NOM:_____COGNOM:_____

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

Problema 1 (Bloc A)

1. En un taller de màgia matemàtica, es dipositen tres monedes no trucades en una caixa i se n'extreu una a l'atzar. Després es llença dues vegades, quina és la probabilitat que surtin dues cares? [0,5 punts]

En aquest cas, com que totes les monedes són no trucades i totes tenen la mateixa probabilitat de ser extretes de la caixa, la probabilitat serà la mateixa que la de que surtin dues cares en llençar una moneda a l'aire dues vegades.

P((c,c)) = P('cara en el primer llançament') · P('cara en el segon llançament') = 1/2· 1/2 = 1/4

Tot seguit se substitueix una de les monedes no trucades, per una de trucada amb dues cares. Ara a la caixa hi ha dues monedes no trucades (amb cara i creu) i una moneda que té cara a ambdues bandes. Tornem a repetir l'experiment d'extreure una moneda a l'atzar.

2a. Si es llença la moneda dues vegades, quina és la probabilitat que surtin dues cares? [0.75 punts]

Anomenem M1:'Extreure una de les monedes no trucades' i MT: 'Extreure la moneda trucada"

P((c,c))= P((c,c)| M1) · P(M1) + P((c,c)| MT) · P(MT) = ¼ · 2/3 + 1 ·1/3 = 1/6 + 1/3 = ½

2b. Si es llença tres vegades, quina és la probabilitat que surtin <u>almenys</u> dues cares? [0,75 punts]

En primer lloc estudiem la situació en una moneda no trucada.

P('almenys dues cares'|M1)= P((c,c,c) ∪ (+,c,c) ∪ (c,+,c) ∪ (c,c,+)) = 4/8 = ½ ja que són conjunts disjunts i en cada cas la probabilitat és 1/8.

En el cas de la moneda trucada sempre sortiran tres cares, per tant la probabilitat d'almenys dues cares és 1.

P('almenys dues cares') = P('almenys dues cares' | M1) · P(M1) + P('almenys dues cares' | MT) · P(MT) = ½ · 2/3 + 1 · 1/3 = 2/3

2c. Finament es llença la moneda tres vegades i s'obtenen tres cares, quina és la probabilitat que la moneda llençada hagi estat la trucada? [1 punt]

P(MT | (c,c,c)) = P(MT ∩ (c,c,c)) / P((c,c,c)) = P((c,c,c)| MT) · P(MT) / P((c,c,c)) = 1/3 : 5/12 = 12/15

La següent activitat es realitza amb dues bosses. Una bossa A conté 2 boles negres i 3 boles blanques i una segona bossa B conté 4 boles negres i 2 boles blanques. Es tria una de les bosses a l'atzar i se n'extreuen tres boles de manera consecutiva (sense reposició). Volem observar el nombre de boles blanques extretes (N)

3a. Trobeu la funció de probabilitat de la variable aleatòria N:"Nombre de boles blanques extretes". [2 punts]

k	P_N (k)
0	1/10
1	9/20
2	2/5
3	1/20

Calculem:

P(N=0) = P((n,n,n)| A) · P(A) + P((n,n,n)|B) · P(B) = 0 · ½ + 4/6 · 3/5 ·2/4 · ½ = 1/10

P(N=1) = [P((b,n,n)| A) + P((n,b,n)| A) + P((n,n,b)| A)] · P(A) + [P((b,n,n)| B) + P((n,b,n)| B) + P((n,n,b)| B)] · P(B)
= [3/5 ·2/4 · 1/3 + 2/5 ·3/4 · 1/3 + 2/5 ·1/4 · 3/3]· ½ + [2/6 ·4/5 · 3/4 + 4/6 ·2/5 · 3/4 + 4/6 ·3/5 · 2/4]· ½ =
= 3 · 1/10 · ½ + 3 · 1/5 · ½ = 9/20

