

NOM: _____ COGNOM: _____

Problema Bloc C

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

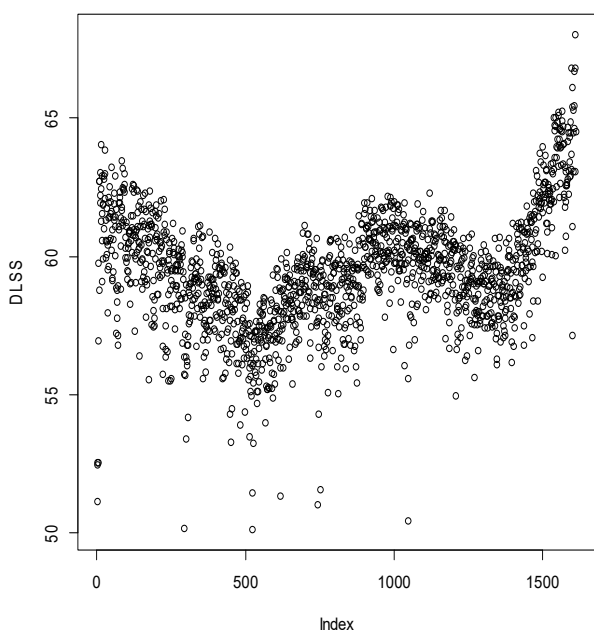
Solució esperada o plantilla. I explicació docent.

Uns companys han comparat el rendiment d'un videojoc quan utilitzem una targeta RTX de NVIDIA amb tecnologia DLSS, i quan no s'utilitza, i han recollit la resposta FPS (fotogrames renderitzats per segon [fps]).

	Min.	1r Qu.	Mediana	Media	3er Qu.	Max
Sense DLSS	25.42	29.91	32.07	32.24	34.08	42.91
Amb DLSS	50.11	58.57	60.03	62.34	63.08	80.40

1) Trobeu a faltar alguna informació rellevant amb aquesta primera descriptiva? Quina? **(0.5 punts)**

El nombre de casos per grup. [R no ho dona perquè, en un estudi ben planificat i executat, no és un resultat, sinó una decisió dels investigadors. Però en general, és la primera informació que volem saber: de quantes dades disposem? Mal si SE ja que no és descriptiva, sinó inferència. Mal si SD. Ja que hi dona els quartils i es pot obtenir el IQS, mesura robusta alternativa a la SD.]



2) Aquest gràfic mostra els valors de FPS pel grup amb DLSS en funció de la variable 'index' —que informa de l'ordre de la recollida de les dades. Raoneu perquè s'aprecia certa variabilitat en FPS (és a dir, perquè obtenim valors diferents). Discutiu també si aquesta seqüència de valors de la resposta us sembla adequada per l'estudi i anàlisi posterior. **(1 punt)**

És habitual trobar certa variabilitat pròpia del sistema: per exemple, el ordinador pot estar més o menys carregat. No necessàriament han de ser errors al prendre les mesures.

[Quines són les fonts de variabilitat és una pregunta molt interessant que sempre us heu de fer: algunes són degudes a una recollida deficient i pobre, altres són inherents a les dades, inevitables i no tenen perquè ser dolentes.]

En aquest gràfic sembla existir un patró temporal que diu que sí, FPS varia segons l'ordre de recollida. Alerta d'una recollida deficient ja que les dades no són independents, primer requisit estadístic: les dades han de ser variables aleatòries independent i idènticament distribuïdes, v.a.i.i.d. [Pot ser no han deixat descansar el sistema,

deixant un temps entre execucions]

