

NOM: _____ COGNOM: _____

Contesteu cada pregunta en el seu lloc. Expliqueu justifiqueu els càlculs.

Problema 1 (A)

En una població on la malaltia X és endèmica, se sap que la prevalença de la malaltia és d'un 12% (és a dir, el 12% de la població té la malaltia). Es disposa d'una prova per a detectar la malaltia, però no és totalment fiable, ja que, la sensibilitat del test és del 90% (és a dir, dóna positiu en un 90% dels casos de persones realment malaltes); però dóna un 5% de falsos positius (és a dir, dóna erròniament positiu en un 5% de persones sanes).

1. Dibuixa l'arbre associat a aquest esdeveniment. (0.5 punts)
2. Si una persona està malalta, quina és la probabilitat que el test li hagi sortit negatiu (fals negatiu)? (0.5 punts)
3. Es demana la taula de probabilitats conjuntes i marginals. A més, també es demana desenvolupar formalment el càlcul de la probabilitat que el test sigui negatiu? (2 punt)
4. Es demana la taula amb les probabilitats condicionades al resultat del test: $P(M|+)$, $P(M|-)$, $P(S|+)$ i $P(S|-)$. Expliqueu breument el resultat que heu trobat. (2 punt)
5. L'especificitat del test és la capacitat del test de donar negatiu en persones sanes. Quina és l'especificitat? (0.5 punts)

6. Quina és la probabilitat que una persona estigui sana si li ha sortit el test positiu? (1 punt)

7. Hi ha dependència o independència entre estar malalt i el resultat del test (justifica la resposta)? (0.5 punts)

En aquesta població hi ha 2 hospitals, X i Y. Aquesta és la funció de probabilitat conjunta en relació al nombre de pacients que ingressen setmanalment:

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	0.02	0.04	0.35
$X = 1$	0.03	0.00	0.10
$X = 2$	0.40	0.05	0.01

8. Quant valen les seves esperances i variàncies? (1.5 punt)

9. Es demana la correlació entre X i Y i interpretar el resultat? (1 punt)

10. Quina és la distribució de la probabilitat del nombre d'ingressos a l'hospital Y en una setmana que ha ingressat a X només un pacient? (0.5 punt)

NOM: _____

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

Problema bloc B

Mentre que mesures sofisticades de control i seguretat en aeroports estan pendents de temes legals, de moment es segueix funcionant, entre altres, amb arcs de seguretat i control del pes dels equipatges. Contesteu les següents preguntes indicant, per a totes les variables necessàries, la definició i model probabilístic (també en les que no es posa explícitament)

Per una part, se sap que al passar per l'arc de seguretat sona una alarma a una mitjana de 5 passatgers per hora. Es demana:

- (1 punt) Definiu unes variables pel nombre de passatgers als que sona l'alarma: una variable de nombre per hora i una altra de nombre per minut. I definiu unes variables pel temps entre passatgers als que sona l'alarma: una pel temps en hores i una altra pel temps en minuts Indiqueu els models probabilístics corresponents i els seus paràmetres

- (1 punt) Calculeu la probabilitat que, en una hora, a cap passatger li soni l'alarma i la probabilitat que li soni com a màxim a 1 passatger

- (1 punt) Calculeu la probabilitat d'estar més de 10 minuts sense que soni l'alarma, i la probabilitat d'estar-ne més de 10 sense que soni l'alarma quan porta almenys 5 minuts sense sonar

- (1 punt) Quin és el temps en minuts que es pot esperar sense que soni l'alarma de l'arc de seguretat? (Justifiqueu el càlcul)

Per una altra part, se sap que el pes de l'equipatge dels passatgers segueix una distribució Normal de mitjana 24 kg i variància 4 kg². Es demana:

- (1 punt) Calculeu la probabilitat que l'equipatge d'un passatger passi de 25 kg

- (1 punt) Calculeu la probabilitat que en un grup de 10 passatgers n'hi hagi 2 amb equipatge amb pes superior a 25 kg

- (1 punt) Calculeu el nombre esperat de viatgers que passaran fins que un tingui equipatge superior a 25 kg

- (1.5 punts) Calculeu la probabilitat que en el grup de 10 passatgers el pes mitjà de l'equipatge superi els 25 kg