P(N=2) = [P((b,b,n)| A) + P((b,n,b)| A) + P((n,b,b)| A)] · P(A) + [P((b,b,n)| B) + P((b,n,b)| B) + P((n,b,b)| B)] · P(B)
= [3/5 ·2/4 · 2/3 + 3/5 ·2/4 · 2/3 + 2/5 ·3/4 · 2/3]· ½ + [2/6 ·1/5 · 4/4 + 2/6 ·4/5 · 1/4 + 4/6 ·2/5 · 1/4]· ½ =
= 3 · 1/5 · ½ + 3 · 1/15 · ½ = 2/5

P(N=3) = P((b,b,b)| A) · P(A) + P((b,b,b)|B) · P(B) = 3/5 ·2/4 · 1/3 · ½ + 0 · ½ = 1/20

3b. Calculeu l'esperança i la desviació típica de la variable aleatòria N. [1 punt]

E(N) = 0 · 1/10 + 1 · 9/20 + 2 · 2/5 + 3 · 1/20 = 1,4
$E(N^2) = 0^2 \cdot 1/10 + 1^2 \cdot 9/20 + 2^2 \cdot 2/5 + 3^2 \cdot 1/20 = 2,5$
$Var(N) = 2,5 - 1,4^2 = 0,54$ i $\sigma_N = 0,7348$

3c. Si el nombre de boles blanques extretes és 2, quina és la probabilitat que la bossa escollida hagi estat l'A? [0,5 punts]

P(A | N=2) = P(A ∩ N=2) / P(N=2) = P(N=2 | A) · P(A) / P(N=2) = [3 · 1/5 · ½] : 2/5 = 3/10 : 2/5 =3/4

4. Considereu la variable aleatòria X, amb la següent funció de densitat:

$$f(x) = \begin{cases} x^2 & 0 \le x \le 1 \\ \dfrac{ax+b}{4} & 1 < x \le \dfrac{7}{3} \\ 0 & altrament \end{cases}$$

Trobeu els valors de a i b per a que f(x) sigui una funció de densitat contínua. [2 punts]

Per tal que sigui una funció de densitat s'ha de complir que:

$\int_{-\infty}^{+\infty} f(x)\,dx = 1$, per tant, en aquest cas $\int_0^1 x^2\,dx + \int_1^{7/3} (\frac{ax+b}{4})\,dx = 1$

$\int_0^1 x^2\,dx = \left[\frac{x^3}{3}\right]_0^1 = 1/3$

$\int_1^{7/3} (\frac{ax+b}{4})\,dx = \left[\frac{ax^2/2+bx}{4}\right]_1^{7/3} = 49a/72 + 7b/12 - (a/8 - b/4) = 5/9a + 1/3b$

Per tant, 5/9a + 1/3b = 2/3 i es té que 5/3 a + b = 2

Per altra banda, perquè sigui contínua es té que $\lim_{x\to 1^+} f(x) = f(1) = 1^1 = 1$, per tant, $\frac{a+b}{4} = 1$

Cal per això resoldre $\begin{cases} \frac{5}{3}a + b = 2 \\ a + b = 4 \end{cases}$, obtenint a=-3 i b=7

Finalment es comprova que per aquests valors f(x) és positiva. (Aquest darrer punt no es penalitza si no s'ha explicitat a l'examen).

Considereu ara la variable T que estudia el temps (en segons) d'un determinat experiment aleatori i que segueix una distribució amb la següent funció de densitat:

$$f(t) = \begin{cases} 3t^2 & \text{si } 0 \le t \le 1 \\ 0 & \text{altrament} \end{cases}$$

5. a) Calculeu la funció de distribució de T (només cal calcular la funció de distribució sense haver de comprovar que f(t) és una funció de densitat) [0,5 punts]

$F(t) = \int_0^t 3x^2\,dx = [x^3]_0^t = t^3$

$$F(t): \begin{cases} 0 & t < 0 \\ t^3 & 0 \leq t \leq 1 \\ 1 & t < 1 \end{cases}$$

b) Calculeu el valor esperat de T. [0,75 punts]

E(T) = $\int_0^1 x \cdot 3x^2 \, dx = \left[\frac{3x^4}{4}\right]_0^1 = \frac{3}{4}$

c) Calculeu la variància de T. [0,75 punts]

E(T^2) = $\int_0^1 x^2 \cdot 3x^2 \, dx = \left[\frac{3x^5}{5}\right]_0^1 = \frac{3}{5}$

