Examen Parcial 1

Bloques 1, 2 y 3

Profesors Dept. EIO

10 de Noviembre de 2020

B1

PROBLEMA 1

1. A una universitat específica, els professors poden ser associats, titulars i catedràtics, i el rector presumeix d'haver assolit la igualtat de gènere perquè aquest és independent de la categoria. Concretament, un 30% són dones associades, un 18% dones titulars i un 9% dones catedràtiques. Deduïu els percentatges corresponents pels professors homes, justificant la resposta. ***(2 punts)***

Puntuació: 2 punts

$$P(C_i \cap D) = (0.3,\, 0.18,\, 0.09)$$

$$P(C_i|D) = \frac{P(C_i \cap D)}{P(D)} = (\frac{0.3}{0.57},\, \frac{0.18}{0.57},\, \frac{0.09}{0.57}) = (0.526,\, 0.316,\, 0.158) = P(C_i|H) = P(C_i)$$

$$P(C_i \cap H) = P(C_i) \cdot P(H) = (0.526,\, 0.316,\, 0.158) \cdot 0.43 = (0.226,\, 0.136,\, 0.068)$$

```
[1] 0.22631579 0.13578947 0.06789474
```

PROBLEMA 2

2. En el sistema de streaming de determinada escola, quan el vídeo funciona bé, l'àudio s'escolta bé amb probabilitat 0.96; i quan l'àudio s'escolta bé, el vídeo funciona bé el 83% de les vegades. Es sap que en el 77% de les sessions el so i la imatge són satisfactoris i permeten fer classe sense problemes. Amb aquesta informació, deduïu la probabilitat que en una sessió no funcioni cap dels dos components. ***(3 punts)***

Puntuació: 3 punts

$$P(A \cap V) = 0.77$$

$$P(A|V) = 0.96$$

$$P(V|A) = 0.83$$

$$P(A) = \frac{P(A \cap V)}{P(V|A)} = \frac{0.77}{0.83} = 0.9277$$

$$P(V) = \frac{P(A \cap V)}{P(A|V)} = \frac{0.77}{0.96} = 0.8021$$

$$P(\bar{A} \cap \bar{V}) = 1 - P(A \cup V) = 1 - (P(A) + P(V) - P(A \cap V)) = 1 - (0.9277 + 0.8021 - 0.77) = 0.0402$$

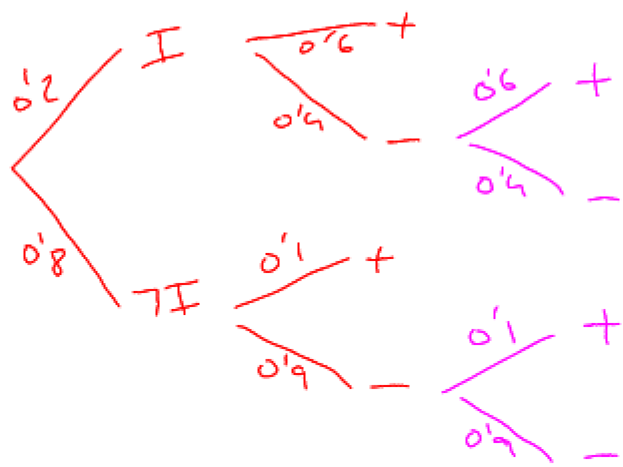
```
[1] 0.04020582
```

PROBLEMA 3

3. A un centre d'atenció primària s'empra una prova diagnòstica ràpida per a certa malaltia vírica que dóna positiu a la prova el 60% de les vegades per a infectats, i el 10% per a no infectats. S'està considerant modificar el protocol, repetint la prova als pacients que hagin donat negatiu la primera vegada; llavors, serà positiu el pacient que ha donat positiu a la primera o a la segona prova. Suposant que hi ha un 20% de pacients realment infectats entre els que visiten el centre, determineu: ***(5 punts)***

a) la proporció de positius que es detecten mitjançant el protocol original, i el mateix amb la modificació

