

NOM: _____ COGNOM: _____

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

Problema 1 (Bloc C)

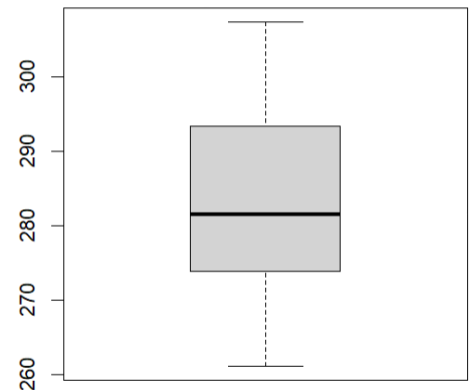
La botiga FRUITES SELECTES ha decidit oferir els seus serveis per internet amb el lliurament de comandes a domicili. Per facilitar la informació a l'usuari, cada producte tindrà a la web una indicació del pes esperat de cada fruita. Un dels productes estrella és la Poma Fuji extra i, per a la posada en marxa de la iniciativa, els propietaris volen fer un estudi per trobar el pes que hauran d'indicar en aquest cas. Per a fer-ho trien una mostra de 30 pomes a l'atzar i en recullen el pes en grams:

$$\sum_{i=1}^{30} x_i = 8484.7835 \quad \text{i} \quad \sum_{i=1}^{30} x_i^2 = 2404449.54$$

1. A partir de les dades anteriors, doneu una estimació puntual de la mitjana del pes (\bar{X}) en grams de les pomes Fuji extra. Doneu també l'error tipus d'aquesta estimació. (1 punt)

2. A la dreta s'adjunta el boxplot de les dades de la mostra. Empleneu la següent taula de manera aproximada. (1 punt)

	Mínim	1r Quartil	Mediana	3r Quartil	Màxim
Pes pomes (g)					



3. Calculeu un interval de confiança al 95% per a la mitjana del pes de les pomes Fuji extra i interpreta el resultat tenint en compte que s'està valorant publicar en el web que el pes mitjà d'una poma Fuji extra és de 280 g (1.5 punts)

4. Es vol també estudiar quina és la variància de les pomes Fuji. Calcula un interval de confiança al 95% per a la variància i interpreta'n el resultat. (1,5 punts)

Com a informació per al consumidor, es vol estudiar si les pomes Golden tenen el mateix pes esperat que les pomes Fuji extra. Per això s'agafa una mostra a l'atzar de 30 pomes Golden i es vol estudiar la diferència de mitjanes entre el pes de pomes Fuji i el pes de les pomes Golden. Les dades per les pomes Golden (G) són les següents:

$$\sum_{i=1}^{30} g_i = 7224.7357 \quad \text{i} \quad \sum_{i=1}^{30} g_i^2 = 1746268.3350$$

5.- Indiqueu de manera justificada si s'ha fet un disseny amb dades independents o amb dades aparellades (0.5 punts)

6.- Calculeu un IC de la diferència de mitjanes amb una confiança del 95% i interpreteu-ne el significat. (2 punts)

Per les pomes Golden dels Alps es fa un tercer estudi. Es tria també una mostra de 30 pomes Golden dels Alps per estudiar si la diferència de mitjanes amb les pomes Fuji i s'ajuden de R per fer-ne l'anàlisi estadístic d'on s'obté el següent output:

```
Two Sample t-test
data: Fuji and GoldenA
t = -0.0374438297, df = 58, p-value = 0.970259708
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-8.07587351713 7.77928920020
```

7. Indica'n l'IC i interpreteu-ne el resultat (1 punt)

8.- Per acabar d'engegar la venda per internet es vol també estudiar la proporció de pomes *naturalment imperfectes* segons el tipus. Per començar, s'ha fet un estudi escollint un mostra de 50 pomes Fuji, entre les quals se n'han trobat 3 de *naturalment imperfectes*. Per altra banda, entre les 60 pomes Golden se n'han comptat 6 de *naturalment imperfectes*. Calculeu l'interval de confiança per a la diferència de les proporcions amb el **90 %** de confiança i interpreteu-ne el resultat. (1.5 punts)

Valors útils pels blocs C i D:

qt(0.975,28)= 2.048	qt(0.975,55)= 2.004	qt(0.975,98)= 1.985	qchisq(0.025,28)= 15.308	qnorm(0,85)= 1,036
qt(0.975,29)= 2.045	qt(0.975,56)= 2.003	qt(0.975,99)= 1.984	qchisq(0.025,29)= 16.047	qnorm(0,9)= 1,036
qt(0.975,30)= 2.042	qt(0.975,57)= 2.003	qt(0.975,100)= 1.984	qchisq(0.025,30)= 16.791	qnorm(0,95)= 1,645
qt(0.975,31)= 2.040	qt(0.975,58)= 2.002	qt(0.975,198)= 1.972	qchisq(0.975,28)= 44.461	qnorm(0,975)= 1,960
qt(0.975,32)= 2.037	qt(0.975,59)= 2.001	qt(0.975,199)= 1.972	qchisq(0.975,29)= 45.722	qnorm(0,99)= 2,326
qt(0.975,33)= 2.035	qt(0.975,60)= 2.000	qt(0.975,200)= 1.972	qchisq(0.975,30)= 46.979	qnorm(0,995)= 2,576

