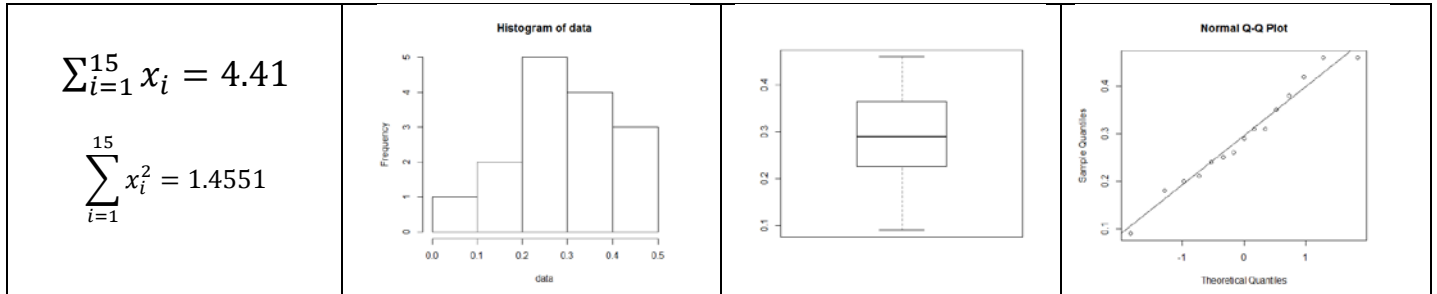


Problema 1 (B4)

Un alumne de la FIB està estudiant com detectar els accessos fraudulents a comptes protegits per contrasenya. El primer mètode que està provant consisteix en detectar els accessos fraudulents a partir de mesurar el temps entre les pulsacions de les tecles a l'hora d'introduir l'usuari i la contrasenya.

En un cas sospitós ha recollit els temps en segons entre pulsacions de les tecles i ha obtingut les següents dades:



1. Calculeu les estimacions puntuals del temps mitjà i de la desviació tipus (1 punt)
2. A partir dels gràfics argumenteu si podem suposar que el temps entre pulsacions segueix una distribució normal. (1 punt)

Per estudis previs, se sap que una persona real utilitza 0.3653 segons de temps mitjà entre pulsacions per introduir les claus d'accés.

3. Per estudiar si el cas sospitós és un possible accés fraudulent es vol contrastar si el temps mitjà entre pulsacions és 0.3653 s (mitjana per una persona real) o no amb un risc de l'1%.
Indiqueu:
 - a) Les hipòtesis, les premisses, la fórmula de l'estadístic i quina és la distribució d'aquest sota la hipòtesi nul·la (1 punt)

b) Calculeu el valor de l'estadístic (0.5 punts)

c) Representeu gràficament els punts crítics, les zones d'acceptació i de rebuig i el valor de l'estadístic (1 punt)

d) A partir de l'estudi i del càlculs realitzats, interpreteu els resultats de la prova d'hipòtesi aplicada sobre el cas sospitós de ser fraudulent. (0.5 punts)

4. a) Calculeu un interval bilateral amb 90% de confiança per la mitjana de temps entre pulsacions en l'usuari fraudulent (2 punts)

b) A partir de l'interval de confiança obtingut, interpreteu els resultats sobre el cas sospitós de ser fraudulent. (0.5 punts)

5. Compareu els resultats dels apartats 3 i 4 (0.5 punts)

6. Un segon mètode que volen explorar és detectar els accessos fraudulents a partir del temps total en introduir l'usuari i la contrasenya. Experimentalment l'alumne de la FIB ha establert que els accessos fraudulents triguen més de 5s en accedir. Dels darrers 80 accessos, 20 han trigat més de 5s en accedir.

a) Calculeu un IC 95% per a aquesta proporció d'accessos que triguen més de 5s en accedir (1.5 punts)

b) Interpreteu el resultat obtingut en l'apartat anterior (0.5 punts)

NOM: _____ COGNOM: _____

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

Problema 2 (B5)

Al laboratori de càlcul on treballem es monitoritza la temperatura dels servidors allotjats. S'està estudiant com es comporta la temperatura a 10 dels ordinadors seleccionats a l'atzar, concretament el canvi després de dotze hores de funcionament. S'ha observat una variació de 46.3 a 52.2 graus en la temperatura mitjana; les diferències entre temperatura inicial i final presenten una desviació tipus de 9 graus.