Puntuació: 2 punts



$$P(+_1) = P(+|I) \cdot P(I) + P(+|\bar{I}) \cdot P(\bar{I}) = 0.6 \cdot 0.2 + 0.1 \cdot 0.8 = 0.2$$
$$P(+_{Modificat}) = P(+_1) + P(+_2|-_1) \cdot P(-_1) = 0.2 + P(+_2|I \cap -_1) \cdot P(I \cap -_1) + P(+_2|\bar{I} \cap -_1) \cdot P(\bar{I} \cap -_1) =$$
$$0.2 + 0.2 \cdot 0.4 \cdot 0.6 + 0.8 \cdot 0.9 \cdot 0.1 = 0.2 + 0.048 + 0.072 = 0.32$$

b) la probabilitat que un pacient estigui infectat si ha donat positiu amb el primer protocol

Puntuació: 1 punt

$$P(I|+_1) = \frac{P(I \cap +_1)}{P(+_1)} = \frac{0.12}{0.2} = 0.6$$

c) ídem amb el protocol modificat

Puntuació: 1 punt

$$P(I|+_{modificat}) = \frac{P(I \cap +_{modificat})}{P(+_{modificat})} = \frac{0.12+0.048}{0.32} = 0.525$$

d) basant-se en els resultats previs, digueu si seria beneficiós aplicar la modificació del protocol.

Puntuació: 1 punt

La probabilitat de que siguin infectats realment es major amb el primer protocol per tant es millor no utilitzar la modificació.

B2

PROBLEMA 1

1. A un sistema informàtic, una variable X relacionada amb la mida dels documents pren valors $x = 10^k$, on $k = 0, 1, \ldots, 6$ (és a dir: 1, 10, 100, ...). La probabilitat que X valgui x és el doble de la probabilitat per a un valor deu vegades més gran (és a dir: $p_X(x) = 2p_X(10x)$), en tots els casos. Determineu : *(2 punts)*

a) la funció de probabilitat de X i l'esperança de la variable.

Puntuació:

- funció de probabilitat: 1 punt

- Esperança: 0.5 punts

x	1	10	100	1000	10000	100000	1000000
p(x)	64/127	32/127	16/127	8/127	4/127	2/127	1/127

$$E(X) = \sum x \cdot p(x) = 1 \cdot \frac{64}{127} + 10 \cdot \frac{32}{127} + 100 \cdot \frac{16}{127} + 1000 \cdot \frac{8}{127} + 10000 \cdot \frac{4}{127} + 100000 \cdot \frac{2}{127} + 1000000 \cdot \frac{1}{127} = 9842.394$$

```
Las probabilidades son:
[1] 64/127 32/127 16/127  8/127  4/127  2/127  1/127

E(X): 9842.394
```

b) Digueu també què val la mediana (quantil 0.5, o percentil 50%), justificant la resposta.

Puntuació: 0.5 punts

(Identificando el valor ya es suficiente sin hacer la tabla)

Per a trobar el quantil .5 haurem de buscar el acumulat i hallar la probabilitat acumulada del 0.5. Per fer-ho anem a realitzar la taula:

x	1	10	100	1000	10000	100000	1000000
p(x)	64/127	96/127	112/127	120/127	124/127	126/127	127/127

Com podem veure, aquest valor correspon quan $X = 1$.

```
M(X): 1
```

PROBLEMA 2

2. Sigui T una variable aleatòria contínua que pren valors en l'interval $[0, 9]$. Un investigador ha presentat la següent funció com la seva funció de distribució: ***(1 punt)***

$$F(t) = \frac{t^2-5t}{(3-t)^2}$$

Discutiu si la proposta és justificable o no, raonadament.

Puntuació: 1 punt

$F(0) = 0$, $F(9) = 1$, pero es negativa desde 0 fins a 5 ja que $t^2 < 5t$. Aixi que no pot ser una funció de distribució.