Var(T) = E(T^2) - E(T)^2 = $\frac{3}{5} - \left(\frac{3}{4}\right)^2$ = 0,0375

NOM:_____

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

Problema bloc B

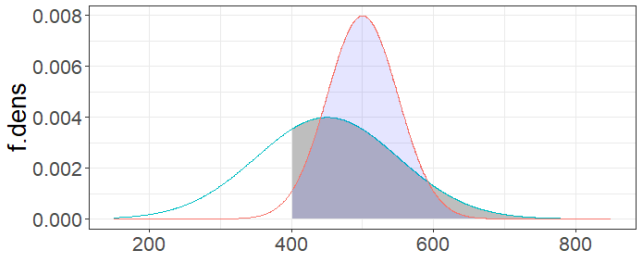
Una investigació sobre un cert tipus d'imatges (reals i generades per IA) estudia característiques com l'índex de nitidesa, la mida i el format dels fitxers corresponents. Per a la nitidesa s'ha establert un índex (un valor més alt indica més nitidesa) amb distribució Normal, que pel cas de la ***nitidesa d'imatges reals (Nr)*** té mitjana 450 i desviació estàndard 100, i pel cas de ***nitidesa d'imatges generades (Ng)*** té mitjana 500 i desviació estàndard 50. La ***mida d'imatges (M)***, tant reals com generades, no té distribució Normal, i té mitjana 100 MB i variància 2500.

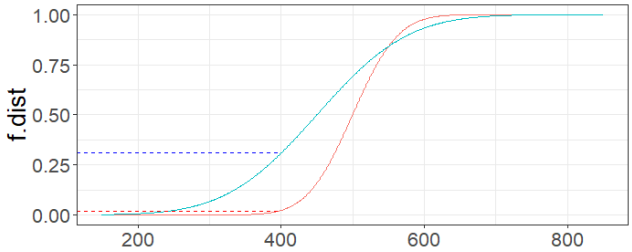
Definiu explícitament variables i models on calgui, i expliciteu formalment els passos en el càlcul de probabilitats

1.- (1 punt) Dibuixeu, pels dos índexs de nitidesa, les dues funcions de densitat en un gràfic, i les dues funcions de distribució en un altre gràfic. Indiqueu, en els dos gràfics, els valors de l'eix horitzontal i només en el cas de les funcions de distribució, els valors de l'eix vertical

Nr és N(μ=450 , σ=100)

Ng és N(μ=500 , σ=50)





2.- (1 punt) Calculeu la probabilitat que per a una **imatge real**, l'índex de **nitidesa** sigui superior a 400. Calculeu-la també per a una **imatge generada** i dibuixeu-les al gràfic anterior

Nr és N(μ=450 , σ=100) Ng és N(μ=500 , σ=50)

P(Nr > 400) = P(Z > (400-450)/100) = P(Z > -0.5) = 1-P(Z<-0.5) = P(Z<0.5) = pnorm(0.5) = 0.6915 69 %

P(Ng > 400) = P(Z > (400-500)/50) = P(Z > -2) = 1-P(Z<-2) = P(Z<2) = pnorm(2) = 0.9772 98 %

(en els gràfics s'ha de veure que la probabilitat és més gran en imatges generades, ja que Ng té mitjana més gran i desviació més petita que fa que acumuli més probabilitat per sobre del 400; 69% i 98% a l'esquerra de 400 en la funció de densitat, i 0.02 (1-0.98) i 0.31 (1-0.69) a la funció de distribució)

3.- (1 punt) Calculeu la probabilitat que l'índex de **nitidesa mitjà** d'una col·lecció de 25 **imatges reals** sigui superior a 400.