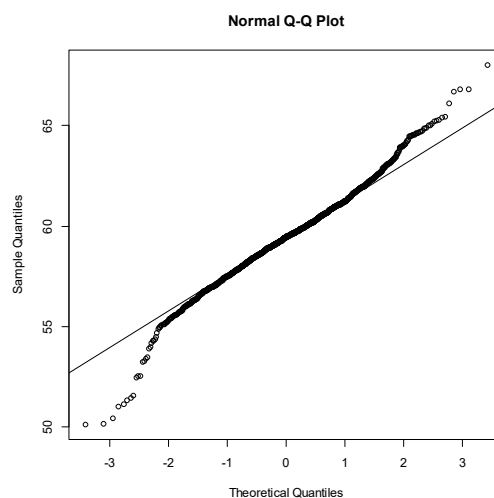
3) Inventeu un valor raonable de la desviació tipus pels FPS cas DLSS i interpreteu. **(0.5 punts)**

[la mitjana estaria al voltant de 60 fps (62 segons la descriptiva anterior) i] la desviació tipus estaria al voltant de 2. Vol dir que la distancia mitjana amb el centre val aproximadament 2 fps. [Exactament, en lloc de la distancia mitjana, és l'arrel quadrada de la mitjana de les distàncies quadrades amb la mitjana.]

4) Interpreteu el QQPLOT de la variable FPS amb DLSS (dreta). **(0.5 punts)**

Els extrems mostren que les cues no ajusten bé al model normal (cues massa pesades amb molts casos).

[Hi ha dades extremes, massa distants. Convindria estudiar perquè.]



Altres companys han avaluat quant més lent és un llenguatge interpretat com Python (P) respecte a un de compilat de C++ (C). Han generat 100 vectors de entre 30.000 i 50.000 elements i els han ordenat amb els dos llenguatges. Finalment han mesurat els respectius temps, i els han transformat, amb el logaritme base 2, creant les variables ℓ_2P i ℓ_2C . La correlació entre les dues mesures val $\text{Corr}[\ell_2P, \ell_2C]=0.9$.

5) Expliqueu si la descripció anterior correspon a un disseny aparellat o de dues mostres independents, i quin és l'avantatge estadístic que ofereix el tipus de disseny aplicat **(1.5 punt)**

La variància de la diferència com si fossin independents: $V(I_2P - I_2C) = V(I_2P) + V(I_2C)$

I si fossin aparellades: $V(I_2P - I_2C) = V(I_2P) + V(I_2C) - 2 \text{Cov}(I_2P, I_2C) = V(I_2P) + V(I_2C) - 2 \text{Corr} \cdot \text{SD}(I_2P) \cdot \text{SD}(I_2C)$

Com la correlació és positiva, reduïrien la variabilitat de la diferència i també de l'error estàndard de la estimació de la diferència mitjana.

→ Les comparacions de les mitjanes serà més precises, tindran menys incertesa.

```
> D <- log2py - log2cmmas  
> summary(lm(D~1))
```

```
Call:  
lm(formula = D ~ 1)
```

```
Residuals:  
    Min       1Q   Median       3Q      Max  
-0.43852 -0.03019 -0.01762  0.00004  0.97831
```

```
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept)    6.3405     0.0132   480.3  <2e-16 ***  
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error on 99 degrees of freedom
```

A la esquerra teniu el codi i els resultats de R de les dades logotransformades [log base 2].

6) Estimeu puntualment l'avantatge de C respecte a P en termes de log-2, desfeu la transformació i interpreteu **(1.5 punt)**

En termes log-2, l'avantatge de C és 6.34 amb un error estàndard de 0.01.

Podem desfer la transformació per la mitjana de les diferències i trobar la mitjana geomètrica del ratio P/C:

$2^{6.34} \approx 81$

C es 81 vegades més ràpid.