PROBLEMA 3

3. Considerem una variable W que només pren 3 valors, -10, 0 i 10. Sabem que la desviació tipus val 9 i que el 50% de les vegades la variable és negativa. Trobar la funció de probabilitat de W i el seu valor esperat. ***(2 punts)***

Puntuació: 2 punts

-10	0	10
0.5	x = 0.163325	$\frac{1}{2}$-x = 0.336675

$$E(W^2) = 100 \cdot 0.5 + 0 \cdot x + 100 \cdot (\frac{1}{2} - x) = 100 - 100x$$

$$E(W) = -10 \cdot 0.5 + 0 \cdot x + 10(0.5 - x) = -10x \to V(W) = 100 - 100x - (-10x)^2 = 81 \to x = 0.163325, E(W) = -1.63325$$

PROBLEMA 4

4. Una empresa fabrica panells solars rectangulars, amb diferent mides d'ample i de llarg, segons el nombre de mòduls (cèl•lules fotovoltaiques) que s'hi instal•len. El fabricant utilitza 2, 3 o 4 mòduls d'ample, i 5, 10 o 15 mòduls de llarg, però no en qualsevol combinació, ni totes les combinacions emprades es venen igual. A la taula següent es mostra, per cada 100 vendes, com es distribueixen segons les dimensions del panell: ***(5 punts)***

	2	3	4
5	25	20	16
10	0	10	21
15	0	0	8

a) Quin és el ample esperat?

Puntuació: 1 punt

$$E(A) = 2 \cdot 0.25 + 3 \cdot 0.3 + 4 \cdot 0.45 = 0.5 + 0.9 + 1.8 = 3.2$$

```
[1] 3.2
```

b) Quin és el llarg esperat?

Puntuació: 1 punt

$$E(L) = 5 \cdot 0.61 + 10 \cdot 0.31 + 15 \cdot 0.08 = 3.05 + 3.1 + 1.2 = 7.35$$

```
[1] 7.35
```

c) Quin és el nombre esperat de mòduls a un panell?

Puntuació: 1 punt

$$E(M) = E(A{\cdot}L) = 5{\cdot}2{\cdot}0.25{+}5{\cdot}3{\cdot}0.2{+}5{\cdot}4{\cdot}0.16{+}10{\cdot}3{\cdot}0.1{+}10{\cdot}4{\cdot}0.21{+}15{\cdot}4{\cdot}0.08 = 2.5{+}3{+}3.2{+}3{+}8.4{+}4.8 = 24.9$$

```
[1] 24.9
```

d) I Quina és la covariància entre ample i llarg?

Puntuació: 1 punt

$$cov(A, L) = E(A \cdot L) - E(A) \cdot E(L) = E(M) - E(A) \cdot E(L) = 24.9 - 3.2 \cdot 7.35 = 1.38$$

```
[1] 1.38
```

e) El signe obtingut és coherent amb el que s'aprecia a les dades? Explica el perquè de la teva resposta

Puntuació: 1 punt

Si es coherent perquè a la taula es veu com més gran és A, major és L.

B3

PROBLEMA 1

1. El quocient intel•lectual (IQ) per a nens de 12 anys s'obté mitjançant certes proves psicològiques, ajustades per a proporcionar un model Gaussià (Normal) amb esperança 100 i desviació tipus 10. Suposem que a una escola es considera que un alumne és d'altes capacitats (AC) si té un IQ superior a 120. Trobeu quina seria la proporció d'alumnes amb AC si l'IQ dels alumnes de 12 anys d'aquesta escola es distribueix segons aquest model. ***(2 punts)***

Puntuació: 2 punts

$$AC = IQ > 120 \; IQ \sim Normal(\mu = 100, \sigma = 10)$$

$$P(IQ > 120) = P(Z > \tfrac{120-100}{10}) = P(Z > 2) = 1 - F_{IQ}(2) = 1 - P(Z \leq 2) = 1 - 0.9772 = 0.0228$$

```
{pnorm(120, mean = 100, sd = 10, lower.tail = F)
## Múltiples opciones
# pnorm(2, mean = 0, sd = 1, lower.tail = F)
# 1 - pnorm(120, mean = 100, sd = 10, lower.tail = T)
# 1 - pnorm(2, mean = 0, sd = 1, lower.tail = F)
}
```

```
[1] 0.02275013
```