N25 és mitjana de 25 Nr i, com és combinació lineal de Normals, també és Normal amb amb μ=450 i σ=100/sqrt(25)=20

P(N25 > 400) = P(Z > (400-450)/20) = P(Z > -2.5) = 1-P(Z<-2.5) = P(Z<2.5) = pnorm(2.5) = 0.9938 99 %

4.- (1 punt) Calculeu la probabilitat que la **mida total** d'una col·lecció de 25 **imatges** sigui inferior als 2250 MB. Indiqueu també quin valor de MB ens assegura que el 50% de les mides totals no el superaran

M25 és suma de 25 M i pel TCL (no és suma de normals però n=25 és prou gran)
és Normal amb μ=100*25=2500 i σ=50*sqrt(25)=250

P(M25 < 2250) = P(Z < (2250-2500)/250) = P(Z < -1) = 1-P(Z<1) = 1 - pnorm(1) = 1-0.8413447 = 0.1586553 16 %

2500 MB és el centre de la distribució de la suma (M25) que és simètrica, per tant és el valor que assegura que la probabilitat de no superar-lo és un 50%

pnorm(0.1) = 0.5398278	pnorm(0.5) = 0.6914625	pnorm(1) = 0.8413447	pnorm(1.5) = 0.9331928	pnorm(2) = 0.9772499	pnorm(2.5) = 0.9937903
qnorm(0.1) = -1.281552	qnorm(0.5) = 0	qnorm(0.01) = -2.32635	qnorm(0.15) = -1.03643	qnorm(0.2) = -0.841621	qnorm(0.25) = -0.67449

Ara estudiarem el format i el nombre d'imatges de fenòmens meteorològics rebudes a la pàgina web d'un canal televisiu. Definiu explícitament variables i models on calgui, i expliciteu formalment els passos en el càlcul de probabilitats

5.- (3 punts) Si es reben una mitjana de 42 imatges cada setmana, calculeu la probabilitat que un dia es rebin 8 imatges

FMset és Poison amb λ=42 **FMdia és el nombre d'imatges rebudes al dia i segueix una Poisson amb λ = 42/7 = 6**

P(FMdia=8) = (6^8 *exp(-6)) / 8! = 4163.352 / 40320 = 0.1032 **10 %**

I quin és el temps esperat en hores entre arribades de dues imatges consecutives?

FMh és Poisson amb λ=6/24 = 0.25 **Th és el temps en hores i segueix Exp amb λ=0.25=1/4**

E(Th) = 1 / λ = 24/6 = 1/0.25 = **4 hores**

I calculeu el nombre d'hores amb que podem assegurar que no arribaran imatges amb una probabilitat de 0.6

P(Th > a) = 0.60 = 1-P(Th<a) = 1-(1-exp(-0.25*a)) = exp(-0.25*a) a=ln(0.60)/-0.25=2.04 **2 hores**

6.-(2 punts) Si per cada 20 imatges que es reben, l'esperança del nombre d'imatges en **format jpeg** és 10, calculeu la probabilitat de trobar-ne 12 d'aquest format en 20 imatges

Jpeg és el "nombre d'imatges en format jpeg"

Jpeg és Bin(n=20,p) amb E(Jpeg) = 10 = n*p = 20*p → p = 10/20 = 0.5

P(Jpeg=12) = $\binom{20}{12}(0.5)^{12}(0.5)^{8}$ = dbinom(12,20,0.5) = = 0.1201 **12.01 %**

= choose(20,12)*(0.5)^12*(0.5)^8 = (20!/12!8!)*(0.5)^20

= 125970*0.000000954

I la probabilitat d'haver de revisar 5 imatges fins observar la primera en format jpeg

