0.068.$$

(h) **(2 punts)**
Si sabem que (exactament) una de les dues peces té un defecte, quina és la probabilitat que les dues peces siguin del proveïdor C?

Solució:
Sigui $D1$ la variable que una de les dues peces té un defecte i C2 la variable que dues peces siguin del proveïdor C. Llavors,

$$\mathrm{P}(\mathrm{C2}|D1) = \frac{\mathrm{P}(D1|\mathrm{C2})\mathrm{P}(\mathrm{C2})}{\mathrm{P}(D1)} = \frac{2 \cdot 0.02 \cdot 0.98 \cdot \frac{1}{6} \cdot \frac{1}{6}}{0.068} = 0.016.$$

NOM:______________________________COGNOMS:____

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

Problema 2 (B2)

Al laboratori d'Arquitectura de Computadors s'estudia un processador que permet solapar instruccions. Inevitablement, alguns parells d'instruccions consecutives no es poden solapar, i es produeix un retard afegit. A continuació teniu la funció de distribució del nombre de cicles de retard per instrucció per aquesta causa, de la que sabem que el seu valor esperat és 2,55:

0	0,21
1	x
2	0,67
4	0,76
6	0,92
8	1,00

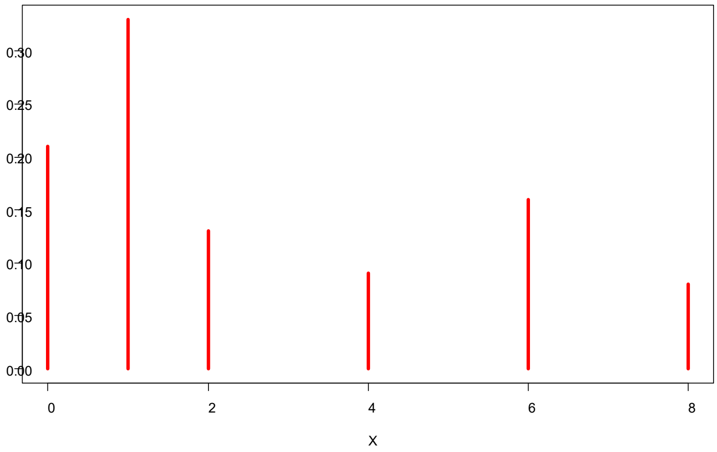
1. Trobeu el valor de la incògnita x, detallant les passes que feu. Dibuixeu també la funció de probabilitat d'aquesta variable aleatòria.

0	0,21
1	x-0,21
2	0,67-x
4	0,09
6	0,16
8	0,08

Obtenim els valors de la funció de probabilitat fent les diferències de valors consecutius (en funció de la incògnita quan s'escau), i imposem que el valor esperat sigui 2,55:

0 0,21 + 1(x-0,21) + 2(0,67-x) + 4 0,09 + 6 0,16 + 8 0,08 = 2,55

D'aquesta equació s'obté que x = 0,54



Considerem ara un cas particular de dues instruccions consecutives i els respectius retards. Sigui X1 la primera instrucció (files) i X2 la segona (columnes), i aquesta la seva funció de probabilitat conjunta:

X1 / X2	0	1	2	4	6
0				0,01	0,05
1		0,01	0,08	0,11	0,10
2	0,03	0,09	0,10	0,05	0,06
4	0,07	0,14	0,04	0,01	
6	0,05				

2. En primer lloc, es demana calcular la probabilitat que la primera instrucció es retardi més cicles que la segona.
   P(X1 > X2) = P(X1=2 & X2=0) + P(X1=2 & X2=1) + P(X1=4 & X2=0) + P(X1=4 & X2=1) + P(X1=4 & X2=2) + P(X1=6 & X2=0) = 0,42

3. Si es sap que entre ambdues instruccions s'ha comptat un total de 6 cicles de retard, quant val ara la probabilitat del mateix esdeveniment?
   P(X1 > X2 | X1+X2=6) = P(X1 > X2 & X1+X2=6) / P(X1+X2=6) = (P(X1=4 & X2=2) + P(X1=6 & X2=0)) / (P(X1=4 & X2=2) + P(X1=6 & X2=0) + P(X1=2 & X2=4) + P(X1=0 & X2=6)) = 0.09 / 0.19 = 0.474

4. Es vol calcular la correlació entre X1 i X2. Coneixem els respectius valors esperats (2,3 i 2,66) i les respectives variàncies (4,4844 i 2,29).
   a. Primer, trobeu la distribució de probabilitat per a la suma de cicles de retard per a les dues instruccions (X1+X2).
   b. Després, trobeu l'esperança i variància de la suma anterior.
   c. Finalment digueu quant valen la covariància i la correlació.
   d. Interpreteu el valor que heu trobat.