7) Estimeu per interval del 95% l'avantatge de C respecte a P en termes de log-2 i desfeu la transformació **(1 punt)**

$\text{mean}(D) \pm t_{99,0.975} \cdot \text{sd}(D) / \sqrt{n} = \text{mean}(D) \pm t_{99,0.975} \cdot \text{s.e.} \approx [6.314, 6.367]$

$\text{Exp}_{\text{base}2}(6.367) \approx 82,5$

$\text{Exp}_{\text{base}2}(6.314) \approx 79,6$

[OK si aproximeu $t_{99,0.975} \approx 1.98$ per la normal ≈ 1.96]

8) Trobeu el valor que hem eliminat ("residual standard error"). Expliqueu en què es diferencien aquest indicador i el anomenat "Std. Error" (què vol dir cadascun?). **(1.5 punt)**

$\text{Std.error} = 0.0132 = s_D / \sqrt{n}$, perquè el model (el més simple de tots, no conté variables explicatives) és només estimar la mitjana de D. Llavors, $s_D = \sqrt{100} \times 0.0132 = 0.132$

El *residual standard error* en aquest cas és la desviació tipus de D, i mesura la dispersió al voltant del valor mitjà [que depèn molt de les particularitats del vector ordenat]. El *Std. Error* és l'error tipus de la mitjana de D, i estima la variabilitat de la mitjana mostral a una mostra de 100 elements com aquesta [les variacions de la mitjana són degudes a l'atzar, dels elements concrets seleccionats per formar part de la mostra].

9) Interpreteu els resultats. **(1 punt)**

Amb una confiança del 95%, C triga entre 79,6 i 82,5 vegades menys.

És un interval prou estret que sembla suficient.

10) Heu trobat un article vell que reproduïx un estudi com el vostre, però els autors comparen Python amb C utilitzant dues mostres, de grandàries respectives 45 i 55. El article no inclou l'anàlisi estadística, només les mitjanes (10 i 3.5, en termes de logaritme en base 2) i les desviacions tipus (3.2 i 1.3). Es demana que calculeu el quocient senyal/soroll amb les dades de l'article i que comenteu sobre el resultat obtingut. **(1 punt)**

La *s pooled* val 2.351 (encara que les desviacions semblen massa diferents entre Python i C per obtenir un valor comú).

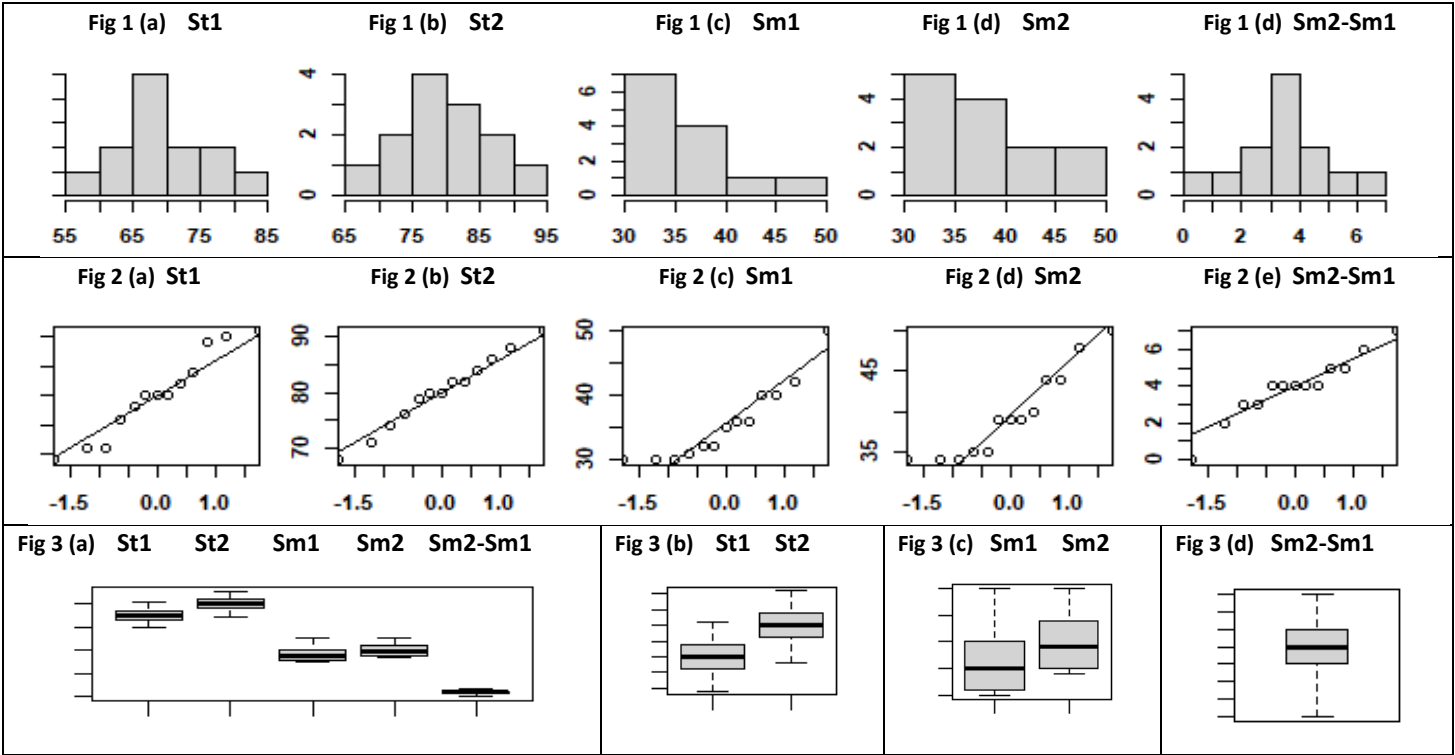
L'indicador *t* (senyal/soroll) val $(10 - 3.5) / 2.351 \sqrt{1/45 + 1/55} = 13.75$

