NOM:_____________________________COGNOM:_______________________________________________________

*Contesteu cada pregunta en el seu lloc. Expliciteu i justifiqueu els càlculs. Cada pregunta val 1 punt.*

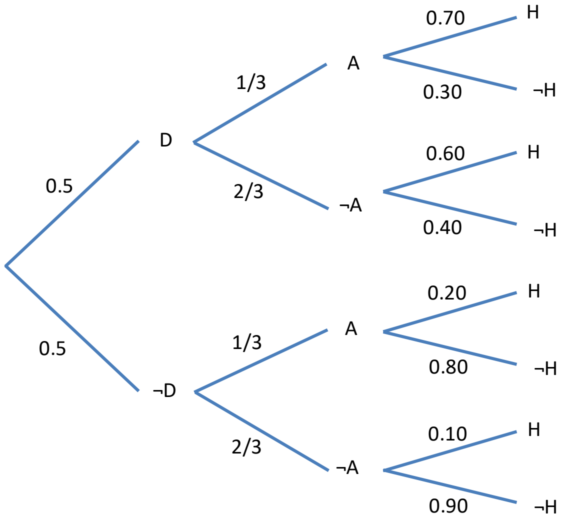
# Problema 1 (A)

El treball d'en David consisteix en estudiar com un agent, potser humà (diguem-ne $A$), potser una intel·ligència artificial ($\neg A$), respon quan se li demana que interpreti si un text determinat ha estat escrit per una persona ($D$) o per un programa informàtic ($\neg D$). La informació que David ha recopilat diu: **1**) un humà identifica correctament un text escrit per humans el 70% de les vegades; **2**) un humà identifica correctament un text escrit per una IA 4 de cada 5 vegades; **3**) un agent no humà identifica correctament un text escrit per humans el 60% de les vegades; **4**) un agent no humà identifica correctament un text escrit per una IA 9 de cada 10 vegades. Anomenarem $H$ al fet que un agent interpreti que un text determinat ha estat escrit per un humà.

Es posa a disposició dels usuaris una finestreta virtual on darrera tenim agents humans i no humans (un humà per cada dos IAs, en principi); els usuaris són els que proporcionen els textos per a ser analitzats. Suposarem que la meitat dels textos han estat escrits per humans.

1. Representeu l'arbre d'esdeveniments i probabilitats de la situació descrita anteriorment; utilitzeu la notació ja definida.

També seria correcte si el primer nivel és A / ¬A i el segon D / ¬D, ja que són independents (les probabilitats no canvien)




2. Poseu a una taula la distribució que teòricament podem esperar per a un conjunt de 90 documents escrits per persones si en David simula el funcionament de la finestreta virtual (justifiqueu rigorosament els càlculs):

| | resposta correcta (text escrit per humà) | resposta incorrecta (text escrit per IA) |
|---|---|---|
| agent humà | $90*P(H\cap A \mid D) = 90*(1/3*0.70) = \mathbf{21}$ | $90*P(\neg H\cap A \mid D) = 90*(1/3*0.30) = \mathbf{9}$ |
| agent IA | $90*P(H\cap\neg A \mid D) = 90*(2/3*0.60) = \mathbf{36}$ | $90*P(\neg H\cap\neg A \mid D) = 90*(2/3*0.40) = \mathbf{24}$ |

3. Suposem ara que s'ignora l'origen del document. Trobeu les probabilitats: **a**) que l'agent sigui humà i respongui que el text és d'origen humà; **b**) que l'agent no sigui humà i respongui que el text és d'origen artificial.

a) $P(A \cap H) = P(A \cap H \mid D)P(D) + P(A \cap H \mid \neg D)P(\neg D) = 1/3*0.70*0.5 + 1/3*0.20*0.5 = \mathbf{3/20}$

b) $P(\neg A \cap \neg H) = P(\neg A \cap \neg H \mid D)P(D) + P(\neg A \cap \neg H \mid \neg D)P(\neg D) = 2/3*0.40*0.5 + 2/3*0.90*0.5 = \mathbf{13/30}$

4. Quina és la probabilitat que el sistema de la finestreta virtual encerti l'origen del text (ho identifiqui correctament)?

Anomenarem OK a l'esdeveniment "el sistema encerta l'origen del text": $OK = D\cap H \cup \neg D\cap\neg H$

$P(OK) = P(D\cap H\cap A) + P(D\cap H\cap\neg A) + P(\neg D\cap\neg H\cap A) + P(\neg D\cap\neg H\cap\neg A) =$

$0.5*1/3*0.70 + 0.5*2/3*0.60 + 0.5*1/3*0.80 + 0.5*2/3*0.90 = \mathbf{0.75}$

5. Si davant el text proporcionat per un usuari a l'atzar el sistema respon que és d'origen humà, quina és la probabilitat que hagi respost una IA?