X1 / X2	0	1	2	4	6	
0				0,01	0,05	0,06
1		0,01	0,08	0,11	0,10	0,3
2	0,03	0,09	0,10	0,05	0,06	0,33
4	0,07	0,14	0,04	0,01		0,26
6	0,05					0,05
	0,15	0,24	0,22	0,18	0,21	1

2	0,04
3	0,17
4	0,18
5	0,25
6	0,19
7	0,10
8	0,07

Els colors indiquen els diferents valors que pren X1+X2. A la dreta, la funció de probabilitat de la suma de cicles.

Tenim els valors esperats de les variables: E(X1) = 2,3, E(X2) = 2,66, i les variàncies: V(X1) = 4,4844, V(X2) = 2,29. A partir d'aquí sabem que E(X1+X2) = 2,3+2,66 = 4,96. En canvi, la variància la calculem des de la funció de probabilitat que hem obtingut:

$E((X1+X2)^2) = 2^2\ 0,04 + 3^2\ 0,17 + \ldots\ 8^2\ 0,07 = 27,04$; $V(X1+X2) = 27,04 - 4,96^2 = 2,4384$

Com que V(X1+X2) = V(X1) + V(X2) + 2 cov(X1,X2),   cov(X1, X2) = (2,4384 – 4,4844 – 2,29)/2 = -2,168

I cor(X1, X2) = cov(X1, X2)/($\sigma_1\ \sigma_2$) = -2,168/√(4,4844 2,29) = -0,6765

El valor negatiu ens diu que si la primera instrucció es retarda més, la segona no es retardarà tant, i viceversa. L'associació és prou forta, donat el valor absolut de la correlació (~2/3).

L'estudi inclou la mesura del temps real que la instrucció triga en executar-se (anomenarem T a aquesta V.A.). S'ha trobat que la següent funció serveix molt bé com a funció de distribució de T: $F_T(t) = a + b\ \cos{t/3}, 0 < t < 3\pi$, on 3π = 9.42... és el màxim temps que es podria observar.

5. Determineu els valors adients per als paràmetres a i b.

$F_T(0) = a + b\ \cos{0/3} = a + b = 0$, la funció de distribució val 0 a l'extrem inferior
$F_T(3\pi) = a + b\ \cos{3\pi/3} = a - b = 1$, la funció de distribució val 1 a l'extrem superior

D'aquestes dues equacions trobem que a = 1/2 i b = -1/2: $F_T(t) = \frac{1}{2} - \frac{1}{2}\cos{t/3}, 0 < t < 3\pi$

6. Digueu també quina de les següents opcions correspon a la variància de la variable aleatòria T i perquè:

a) $\frac{1}{6}\int_0^{3\pi} x^2 \cos{x/3}\,dx - \mu_T^2$	b) $\frac{1}{6}\int_0^{3\pi} x \cos{x/3}\,dx - \mu_T^2$	c) $\frac{1}{6}\int_0^{3\pi} x^2 \cos{x^2/3}\,dx - \mu_T^2$
d) $\frac{1}{6}\int_0^{3\pi} x^2 \sin{x/3}\,dx - \mu_T^2$	e) $\frac{1}{6}\int_0^{3\pi} x \sin{x/3}\,dx - \mu_T^2$	f) $\frac{1}{6}\int_0^{3\pi} x^2 \sin{x^2/3}\,dx - \mu_T^2$

Emprarem l'expressió V(T) = E(T^2) - μ_T^2. Clarament, només hem d'identificar quina de les integrals correspon a E(T^2). La derivada de $F_T(t)$ és 1/3 1/2 sin $t/3$, i E(T^2)= $\int\ x^2 f_T(x)dx$. Per tant, l'expressió correcta és la d).

NOM:______________________________COGNOMS:____

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

Problema 3 (B3)

En les darreres setmanes, els servidors de la Universitat Autònoma de Barcelona (UAB) han sofert un ciberatac de tipus *ransomware*. En aquest problema oferim una sèrie de qüestions relacionades amb aquest fet que no es basen en dades reals.

1. Hi ha 8 servidors a la UAB que contenen dades personals dels diferents col·lectius que conformen la universitat. La probabilitat de que els atacants accedeixen a les dades personals d'un servidor en un intent concret és 0.2.