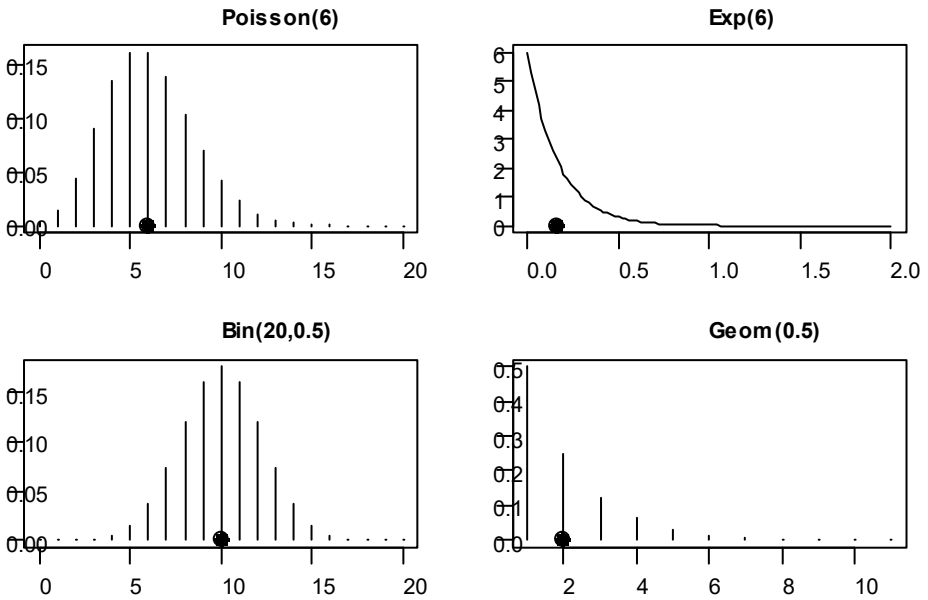
Jpeg1 és el "nombre d'imatges a revisar fins la primera en format jpeg"

Jpeg1 és Geom(p=0.5)

P(Jpeg1=5) = $(0.5)^{4}(0.5)$ =0.5^5 = **0.0312** **3.1 %**

= dgeom(5-1,0.5)

7.- (1 punt) Dibuixeu de forma aproximada les 4 funcions de probabilitat o de densitat de cadascun dels models implicats en les preguntes 5 i 6 anteriors. Indiqueu el model i valor de paràmetre concret i senyaleu el valor de l'esperança en cada gràfic



Els valors de les esperances són 6, 1/6, 10 i 2, respectivament.

NOM:_____COGNOM:_____

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

Problema 3 Bloc C

1. A partir d'una mostra de fitxers volem donar un interval de confiança per a l'esperança de la variable "temps de compressió", assumint distribució Normal. Si la mostra obtinguda és de mida 10 amb mitjana de 7.2, i considerem que la variància poblacional és 1.4, l'interval calculat al 95% és: (1 punt)

Calculem l'interval de confiança fent servir el valor de la variança poblacional i la mida mostral indicada i $z_{1-\alpha/2}$=1.96:

7.2 ± 1.96 (SQRT(1.4)/SQRT(10)) = **[6.4666, 7.9333]**

2. En dos programes de compressió diferents que tenen una durada amb distribució normal de sigma= 6 segons i sigma= 4 segons respectivament, realitzem una operació de mostreig amb 30 casos per al primer programa i n_2 per al segon. Si volem estimar la diferència en la mitjana dels temps, quina ha de ser la mida de la mostra del 2n programa per tal que l'error tipus d'aquest estudi sigui, com a màxim, 1.25? Justifiqueu el procés de càlcul. (2 punts)

L'estadístic emprat és la diferència de mitjanes de dues mostres independents: y1 – y2
V(y1 – y2) = $6^2/30 + 4^2/n2$. L'error tipus ha de ser com a màxim 1.25, és a dir:
$6^2/30 + 4^2/n2 \leq 1.25^2 = 1.5625$
$16/n2 \leq 0.3625$
$n2 \geq 16/0.3625 = 44.14$
La mida de la segona mostra (n2) ha de ser almenys **45**.

3. En 12 fitxers, la diferència D entre els temps d'execució de dos programes de compressió de fitxers ha estat de mitjana 3.1 segons i desviació 4. Acceptant que D ~ Normal, per saber com són de diferents, calculeu l'interval de confiança al 99%, i plantegeu si és raonable pensar que els dos compressors triguen el mateix en mitjana. (1 punt)

IC(μ, 99%) = 3.1 ± $t_{11,\ 0.995}$ 4/SQRT(12) = 3.1 ± 3.109·1.1547 = [**-0.48, 6.68**]

Primer intent: el valor 0 (equivalent a dir que les mitjanes dels dos compressors són iguals) es troba a l'interior d'aquest IC, llavors es podria dir que amb confiança 99% és una situació versemblant.