És a dir: la diferència entre mitjanes (6.5) és quasi 14 vegades més gran que l'error mostral de la diferència de mitjanes mostrals (0.47)

Ja sabíem que C és molt més ràpid que Python, per tant aquest resultat no és sorprenent. Si obtenim l'interval de confiança d'aquestes dades, seria aproximadament 6.5 ± 1 que, comparat amb el nostre disseny aparellat, és bastant menys precís.

Problema D

Per comparar 2 aparells (1 i 2) d'extreure Suc a taronges tenim m.a.s. dels grams extrets en taronges senceres i partides: St1 i St2 són els grams en 26 taronges senceres (13 en cada aparell); i Sm1 i Sm2 en 13 taronges partides (una meitat a cada aparell). També es té el pes de les 26 taronges no partides. Alguns dels resultats per St1, St2, Sm1, Sm2, i Sm2-Sm1 són:
mean(St1)=70.1 sd(St1)=7.2 mean(St2)=80.1 sd(St2)=6.6 mean(Sm1)=35.7 sd(Sm1)=6.0 mean(Sm2)=39.6 sd(Sm2)=5.4 mean(Sm2-Sm1)=3.9



I resultats pels grams i també el pes de les 26 taronges senceres:



I resultats per diversos models a partir de les dades recollides:

MODEL A	
<code>lm(Y~1) # Y és Sm2-Sm1</code>	
Coefficients:	
Estimate Std. Error t value Pr(> t)	
(Intercept) 3.9231 0.4865 8.064 3.47e-06 ***	
Residual standard error: 1.754 on 12 degrees of freedom	
MODEL B	
<code>lm(Y~Aparell) # Y és St1 i St2 junts. "Aparell" és 1 o 2 segons l'utilitzat per extreure</code>	
Coefficients:	
Estimate Std. Error t value Pr(> t)	
(Intercept) 70.077 1.918 36.535 <2e-16 ***	
Aparell2 10.000 2.713 3.687 0.00116 **	
Residual standard error: 6.916 on 24 degrees of freedom	
Multiple R-squared: 0.3615, Adjusted R-squared: 0.3349	
MODEL C	
<code>lm(Y~Pes) # Y és St1 i St2 junts</code>	
Coefficients:	
Estimate Std. Error t value Pr(> t)	
(Intercept) 15.01486 4.41767 3.399 0.00236 **	
Pes 0.27192 0.01983 13.711 7.55e-13 ***	
Residual standard error: 2.912 on 24 degrees of freedom	
Multiple R-squared: 0.8868, Adjusted R-squared: 0.8821	
<code>predict(modelC,data.frame("Pes"=250),interval="prediction"):</code>	<code>fit=82.99386 lwr=76.75405 upr=89.23367</code>
<code>predict(modelC,data.frame("Pes"=250),interval="confidence"):</code>	<code>fit=82.99386 lwr=81.31765 upr=84.67007</code>