$P(\neg A \mid H) = P(H \mid \neg A)P(\neg A) / P(H)$

$P(H \mid \neg A) = P(H \cap D \mid \neg A) + P(H \cap \neg D \mid \neg A) = P(H \mid \neg A \cap D)P(D) + P(H \mid \neg A \cap \neg D)P(\neg D) =$

$0.60*0.50 + 0.10*0.50 = 0.35$

Similarment: $P(H \mid A) = 0.70*0.50 + 0.20*0.50 = 0.45$

$P(\neg A) = 2/3$; $P(A) = 1/3$ (és independent de D)

$P(H) = P(H \mid A)P(A) + P(H \mid \neg A)P(\neg A) = 0.45*1/3 + 0.35*2/3 = 23/60$

Llavors, $P(\neg A \mid H) = 0.35*2/3 / (23/60) = \mathbf{0.609}$

En David per la nit és un programador de jocs, i està configurant un personatge amb dos atributs: *potència* de l'atac (horitzontal) i *resistència* de l'atac (vertical). Els dos atributs es mesuren amb un número enter entre 1 i 5, i la resistència esperada hauria de ser 2.8. Es vol que els atributs es comportin d'una determinada manera que s'implementa amb una taula de freqüències (a la dreta). Però dues freqüències s'han perdut.

| resistència \ potència | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | | | 1/40 | 1/40 | 3/40 |
| 2 | | 1/20 | 3/40 | * | 1/20 |
| 3 | 1/40 | 1/10 | # | 1/40 | 1/40 |
| 4 | 1/20 | 3/40 | 1/20 | | |
| 5 | 1/20 | 1/40 | | | |

6. Reconstruïu la taula de freqüències, justificant els passos.

Tenim (1) tota la taula ha de sumar 1, i (2) el valor esperat de R és 2.8. Anomenarem $x$ a P(3,3) [#] i $y$ a P(2,4) [*]

$$29/40 + x + y = 1$$

$$5/40*(1) + (7/40 + y)*(2) + (7/40 + x)*(3) + 7/40*(4) + 3/40*(5) = 2.8$$

Si resolem el sistema, s'obté:

$$x = 7/40$$

$$y = 1/10$$

7. Quant val la potència esperada? Trobeu també la desviació tipus de la resistència i interpreteu.

E(P) = 5/40*(1) + 10/40*(2) + 13/40*(3) + 6/40*(4) + 6/40*(5) = **2.95**
$E(R^2)$ = 5/40*(1) + 11/40*(4) + 14/40*(9) + 7/40*(16) + 3/40*(25) = 9.05
V(R) = 9.05 – 2.8*2.8 = 1.21
$\sigma_R$ = **1.1**

Indica que la resistència de l'atac pot variar de la mitjana típicament al voltant de 1.1 punts

8. Cada atac representa un consum de l'energia del personatge proporcional a la seva potència i a la seva resistència. A efectes pràctics, considerem que el consum es mesura pel producte d'ambdós atributs. Trobeu la distribució de probabilitat de l'energia consumida en un atac.

| | |
|---|---|
| 3 | 2/40 |
| 4 | 5/40 |
| 5 | 5/40 |
| 6 | 7/40 |
| 8 | 7/40 |
| 9 | 7/40 |
| 10 | 3/40 |
| 12 | 3/40 |
| 15 | 1/40 |

A l'esquerra, tots els valors que podem trobar fent el producte de potencia i resistència, amb probabilitat major que 0.

A la dreta, funció de probabilitat, sumant les probabilitats que corresponen al mateix valor de consum.

Per exemple, P(E=8) = P(2,4) + P(4,2) = 1/10 + 3/40 = 7/40

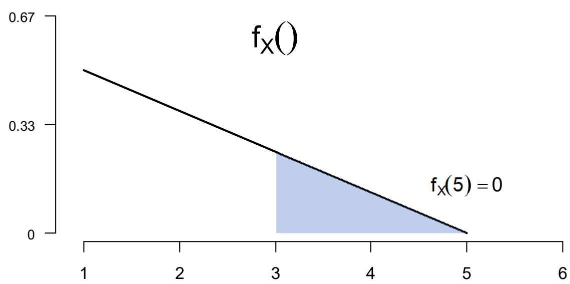
9. Com és la relació entre la potència d'un atac i la resistència? Quin indicador ens quantifica la relació existent? Calculeu quant val en aquest cas.