PROBLEMA 2

2. Una classe té 25 alumnes de 12 anys d'edat, independents entre ells. Digueu quin és el model per a descriure el nombre d'alumnes amb AC, i calculeu manualment les probabilitats dels primers valors d'aquesta variable (és suficient amb arribar a una probabilitat acumulada que sobrepassi el 97.5%). Interpreteu el que heu trobat. ***(2 punts)***

a) Digueu quin és el model per a descriure el nombre d'alumnes amb AC.

Puntuació: 0.5 punts

$$X \sim Binomial(n = 25, \; p = 0.0228)$$

b) calculeu manualment les probabilitats dels primers valors d'aquesta variable (és suficient amb arribar a una probabilitat acumulada que sobrepassi el 97.5%). Interpreteu el que heu trobat.

Puntuació: 1.5 punts

$$P(X=0) = \binom{25}{0} \cdot 0.0228^0 \cdot 0.9772^{25} = \frac{25!}{0!(25-0)!} \cdot 0.0228^0 \cdot 0.9772^{25} = 0.5618$$

$$P(X=1) = \binom{25}{1} \cdot 0.0228^1 \cdot 0.9772^{24} = \frac{25!}{1!(25-1)!} \cdot 0.0228^1 \cdot 0.9772^{24} = 0.3277$$

$$P(X=2) = \binom{25}{2} \cdot 0.0228^2 \cdot 0.9772^{23} = \frac{25!}{2!(25-2)!} \cdot 0.0228^2 \cdot 0.9772^{23} = 0.0917$$

x	0	1	2
$P(X = x)$	0.5618	0.3277	0.0917
$P(X \leq x)$	0.5618	0.8895	0.9812

Amb una probabilitat del 98%, en una aula de 25 alumnes no hi ha més de 2 alumnes amb AC.

```
# Calculamos los valores de la función de probabilidad
{d0 <- dbinom(x = 0, size = 25, p = 0.0228);
d1 <- dbinom(x = 1, size = 25, p = 0.0228);
d2 <- dbinom(x = 2, size = 25, p = 0.0228);
cat("La funció de densitat és:\n\t p(X = 0) =", d0,
    "\n\t p(X = 1) =", d1, "\n\t p(X = 2) =", d2)
# Calculamos los valores acumulados de la función de distribución
p0 <- pbinom(q = 0, size = 25, p = 0.0228);
p1 <- pbinom(q = 1, size = 25, p = 0.0228);
p2 <- pbinom(q = 2, size = 25, p = 0.0228);
cat("\n\nLa funció de distribució és:\n\t p(X <= 0) =", p0,
    "\n\t p(X <= 1) =", p1, "\n\t p(X <= 2) =", p2)
# Calculamos el quantil 97.5%
q <- qbinom(p = 0.975, size = 25, prob = 0.0228)
cat("\n\nEl valor de x que conté el 97.5% és de", q)}
```

```
La funció de densitat és:
     p(X = 0) = 0.5618061
     p(X = 1) = 0.3277011
     p(X = 2) = 0.09175093

La funció de distribució és:
     p(X <= 0) = 0.5618061
     p(X <= 1) = 0.8895072
     p(X <= 2) = 0.9812581

El valor de x que conté el 97.5% és de 2
```

PROBLEMA 3

3. Una determinada versió del sistema operatiu XYZ llença avisos als seus usuaris quan es produeix una *"minor update"*, que no requereix un *restart* immediat, encara que el sistema força el *restart* si amb el temps han arribat 10 *"minor updates"*. Els avisos són independents entre sí, i es produeixen en mitjana un cada 45 dies. *(4 punts)*

a) Quin és el model de probabilitat de la variable "Nombre de *"minor updates"* en un període de 6 mesos"?