Contesteu les següent preguntes, justificant la resposta (si cal indicant en quina figura o model dels resultats anteriors us baseu).

1.- (1 punt) Indiqueu i justifiqueu si la comparació de mitjanes de grams de suc extret en taronges senceres és un estudi de mostres aparellades o independents. I també quan la comparació és de mitjanes de grams de suc extret en mitges taronges.

Les dades de suc extret de taronges senceres són independents ja que són 13 taronges diferents que es passen per cada aparell.

Les dades de suc de taronges partides són aparellades ja que cada taronja es parteix i cadascuna de les dues meitats es passen per un dels aparells.

2.- (1.5 punts) Per concloure si els dos aparells extreuen una quantitat de suc comparable o no, indiqueu i quantifiqueu la diferència mitjana de grams de suc extrets (amb el seu error tipus) pel cas de taronges senceres i pel cas de mitges taronges. Interpreteu i justifiqueu en quin dels resultats donats baseu la resposta.

En el model de mostres independents en taronges senceres (resultats del MODEL B):

10 gr és l'estimació del canvi de **més suc extret en mitjana per l'aparell 2** (canvi a Aparell2 enlloc del 1)

2.713 gr és l'error tipus d'aquesta estimació

En el model de mostres aparellades en mitges taronges (resultats del MODEL A):

3.9 gr és l'estimació del canvi de **més suc extret en mitjana per l'aparell 2** (la diferència és de l'aparell 2 menys el 1)

0.4865 gr és l'error tipus d'aquesta estimació

3.- (1.5 punts) Indiqueu —tant pel cas del model aplicat a mitges taronges com a senceres— les premisses i si es compleixen o no justificant-ho segons els resultats i figures donades.

En el model de mostres independents en taronges senceres, una primera premissa és que St1 i St2 siguin **mostres aleatòries simples** (m.a.s. de l'enunciat). També que compleixin **normalitat**, que es veu en les figures 1(a) i 1(b) amb histogrames força semblants al model Normal, i en les figures 2(a) i 2(b) amb qqnorm força alineats seguint el model normal. I també les dues mostres independents tenen una **variabilitat semblant** i força simetria (valors de sd i gràfics 3(a) i 3(b))

En el model de mostres aparellades en mitges taronges, una primera premissa és que la mostra de 13 taronges que partim sigui **aleatòria simple** (m.a.s. de l'enunciat). I també que compleixi **normalitat**, que es veu en la figura 1(e) amb histograma força semblant al model Normal reforçat en la figura 3(d) amb boxplot molt simètric, i en les figures 2(e) amb qqnorm força alineat. Les premisses cal que es **compleixen en la diferència i no en cadascuna**

4.- (1 punt) Comenteu globalment la comparativa entre els dos aparells d'extreure suc de taronges segons les dues proves. Indiqueu i justifiqueu, en quina de les 2 proves, l'interval de confiança de la mitjana seria més precís.

Tant la prova aparellada com la independent compleixen les premisses, i mostren uns resultats de **més extracció de suc usant l'aparell 2**. En el cas de taronges s'extreu entorn uns 10 gr més de suc, i en taronges partides entorn uns 4 gr més.