NOM: _____ COGNOM: _____

Problema 2. Bloc D

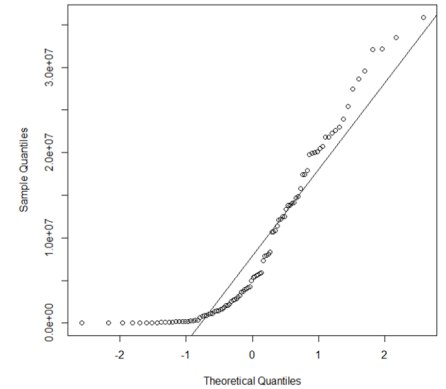
Tots els apartats valen 1 punt excepte 4 i 7 que valen 1.5pts.

L'algorisme estàndard per multiplicar matrius té un cost de $O(M^3)$, on M és el nombre de files/columnes. A la dècada dels 1960, Volker Strassen va presentar un algorisme amb cost $O(M^{\log_2 7})$. Uns estudiants volen comparar empíricament els dos algorismes. Han generat matrius quadrades de mida aleatòria M entre 256 i 2048, amb valors obtinguts de la instrucció de R `runif()`, i han multiplicat cada matriu per sí mateixa amb els 2 algorismes. Han repetit el procediment $n=100$ vegades. Han pres el temps, per cada algorisme (estàndard: "ntime", Strassen: "stime"), en cada producte de matrius en microsegons (10^{-6} s).

1. Quins són els elements, o unitats, de la població objectiu de l'estudi? Com són les observacions de la mostra? Podem considerar que el requisit de mostra a l'atzar es compleix, i perquè sí, o perquè no?

2. Els estudiants han fet el gràfic de la dreta utilitzant la diferència `ntime-stime`.

- A) Expliqueu què han fet, què pretenen comprovar i amb quin objectiu.
- B) Interpreteu el gràfic.
- C) Dibuixeu com seria l'histograma d'aquestes dades.

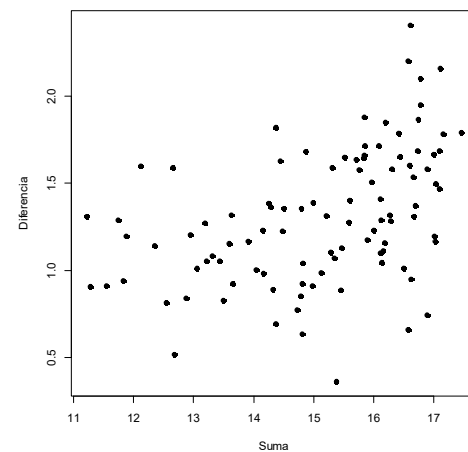


3. En el context de l'estudi anterior, comenteu les diferències entre tenir un efecte en termes additius o en termes multiplicatius.

4. A partir de la següent sortida de R indiqueu quin és el paràmetre analitzat, quant val la seva estimació i el seu error o incertesa (i construïu un IC al 95% de confiança), i treieu-ne la conclusió

```
lm(formula = log(ntime) - log(stime) ~ 1)
Residuals:
Min      1Q  Median 3Q      Max
-0.409 -0.182  0.025  0.154  0.435
Coefficient:  Estimate Std. Error  t value
(Intercept)    0.50192    0.02182    23.00
Residual standard error: 0.2182 on 99 degrees of freedom
```

5. Interpreteu el grafic Bland-Altman de les diferències “ $\log(\text{ntime}) - \log(\text{stime})$ ” en front de $\text{Suma} = “[\log(\text{ntime}) + \log(\text{stime})]/2”$.



6. A partir de la següent sortida de R, digueu quin és el model estadístic que volem aplicar; expliciteu quines són les variables que hi intervenen i amb quin paper cadascuna d’elles; i digueu també quins són i què signifiquen els paràmetres del model. *No contesteu aquí amb resultats numèrics (llegiu la següent pregunta)*

```
lm(formula = (log(ntime) - log(stime)) ~ log(M))
Residuals:
  Min       1Q   Median       3Q      Max
-0.9694 -0.2346  0.0038  0.2287  0.955
Coefficients: Estimate Std. Error t value
(Intercept) -0.85608    0.43936   -1.948
log(M)       0.31271    0.06331    4.940
Residual standard error: 0.3486 on 98 degrees of freedom
Multiple R-squared: 0.1993, Adjusted R-squared: 0.1912
```

7. A partir dels resultats de l’anterior sortida de R, indiqueu els valors estimats pels paràmetres, amb les respectives mesures d’incertesa, i interpreteu.

8. Interpreteu el R^2 del model de l’anterior sortida de R.

9. Utilitzeu el model de l’anterior sortida de R per obtenir una predicció per a multiplicar matrius de mida 1000 i de mida 2000, i expliqueu què heu obtingut.