- (1.5 punts) Calculeu el pes màxim que podem observar amb un error del 10%

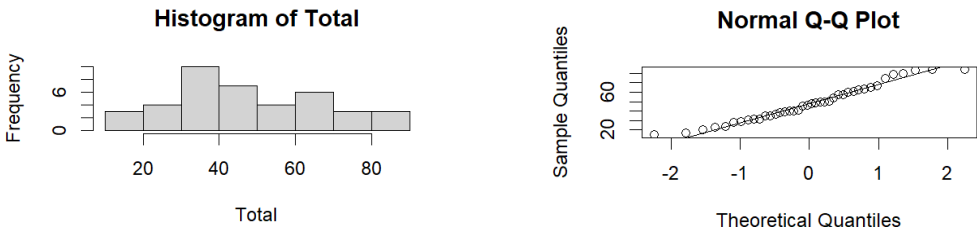
pnorm(0.1) = 0.5398278	pnorm(0.5) = 0.6914625	pnorm(0.6) = 0.7257469	pnorm(0.9) = 0.8159399	pnorm(1.0) = 0.8413447	pnorm(1.6) = 0.9452007
qnorm(0.1) = -1.281552	qnorm(0.2) = -0.8416212	qnorm(0.5) = 0	qnorm(0.6) = 0.2533471	qnorm(0.8) = 0.8416212	qnorm(0.9) = 1.281552

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

Problema 3 (Bloc C)

A Catalunya cada mes de març més de cent mil alumnes participen a les proves Cangur de resolució de problemes. Hi participa alumnat des de 5è de primària fins a 2n de batxillerat i cicles formatius. Per l'edició de 2023, s'ha escollit una mostra a l'atzar de 40 participants de 5è de primària i s'han obtingut les següents dades sobre la puntuació Total (T):

$\sum_{i=1}^{40} t_i = 1920,5$ i $\sum_{i=1}^{40} t_i^2 = 107.111,5$



1. A partir de les dades anteriors, doneu una estimació puntual de la mitjana de puntuació Total (T) de la prova de 5è de primària. Quin és l'error estàndard d'aquesta estimació? (1,5 punts)

2. A partir dels gràfics, argumenteu si podem suposar que la puntuació Total (T) de la prova de cinquè de primària segueix una distribució normal. (1 punt)

3. A continuació podeu trobar dos outputs d'R. Un correspon a l'IC de la mitjana de la puntuació total amb un 90% de confiança i l'altre amb un 99%. Indiqueu de manera raonada quin és cadascun (1 punt)

<div>OUTPUT A - One Sample t-test</div> <div>data: Total</div> <div>t = 15.534, df = 39, p-value < 2.2e-16</div> <div>alternative hypothesis: true mean is not equal to 0</div> <div>?? percent confidence interval:</div> <div>42.80476 53.22024</div> <div>sample estimates:</div> <div>mean of x ??</div>	<div>OUTPUT B -One Sample t-test</div> <div>data: Total</div> <div>t = 15.534, df = 39, p-value < 2.2e-16</div> <div>alternative hypothesis: true mean is not equal to 0</div> <div>?? percent confidence interval:</div> <div>39.64267 56.38233</div> <div>sample estimates:</div> <div>mean of x ??</div>
---	--

IC(μ_T , 90%) =
IC(μ_T , 99%) =

4. Calculeu un interval de confiança (IC) del 95% per a la mitjana de la puntuació Total de la prova de cinquè de primària. Recordeu justificar les premisses. (1 punt)

Ara s'afegeixen també les dades del nivell de 2n de batxillerat perquè es vol estudiar si hi ha diferències entre els dos nivells respecte la variabilitat de les dades. En concret, s'escull també una mostra de 40 alumnes de 2n de batxillerat en què la seva puntuació total (TB) segueix una distribució normal i que té 172,1458 per variància mostral.

5.- Calculeu l'interval de confiança del 95% per al quocient de variàncies entre el total de puntuació de la prova de cinquè de primària i 2n de batxillerat. Argumenteu si podem concloure que tenen o no la mateix variància. Interpreteu en termes de desviacions tipus. (2 punts)

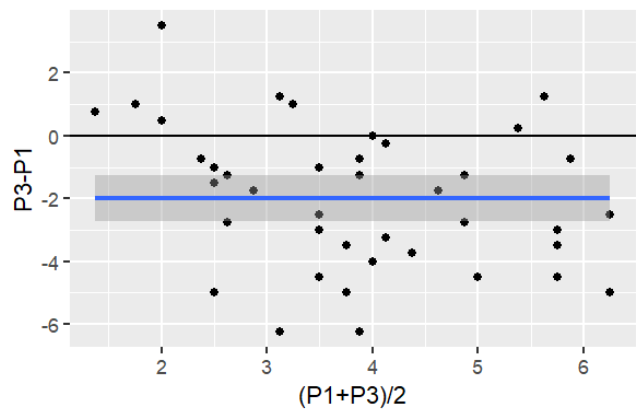
En el nivell de 2n de batxillerat la prova consta de tres parts amb deu problemes a cada part. A més de la puntuació total, per cada alumne es recull la puntuació de cadascuna de les parts sobre un puntuació de 10 punts. Es vol estudiar la diferència de mitjanes entre la puntuació obtinguda a la primera (P1) i a la tercera part (P3).