L'IC de la prova aparellada seria més precís. L'amplada és $2 * Std_error * qt()$, i Std_error o **error tipus en aparellada és molt menor**

($0.4865 < 2.713$), tot i que el valor de qt és una mica major per menys graus de llibertat ($qt(0.975,12) > qt(0.975,24)$) ja que en un cas són 13 taronges partides i en l'altre 26 taronges senceres

Per altra part, a partir del model de relació lineal entre els grams de suc extrets i el pes de les taronges senceres:

5.- (1 punt) Indiqueu quina és l'equació de la recta que relaciona els grams de suc extrets amb el pes de la taronja. Interpreteu els coeficients de la recta i justifiqueu en quin resultat anterior baseu la resposta.

La recta és (resultats a MODEL C com a estimadors de l'intercept o ordenada a l'origen i del coeficient que afecta al Pes):

Suc = 15.01486 + 0.27192 * Pes

El coeficient del terme independent ($b_0=15.015$ gr) indicaria els grams extrets quan Pes=0, tot i que en aquest cas no té sentit interpretar la recta per aquests valors de Pes molt allunyats del rang de pesos usats per ajustar-la.

El coeficient del pendent de la recta ($b_1=0.27$ gr) indica que per cada gram més de pes, s'extreuen uns 0.27 gr més de suc.

6.- (1.5 punt) Per quantificar quant bo és aquest model per fer prediccions, indiqueu quin és el coeficient de determinació del model, i interpreteu-lo relacionant-lo amb el càlcul de la correlació entre pes i grams extrets i justifiqueu en quin resultat anterior baseu la resposta. I també indiqueu la desviació residual del model

$R^2 = 88.68 \%$ És el coeficient de determinació (a MODEL C Multiple R-squared)

i indica que quasi un 90% de la variabilitat dels grams de suc extrets s'expliquen pel pes.

$\sqrt{0.8868} = 0.94$ (+0.94 ja que el pendent de la recta és positiu a la figura 4(a)) és la correlació entre Pes i Suc

i indica la forta relació positiva entre els grams de Pes i de Suc

2.912 és la desviació residual del model (a MODEL C Residual Standard Error)

7.- (1 punt) Calculeu, i indiqueu-ne el càlcul, una predicció del suc extret per una taronja de 250 grams. I doneu uns intervals, amb limit inferior i superior al 95% de confiança, per a la predicció com a puntual o com a valor esperat o mitjana, i compareu-los

Suc_250 = 15.01486 + 0.27192 * 250 = 82.99 grams

[76.75405, 89.23367] és l'interval per a la predicció puntual (lwr i upr a MODEL C predict amb interval="prediction")

[81.31765, 84.67007] és l'interval per a la predicció puntual (lwr i upr a MODEL C predict amb interval="confidende")

En el cas de l'interval per a la predicció com a valor puntual és més ample, ja que hi ha més incertesa, comparat amb la predicció com a valor esperat o mitjana que té menys error per calcular l'interval

8.- (1.5 punts) Per aquest model, indiqueu quines són les premisses i si es compleixen o no justificant-ho segons els resultats i figures anteriors.

En el model lineal entre Pes i Suc, unes primeres premisses són que es tracti d'una **mostra aleatòria simple** (m.a.s. de l'enunciat) i que hi hagi **linealitat** entre el Pes i el Suc (comprovable a al figura 4(a) on els punts s'ajusten força al llarg de la recta)

Per altra part cal que es compleixin les premisses de normalitat, homocedasticitat i independència. La **normalitat** es pot comprovar a la figura 4(b) en l'histograma força simètric i forma de campana, i en el Normal QQ amb força alineació. En els dos altres gràfics de Residuals de la figura 4(b) es pot comprovar la **homocedasticitat** (en els residus no hi ha zones indicant més proximitat a la recta i altres allunyats, com també es veu a la figura 4(a), per tant la variabilitat al llarg de la recta és homogènia). I també la **independència** doncs no hi ha cap patró entre els valors de residus que n'indiqués alguna dependència.