Puntuació: 1 punt

U: nombre de minor updates en sis mesos

$$U \sim Poisson(\lambda = 6/1.5)$$

b) Calculeu la probabilitat que un usuari que ignora els avisos es trobi amb un *restart* forçat abans de 6 mesos des de l'últim (assumiu que tots els mesos són de 30 dies).

Puntuació: 1.5 punts

U: *Nombre de minor updates en 6 mesos* $\sim$ *Poisson*$(\lambda = 4)$

$$P(U \geq 10) \text{ Si hi ha 10 o més en 6 mesos es que hi haurà un restart forçat.}$$

$$P(U \geq 10) = 1 - P(U < 10) = 1 - F_U(9) = 1 - (\tfrac{0.4^0 \cdot e^{-0.4}}{0!} + \tfrac{0.4^1 \cdot e^{-0.4}}{1!} + \tfrac{0.4^2 \cdot e^{-0.4}}{2!} + \tfrac{0.4^3 \cdot e^{-0.4}}{3!} + \tfrac{0.4^4 \cdot e^{-0.4}}{4!} +$$

$$\tfrac{0.4^5 \cdot e^{-0.4}}{5!} + \tfrac{0.4^6 \cdot e^{-0.4}}{6!} + \tfrac{0.4^7 \cdot e^{-0.4}}{7!} + \tfrac{0.4^8 \cdot e^{-0.4}}{8!} + \tfrac{0.4^9 \cdot e^{-0.4}}{9!}) = 1 - 0.992 = 0.008$$

Els valors de les probabilitats s'han buscat en la taula de la Poisson.

```r
ppois(9, lambda = 4, lower.tail = FALSE)
```

```
[1] 0.008132243
```

c) Podeu averiguar quin és el temps màxim (en mesos, amb un error de l'1%) que pot passar sense rebre cap avís de *"minor updates"* ?

Puntuació: 1.5 punts

T: *temps fins al proper avís en mesos*

$$T \sim Exponencial(\lambda = 1/1.5) \; P(T < t) = 0.99 \; P(T < t) = 1 - e^{-\lambda \cdot t} = 1 - e^{\frac{-1}{1.5} \cdot t} = 1 - e^{\frac{-t}{1.5}} = 0.99$$

$$e^{\frac{-t}{1.5}} = 1 - 0.99 \; \frac{-t}{1.5} = log(1 - 0.99) \; \frac{t}{1.5} = -log(1 - 0.99) \; t = -log(1 - 0.99) \cdot 1.5 = 6.90776$$

```r
p = 0.99
-log(1-p)*1.5
```

```
[1] 6.907755
```

PROBLEMA 4

4. Reprenem la variable X del primer exercici de B2, la mida d'uns documents d'aquest SO. Admetem que el valor esperat és 10000 aprox., i la desviació tipus 90000 aprox. Es vol determinar la mida total màxima, amb un marge d'error de l'1%, per a un conjunt de n documents seleccionats a l'atzar. *(2 punts)*

a) Resoleu per al cas concret de 500 documents.

Puntuació: 1 punt

La mida total de $n = 500$ documents és $MT \sim Normal(\mu = 10000 \cdot n, \sqrt{n} \cdot 90000) = Normal(\mu = 5e^6, \sigma = 2012461)$

Fita (1%) $= \mu + Z_{0.99} \cdot 2012461 = 5e^6 + 2.33 \cdot 2012461 = 9689035$

```
n = 500
Ex = 10000
sd = 90000
mu = Ex*n
sigma = sqrt(n)*sd

{cat("Valor teòric:", mu + qnorm(0.99) * sigma)
cat("\nValor aproximat:", mu + 2.33 * sigma)}
```

```
Valor teòric: 9681685
Valor aproximat: 9689035
```

b) Discutiu breument la problemàtica per a aquesta situació

Puntuació: 1 punt

La distribució és extremadament asimètrica, però per TCL i amb una n suficientment gran podriem utilitzar la convergència a la Normal. Aquesta n hauria de ser molt gran, de l'ordre de les centenes o milers.