6.- Justifiqueu que s'ha fet un disseny amb dades aparellades. (1 punt)

7.- Calculeu l'IC de la diferència de mitjanes (D) entre la puntuació obtinguda a la primera (P1) i a la tercera part (P3) amb $D=P3-P1$ i una confiança del 95% (1,5 punts)

```
> sum(D)      [1] -79.25      > sum(D^2)      [1] 359.6875
```

8. Indica quin gràfic hi ha a continuació i interpreteu-ho en el context de l'estudi amb disseny amb dades aparellades realitzat (1 punt)



qt(0.950,39)=1.6849	qt(0.975,39)= 2.0227	qf(0.975,39,39) =1.8907	qf(0.025,39,39) =0.5289
qt(0.950,40)=1.6839	qt(0.975,40)=2.0211	qf(0.975,40,40) = 1.8752	qf(0.025,40,40) = 0.5332
qt(0.950,41)=1.6829	qt(0.975,41)=2.0195	qf(0.975,41,41) = 1.8604	qf(0.025,41,41) = 0.5375

Problema D

Un equip està treballant en un estudi per relacionar el rendiment acadèmic amb el mode de transport utilitzat (T) per arribar a la universitat. Demanant a companys si volen respondre l'enquesta, obtenen 38 participants que declaren arribar amb transport privat (T='priv') i 54 que utilitzen transport públic (T='pub'). Tots aporten el temps mitjà per trajecte des de casa (M, en minuts), i la nota mitjana de les assignatures cursades l'any passat (Q, entre 0 i 10).

Els estudiants de l'equip fan una sèrie d'anàlisis amb R, que teniu aquí (DF és el data.frame que conté les dades):

```
t.test(M~T,DF,var.equal=TRUE)

t = 10.783, df = 90, p-value < 2.2e-16
alternative hypothesis: true difference in means between group pub and group priv is not equal to 0
95 percent confidence interval:
 20.30471 29.47696

sample estimates:
mean in group pub mean in group priv
      56.25926      31.36842
```

Quadre 1

	Estimate	Std. Error	t value	Pr(> t)			
(Intercept)	4.09422	0.49832	8.216	x x x x	[a]	[b]	
M	0.02902	0.01021	2.841	x x x x	[c]	[d]	[e]

Residual standard error: 1.599 on 90 degrees of freedom

Multiple R-squared: 0.08232, Adjusted R-squared: 0.07212

F-statistic: 8.073 on 1 and 90 DF, p-value: 0.005557

Quadre 2

	Estimate	Std. Error	t value	Pr(> t)			
(Intercept)	5.851141	0.873439	6.699	x x x x	[h]	[i]	
M	0.001661	0.015061	0.110	x x x x	[j]	[k]	[l]
Tpriv	-1.208140	0.499328	-2.420	x x x x	[m]	[n]	[o]

Residual standard error: 1.558 on 89 degrees of freedom

Multiple R-squared: 0.139, Adjusted R-squared: 0.1196

F-statistic: 7.181 on 2 and 89 DF, p-value: 0.001284

Quadre 3

- 1) Primer, escriviu una proposta explícita de l'objectiu formal de l'estudi, assenyalant quin és el paper de cadascuna de les variables. Quines són quantitatives, i quines són categòriques? (1pt)
- 2) És un estudi experimental o observacional? Si és experimental, digueu com s'han assignat les unitats a la intervenció. Si no ho és, justifiqueu la vostra opció. (1pt)
- 3) Mireu el Quadre 1: expliqueu l'estudi estadístic realitzat i interpreteu els resultats obtinguts. (1pt)

- 4) Quins són els dos models estadístics darrera els resultats presents als quadres 2 i 3? (0.5 pts)
- 5) Quin és el significat dels estimadors que surten a la línia “Intercept” ([a] i [b] / [h] i [i])? (1.5 pts)
- 6) Comenteu els indicadors de la línia “M” ([c], [d] i [e] / [j], [k] i [l]). Expliqueu les implicacions que se’n deriven, amb referència al grau d’evidència estadística que poden aportar. (1.5 pts)
- 7) Quina informació porta la línia del Quadre 3 “Tpriv”. Segons aquest resultat, quina interpretació fem en el àmbit de l’estudi? (1.5 pts)
- 8) Expliqueu breument la variació que s’ha produït entre els indicadors [f] i [p], i entre [g] i [q]. Perquè l’un ha disminuït i l’altre ha augmentat? (1 pt)
- 9) Escriviu una breu conclusió global: què creiem que està relacionat amb el rendiment escolar i quins arguments (estadístics) tenim? (1 pt)