Segons es veu a la taula, quan creix la potència de l'atac, la resistència té tendència a disminuir i viceversa. Per tant la relació és negativa. Per quantificar la força de la relació podem trobar la covariància o, millor, la correlació.

Cov(P, R) = E[P·R] – E[P] E[R] = 7.325 – 2.8·2.95 = **–0.935** (observem el signe negatiu)
(E[P·R] es pot calcular pel valor esperat del consum d'energia, apartat anterior)

Cor(P,R) = Cov(P, R) / ($\sigma_P$ $\sigma_R$) = –0.935 / 1.1 / 1.223724 = **–0.6946** (correlació prou alta: és una relació forta)

10. Des d'un altre punt de vista, en David pensa que podria implementar tant la potència com la resistència amb una funció de densitat com la de la figura de la dreta. Suposant que en un determinat atac aquests dos atributs fossin independents, calculeu la probabilitat que tant resistència com potència prenguin valors menors que 3.




$P(R<3 \cap P<3) = P(X<3)^2 = (1–P(X>3))^2$

Geomètricament, es pot deduir que P(X>3) = ¼. O be, trobar l'àrea del triangle per sobre de 3, si prèviament deduïm l'equació de la recta $f_X()$ i avaluem a 3.
O be, comprovar que $f_X(3)$ és la meitat de $f_X(1)$, que és l'alçada del triangle gran, i que perquè aquest tingui àrea 1 l'alçada ha de ser ½. O integrar $f_X()$ des de 3 fins a 5. En qualsevol cas, la solució demanada és $(1–1/4)^2$ = **0.5625.**

NOM:________________________________________________________________________________

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

# Problema B

**EN CADA APARTAT DEFINIU EXPLÍCITAMENT LES VARIABLES ALEATÒRIES I EL MODEL QUE ELS HI CORRESPON**

Una comercialitzadora d'ous assegura que, en mitjana, en el procés de triatge i empaquetat cada hora es detecta un ou de categoria superior.
(2 punts) Calculeu la probabiliat d'estar més de dues hores sense detectar cap ou de categoria superior, i la probabilitat de només trobar un ou de categoria superior en les 8 hores d'una jornada laboral.

H és hores entre ous detectats E(H) = 1 H és Exp(λ=1)
N és nombre d'ous cada hora E(N) = 1 N és Poi(λ=1)
**P(H>2) =** 1-P(H<=2) = 1 – (1 – exp(-1*2)) = exp(-2) = **0.135 (14%)**
( o bé N2h és nombre d'ous cada 2 hores E(N2h) = 2 N2h és Poi(λ=2), llavors **P(N2h=0)=** $(2)^0$ * exp(-2) / 0! = **0.135** )

N8 és nombre d'ous de categoria superior en 8 hores N8 és Poi(λ=8)
**P(N8=1) =** $(8)^1$ * exp(-8) / 1! = 8 * exp(-8) **= 0.0027 (0.3 %)**
(en R dpois(1,8) = 0.0027 o bé ppois(1,8)-ppois(0,8) = 0.0030 - 0.00033 = 0.0027)

La comercialitzadora assegura que la mitjana del pes dels ous de categoria superior és de 88 g amb variància 100
(1 punt) Calculeu la probabilitat que la mitjana del pes d'aquests ous en caixes de 25 sigui inferior a 85 g

Pm25 és mitjana del pes d'ous de categoria superior en caixes de 25
Pm25 és Normal amb μ=88 i $\sigma=10/\sqrt{25}=2$

**P(Pm25 < 85) =** P(Z < (85-88)/2) ) = P(Z<-1.5) = 1 – P(Z<1.5) = 1 – pnorm(1.5) = 1 – 0.933 = **0.067 (aprox 7 %)**
(en R pnorm(-1.5) o bé 1-pnorm(1.5))

(1 punt) Calculeu la probabilitat que la suma del pes d'aquests ous en caixes de 25 superi els 2 kg i quart

Sm25 és suma del pes d'ous de categoria superior en caixes de 25
Sm25 és Normal amb μ=88*25=2200 i $\sigma=10*\sqrt{25}=50$)