Segon intent: Si no voleu que el mot "raonable" us enredi, tingueu en compte que aquest resultat s'ha trobat amb un procediment "molt generós" (que només falla una de cada 100 vegades) que resulta en un interval molt ample. És evident que la noció del que és *raonable* o *versemblant* està lligat al grau de confiança emprat: aquí, per exemple, un IC al 95% deixaria fora el valor 0. En general, no hauríem d'utilitzar ICs per posar a prova valors concrets.

4.Escriu la definició, en paraules, de la variància de la v.a. X i la variància de la variable mitjana mostral. Eviteu posar només fórmules. Poseu un exemple de cadascun d'ells. (1 punt)

La variància de la variable aleatòria X és la mesura de dispersió de la distribució de X, i la variància de la variable mitjana mostral és la mesura de dispersió del paràmetre estimat.

Un exemple de la variància de la v.a. X pot ser la variància calculada per a la variable temps de compressió en una mostra particular de n observacions. I la variància de la variable mitjana mostral es pot obtenir calculant la variança per a un conjunt de mitjanes mostrals obtingudes de diferents mostres (tenim una mitjana mostral per cada mostra considerada).

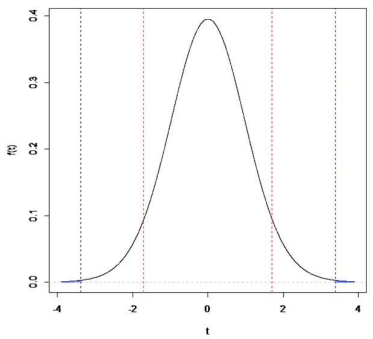
Estudiarem el temps de compressió mesurat en 30 documents i realitzem 3 càlculs amb els següents resultats:

A: t = -3.3853, df = 29, p-value = 0.002058 alternative hypothesis: true mean is not equal to 8 90 percent confidence interval: 6.758417 7.588250 sample estimates: mean of x 7.173333	B: t = -3.3853, df = 29, p-value = 0.002058 alternative hypothesis: true mean is not equal to 8 95 percent confidence interval: 6.673901 7.672766 sample estimates: mean of x 7.173333	C: t = -3.3853, df = 29, p-value = 0.002058 alternative hypothesis: true mean is not equal to 8 99 percent confidence interval: 6.500241 7.846426 sample estimates: mean of x 7.173333

5. Què són els valors t de l'output de R i perquè són iguals? (1 punt)

Les t són les ràtios senyal/soroll, on el senyal és la diferència entre la mitjana de la mostra i el valor 8, i el soroll l'error estàndard de la mitjana, indicant que el senyal és més de 3 vegades superior a la magnitud del soroll. Però ni la mitjana ni l'error estàndard canvien quan es fan servir diferents graus de confiança, com aquí es fa (d'esquerra a dreta, 90, 95 i 99%).

6. Com s'ha calculat el p-valor de la sortida A. Feu un gràfic i representeu -aproximadament- aquest p-valor. (1 punt)



L'estadístic t de la prova segueix una distribució t-Student amb 29 graus de llibertat quan les premisses són les correctes i quan el valor esperat de la variable realment val 8. Com la mitjana observada és 7.17, l'estadístic pren el valor -3.3853 (punt que separa l'àrea blava de l'esquerra del centre de la distribució). El P valor és la probabilitat d'observar un resultat més extrem, per tant més petit que -3.3853 o més gran que 3.3853. Les dues àrees blaves sumades representen aquesta probabilitat 2*pt(-3.3853, 29)= 0.002 (associada a mostres d'una variable amb valor esperat igual a 8). En el gràfic, en vermell s'indica el valor crític a la distribució t amb 29 graus de llibertat i les línies negra i blava marquen l'àrea del p-valor per l'esquerra i per la dreta (que es correspon amb els valors de l'estadístic t i -t.