   A) Digues quin és el nombre esperat de servidors que seran accedits. **(1 punt)**

   X="Número de servidors que accedeixen" ~ B(n=8, p=0.2)

   $$E(X) = n \cdot p = 8 \cdot 0.2 = \mathbf{1.6}$$

   B) Quina és la probabilitat que els atacants accedeixin a més de 2 servidors? **(1 punt)**

   $$P(X > 2) = 1 - P(X \leq 2) = (Taules) = 1 - 0.7969 = \mathbf{0.2031}$$

   C) Els atacants adopten una estratègia més agressiva: si fallen en un intent sobre un servidor, l'algoritme s'actualitza i es torna a provar en un nou intent. Assumint que la probabilitat d'accedir al servidor és manté constant en qualsevol intent, i que l'èxit per als ciberatacants és independent en cada intent, quina és la probabilitat de que aconsegueixen les dades dels 8 servidors en exactament 20 intents? **(1 punt)**

   Y="Nombre d'intents necessaris per assolir 8 èxits" ~ BN(r=8, p=0.2)

   $$P(X = 20) = \binom{19}{7} \cdot 0.2^8 \cdot 0.8^{12} = 50388 \cdot 2.56 \cdot 10^{-8} \cdot 0.06871948 = \mathbf{0.00886}$$

2. Una de les accions més emprades pels hackers és enviar correus *spam* a les víctimes potencials amb un enllaç maliciós. Aquesta acció la realitzen durant una hora. Durant aquest temps, 400 correus són enviats a professors, i 1600 a alumnes. La probabilitat que un professor cliqui en un d'aquests links és 0.005, mentre que la probabilitat que cliqui un alumne és 0.01.

   A) Quina és la probabilitat de que exactament 2 professors cliquin en un enllaç maliciós? **(1.5 punts)**

   N_p: "Nombre de professors que clicaran en un enllaç maliciós en 1 hora" ~ B(n=400, p=0.005)~P(λ=2)

   $$Solució\ per\ Binomial \rightarrow\ P\left(N_p = 2\right) = \binom{400}{2} \cdot 0.005^2 \cdot 0.995^{398} = \mathbf{0.27}$$

   $$Solució\ per\ Poisson\ \rightarrow\ P\left(N_p = 2\right) = e^{-2} \cdot \frac{2^2}{2!} = \mathbf{0.27}$$

   B) Quin és el nombre màxim de persones (professors o alumnes) que clicaran en un enllaç maliciós amb una probabilitat com a mínim de 0.90? **(1 punt)**

   N_a: "Nombre d'alumnes que clicaran en un enllaç maliciós en 1 hora" ~ B(n=1600, p=0.01) ~ P(λ=1600·0.01=16)

   N: "Nombre de persones que clicaran en un enllaç maliciós en 1 hora"

   $N = N_p + N_a \sim P(\lambda=18)$

   $$P(N \leq q_{0.90}) = 0.9 \rightarrow (Taules) \rightarrow q_{0.90} = \mathbf{24}$$

   C) Un professor despistat acaba de clicar en un enllaç maliciós. Quin és el temps esperat fins que un altre professor cometi el mateix error? **(1 punt)**

   T="Temps esperat entre dos *clicks* en enllaços maliciosos de professors"~Exp(λ=2)

   E(T)=1/2 hora=**30 minuts**

3. A banda dels 8 servidors amb dades personals mencionats prèviament, hi ha un total de 200 servidors repartits pels diferents departaments de la UAB amb informació de docència i recerca que es fa a la universitat. S'estima que després de l'atac la informació perduda en cadascun d'aquests servidors es distribueix segons una distribució que té com a valor esperat 1 GB i desviació estàndard 2 GB. *Nota: podeu pensar que es tracta d'una distribució amb una llarga cua per la dreta.* **(1 punt)**

   A) Quina distribució segueix la quantitat d'informació total perduda (en GB) en aquests 200 servidors? **(1 punt)**

   Pel TCL:

   S="GB totals perduts en els 200 servidors"$\sim \boldsymbol{N(\mu = 200\ GB, \sigma = \sqrt{200 \cdot 2^2} = 28.28\ GB)}$

   B) Quina és la probabilitat de que la informació perduda sigui superior a 220 GB? **(1 punt)**

   Z~N(0,1)

$$P(S > 220) = P\left(\frac{S-\mu}{\sigma} > \frac{220-200}{28.28}\right) = P(Z > 0.707) = 1 - P(Z < 0.707) = (Taules) \approx 1 - 0.7611 = \mathbf{0.2389}$$

   C) Es vol afitar inferiorment la quantitat d'informació total perduda, amb un marge d'error del 1% (és a dir: amb probabilitat 99%, la informació perduda és superior a ... GB). Troba el valor d'aquesta fita. **(1.5 punts)**

$$P(S > a) = 0.99 \rightarrow P\left(\frac{S-\mu}{\sigma} > \frac{a-200}{28.28}\right) = 0.99 \rightarrow P\left(Z > \frac{a-200}{28.28}\right) = 0.99 \rightarrow \frac{a-200}{28.28} = z_{0.01} \xrightarrow{Taules} a = 200 + 28.28 \cdot (-2.33) = \mathbf{134.2}$$