**P(Sm25 > 2250) =** 1-P(Sm25<2250) = 1 - P(Z < (2250-2200)/50) ) = 1 - P(Z<1) = 1-pnorm(1) = 1-0.8413= **0.1587 (aprox 16 %)**

| | | | | | |
|---|---|---|---|---|---|
| pnorm(1.96) = 0.975 | pnorm(1.5) = 0.933 | pnorm(1.25) = 0.894 | pnorm(1) = 0.841 | pnorm(0.5) = 0.691 | pnorm(0.25) = 0.599 |
| qnorm(0.975)=1.96 | qnorm(0.95)=1.645 | qnorm(0.90)=1.282 | qnorm(0.10) = -1.282 | qnorm(0.05) = -1.645 | qnorm(0.025) = -1.96 |
| pbinom(10,25,0.4)=0.586 | pbinom(11,25,0.4)=0.732 | pbinom(12,25,0.4)=0.846 | pbinom(13,25,0.4)=0.922 | pbinom(14,25,0.4)=0.966 | pbinom(15,25,0.4)=0.987 |
| pbinom(10,25,0.5)=0.212 | pbinom(11,25,0.5)=0.345 | pbinom(12,25,0.5)=0.5 | pbinom(13,25,0.5)=0.655 | pbinom(14,25,0.5)=0.788 | pbinom(15,25,0.5)=0.885 |
| pbinom(10,25,0.6)=0.034 | pbinom(11,25,0.6)=0.078 | pbinom(12,25,0.6)=0.154 | pbinom(13,25,0.6)=0.268 | pbinom(14,25,0.6)=0.414 | pbinom(15,25,0.6)=0.576 |

Una altra comercialitzadora d'ous assegura que en el seu cas la distribució del pes dels ous de categoria superior és Normal amb mitjana 88 g i desviació 8 g
(1 punt) Calculeu la probabilitat que un ou de categoria superior superi el 90 g

Pes és Normal amb μ=88 i σ=8

**P(Pes>90) =** 1-P(Pes<90) = 1 - P(Z<(90-88)/8) = 1 - P(Z<0.25) = 1 - pnorm(0.25) = 1- 0.599 = **0.40 (40%)**

(1 punt) Calculeu la probabilitat que un ou de categoria superior estigui entre 84 i 90 gr

**P(84 < Pes < 90) =** P(Pes<90) – P(Pes<84) = P(Z<(90-88)/8) – P(Z<(84-88)/8) = P(Z<0.25) – P(Z<-0.5) = P(Z<0.25) – (1-P(Z<0.5)) =
= pnorm(0.25) – (1-pnorm(0.5)) = 0.599 – (1-0.691) = **0.29 (29%)**
(en R pnorm(0.25)-pnorm(-0.5) o bé pnorm(0.25)-(1-pnorm(0.5)))

(1 punt) Calculeu el pes que un ou de categoria superior d'aquesta comercialitzadora superarà amb una prob del 5%

P(Pes > max) = 0.05
**P(Pes < max) = 0.95** = P(Z < (max-88)/8)) → (max-88)/8 = 1.645 → max=(1.645*8)+88 max = **101.16 g**
qnorm(0.95) -> 1.645

(2 punts) Inspeccionant aquests ous d'aquesta comercialitzadora, calculeu la probabilitat que es detecti el primer amb pes superior a 90 gr abans del tercer inspeccionat. Trobeu també l'esperança del nombre d'ous a inspeccionar fins detectar cinc amb pes superior a 90 gr

Fins1r és Geom amb p=0.4
**P(Fins1r<3) =** P(Fins1r=1) + P(Fins1r=2) = 0.4 + 0.6*0.4 = 0.4 + 0.24 = **0.64 (64%)**
( o bé P(Fins1r<=2) = 1 – $(0.6)^2$ = 0.64 )
(En R dgeom(0,0.4)+dgeom(1,0.4) o pgeom(k-1,0.4) = pgeom(1,0.4) )

Fins5e és BinNeg amb r=5 i p=0.4
**E(Fins5e) =** r/p = 5 / 0.4 = **12.5 (fins el 13è no esperem haver detectat el cinquè amb pes superior a 90 gr)**

(1 punt) En caixes de 25 ous de categoria superior d'aquesta comercialitzadora, calculeu la probabilitat que cap d'ells superi els 90 g, i la probabilitat que més de la mitad superin els 90 g

N és nombre d'ous de pes superior a 90 g en caixes de 25
N és Binomial amb n=25 i p=0.4

**P(N=0) =** $\binom{25}{0} 0.4^0\ 0.6^{25} =$ **0.0000028** (0.0003 %) (en R dbinom(0,25,0.4) )

**P(N>=13) =** 1-P(N<=12) = 1- pbinom(12,25,0.4) = 1 - 0.8462 = **0.1538 (15%)**