7. S'obtenen 60 compressions realitzades, en 52 no es supera el llindar de 8 segons. Indiqueu l'estimació puntual i un interval de confiança al 90% pel percentatge de compressions que no superen el llindar. Justifiqueu com feu el càlcul i interpreteu el que heu trobat (2 punts)

L'estimació puntual de π, P = 52/60 = 0,875
Per fer el càlcul de l'IC ens cal buscar primer l'error estàndard (se):
se = sqrt((0.875*0.125)/60) = 0.0434 i $z_{0.95}$=1.645. Obtenim que l'interval de confiança és
P – 1.645*0.0434 = 0.875 – 0.071393= 0.804; P + 1.645*0.0434 = 0.875 + 0. 071393= 0.946
Interpretació: Podem pensar, amb una confiança del 90%, que la proporció real de temps d'aquest programa de compressió que no arriben als 8 segons està entre 80.4% i 94.6%

[Resolució alternativa: Si prenem el valor de P=0.5 per fer el càlcul de la se trobem se = sqrt(0.5*0.5/60) =0.0646) i obtenim l'IC de [0.769; 0.981]. Aquesta solució no és òptima, ja que l'estimació de π (0.875) no es un valor proper al centre.]

8. En les mesures preses a les 60 compressions hem obtingut un valor $s^2 = 2^2$. Calculeu l'interval de confiança del 95% per a la desviació tipus del temps de compressió (0.5 punts), i l'interval de confiança del 90% per a la variància (0.5 punts).

A partir de la fórmula: $IC(\sigma^2, 0.95) = \left[\frac{s^2(n-1)}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{s^2(n-1)}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$

Fent servir els valors de qchisq(0.975, 59)=82.11741 i qchisq(0.025, 59)= 39.66186 per al càlcul de l'IC del 95% tenim que: (4*59)/82.12=2.873843 ; (4*59)/39.66=5.95058
L'IC del 95% per a la variància: [2.874, 5.951]
Per a obtenir l'IC de la desviació tipus s'ha de fer l'arrel quadrada per l'interval per a sigma: [**1.695, 2.439**]

Similarment per trobar l'IC del 90% per a la variància: [**3.028**, **5.575**]
Valors pel al càlcul del IC del 90%: qchisq(0.05, 59)= 42.33931; qchisq(0.95, 59)= 77.93052.

NOM:_____

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

Problema bloc D

Ja hem treballat el tema de la qualitat dels vins, ara ens ocuparem d'un altre plaer gastronòmic: el pernil. Disposem d'una base de dades amb avaluacions de diferents experts de 250 marques de pernil: la qualitat Q és un número entre 0 i 10; el tipus T pot ser 1) bodega, 2) pinso (*recebo*), 3) gla (*bellota*); també es recull el pes P en kilograms de la peça, i finalment l'origen R (Ext: Extremadura, Alt: altre). Volem analitzar la relació d'aquests factors amb la qualitat del pernil.

A

```
            Estimate Std. Error
(Intercept)   5.5192     0.1028
Residual standard error: 1.625 on 249 deg. of fr.
```

B

```
            Estimate Std. Error
(Intercept)   4.9382     0.1326
T2            0.6972     0.2121
T3            1.8785     0.2503
Residual standard error: 1.471 on 247 deg. of fr.
Multiple R-squared:  0.1876,
F-statistic: 28.52 on 2 and 247 DF,  p-value: 7.168e-12
```

C

```
            Estimate Std. Error
(Intercept)   3.2959     0.9218
T2            0.6758     0.2114
T3            1.8989     0.2494
P             0.2342     0.1301
Residual standard error: 1.464 on 246 deg. of fr.
Multiple R-squared:  0.1982,
F-statistic: 20.27 on 3 and 246 DF,  p-value:
9.028e-12
```

D

```
            Estimate Std. Error
(Intercept)  3.29784    0.92351
T2           0.67926    0.21214
T3           1.90608    0.25105
P            0.23744    0.13079
RExt        -0.05607    0.18715
Residual standard error: 1.467 on 245 deg. of fr.
Multiple R-squared:  0.1985,
F-statistic: 15.17 on 4 and 245 DF,  p-value: 4.298e-11
```

1. Expliqueu que hem obtingut en el output A: quin és el model, quines variables intervenen, què resultats hem de destacar (no cal que trobeu intervals de confiança si no es demanen, però utilitzeu alguna mesura de la incertesa de l'estimació quan sigui oportú).
   El model és el més simple possible: no hi ha cap variable explicativa de la resposta Q. Es tracta d'estimar la qualitat mitjana (5.52), i l'error tipus de la mitjana val 0.10. La valoració mitjana d'un pernil és 5.52 (SE=0.10), i la desviació tipus de la valoració en la base de dades és 1.625 punts.

2. D'acord amb la sortida que veiem al output B, quins canvis s'han introduït al model? Com s'interpreten les noves estimacions? Són canvis rellevants, és a dir, poden explicar (part de) la valoració del pernil? Justifiqueu la resposta.
   Hem introduït com a variable explicativa el tipus T. El valor de referència és la categoria 1 (bodega): el terme *Intercept* s'interpreta com la valoració mitjana per a aquest tipus (4.94, SE=0.13). El coeficient de T2 és l'increment de valoració per a un pernil de pinso (+0.70, SE=0.21), i el de T3 l'increment per a un pernil de gla respecte al de bodega (+1.88, SE=0.25). La desviació tipus residual trobada (desviació estimada dins de cada categoria) ha baixat a 1.47 punts. El tipus de pernil explica el 18.76% de la variabilitat observada de la valoració Q.

3. Comenteu de manera semblant el que es veu al output C. Perquè ha canviat tant l'estimació del terme "Intercept"?
   Afegim al model el pes del pernil. *L'Intercept* ha canviat sensiblement (3.30, SE=0.92), perquè en aquest model és tracta de la valoració base per a un pernil de bodega de pes 0Kg. Els coeficients T2 i T3 són molt semblants als del model anterior, i el pes té un coeficient 0.2342 (SE=0.13), que significa l'increment en la valoració per Kg de pes. La desviació residual és pràcticament igual, i el coeficient de determinació ha pujat un 1% escàs (19.82%). Per tant, el pes del pernil no sembla un factor molt rellevant en la qualitat.

4. Continueu amb l'analítica del output D.
   Afegim el factor Origen: la referència és "Altre", però veiem que el canvi si l'origen és "Extremadura" és molt insignificant (-0.056, SE=0.19). A més a més, els altres coeficients tenen variacions molt petites: aquest model és pràcticament igual a l'anterior perquè l'origen no sembla afectar a la qualitat del pernil.

5. Si tenim dues peces del mateix pes i origen, però una és bodega i l'altra és pernil de gla, quina és la diferència de valoració mitjana entre ambdues peces? Responeu amb un (o varis) interval(s) de confiança al 95%.

La diferència és: 1.8785 ± 1.96×0.2503 = (1.39, 2.37), utilitzant el model més simple que incorpora el tipus. Amb altres models:
C: 1.8989 ± 1.96×0.2494 = (1.41, 2.39), suposant mateix pes, i D: 1.9061 ± 1.96×0.2510 = (1.41, 2.40), mateix pes i origen.

6. Podem dir que una peça més grossa té més possibilitats de rebre millors valoracions? I una peça d'Extremadura? Raoneu les respostes, i quantifiqueu els "efectes" de cadascú dels factors.
   Si calculem els intervals de confiança per als coeficients associats amb l'últim model:
   P: 0.23744 ± 1.96×0.13079 = (-0.019, 0.4938)
   Veiem que el 0 hi és a dins de l'interval, el que indicaria que no és pot descartar que el pes no jugui cap paper en la valoració. En tot cas, potser s'hi veu una tendència que es tindria que demostrar amb una mostra aleatòria apropiada.
   R: -0.05607 ± 1.96×0.18715 = (-0.42, 0.31)
   Si la peça prové d'Extremadura no implica un canvi clar en cap direcció, tant podria baixar 0.4 com pujar 0.3.

7. Perquè al primer anàlisi no tenim un indicador "`Multiple R-squared`"? Què ens diu a les altres anàlisis?
   El R^2 és la quantitat (percentual) de variabilitat de la variable resposta Q explicada por diversos factors. Si no hi ha cap factor al model no es pot descomposar la variabilitat entre explicada i residual. Com hem dit abans, el tipus de pernil pot explicar quasi un 19%, i amb el pes pot pujar un 1%, però l'origen no aporta res.