Per altra banda, sabem que 4 dels 10 servidors escollits disposen d'un sistema de refrigeració instal·lat recentment: en aquests l'increment mitjà de la temperatura ha estat de 3.2 graus.

1. Determina a partir de la informació disponible en quant ha augmentat en mitjana la temperatura als altres sis servidors. (1pt)
2. Considerant la mostra del 10 servidors, creieu que es pot dir que després de dotze hores ha augmentat la temperatura mitjana dels servidors? Justifiqueu la resposta i estimeu amb interval de confiança del 95% l'increment esperat de la temperatura. (2pt)
3. La desviació tipus de l'increment de temperatura al grup de 4 servidors ha estat de 5 graus, i de 7.5 graus a l'altre grup. Justifiqueu formalment si hi ha motius per pensar que els servidors amb el sistema de refrigeració més recent tenen menor variabilitat en el canvi de temperatura que els altres servidors. Aclariu les premisses necessàries per la prova que utilitzeu. (1.5pt)
4. Es vol trobar resposta també a la qüestió de si el nou sistema de refrigeració és més efectiu que l'antic. És a dir: l'increment de temperatura és igual en ambdós sistemes, o un permet moderar millor la temperatura? Independentment del que heu trobat a la pregunta anterior, assumiu que els dos sistemes comparteixen una variància comuna en la variable "increment de temperatura". (2pt) [si no heu trobat la solució de l'apartat 1, preneu un valor qualsevol "raonable"]

5. Expliqueu els resultats trobats als apartats anteriors, segons l'evidència que proporciona la mostra disponible. (1.5pt)

6. El fabricant dels sistemes de refrigeració disposa de molta més informació dels seus clients, i manté una base de dades dels resultats de les proves efectuades, encara que només pot saber si l'increment de temperatura ha estat superior o inferior a 8 graus, després de 12 hores d'observació. En resum, aquesta és la informació recollida:

	Increment de T inferior a 8°	Increment de T superior a 8°
Sistema anterior	270	115
Sistema nou	95	28

Podem dir amb les dades del fabricant si hi ha cap diferència entre els sistemes? Feu la prova basada en l'estadístic chi-quadrat. Justifiqueu i interpreteu la conclusió. (2pt)

Marginals	Increment de $T \leq a 8^\circ$	Increment de $T > a 8^\circ$	
Sistema anterior	270	115	
Sistema nou	95	28	

Valor esperat	Increment de $T \leq a 8^\circ$	Increment de $T > a 8^\circ$
Sistema anterior		
Sistema nou		

NOM: _____ COGNOM: _____

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

Problema 3 (B6)

Un empresa té dos proveïdors principals (A i B) de PCs, amb unes rebaixes del **preu per unitat** segons el nombre de PCs que es compren. Es decideix analitzar el preu per unitat que va aplicar cada proveïdor en 10 comandes de l'any passat a cadascun i que van ser de 1, 10, 15, 20, 25, 30, 35, 40, 45, i 50 unitats.

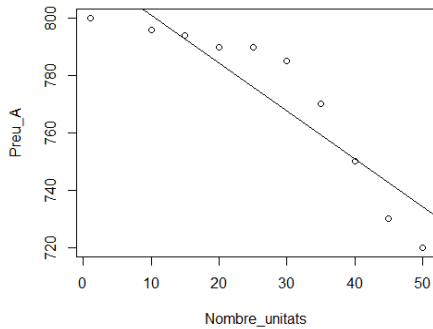
Un equip de l'empresa analitza les dades del proveïdor A i presenta els següents resultats:

Nombre d'unitats comprades en 10 comandes del darrer any: `Nombre_unitats <- c(1, 10, 15, 20, 25, 30, 35, 40, 45, 50)`

Preu unitari aplicat en cadascuna de les 10 comandes indicades: `Preu_A <- c(800,796,794,790,790,785,770,750,730,720)`

	Mitjana	Desviació tipus	Min	Max
Preu A	772.5	29.09	720	800
Nombre unitats	27.1	15.84	1	50

$$\text{cov}(\text{Preu_A}, \text{Nombre_unitats}) = S_{\text{Preu A, Nombre unitats}} = -419.72$$

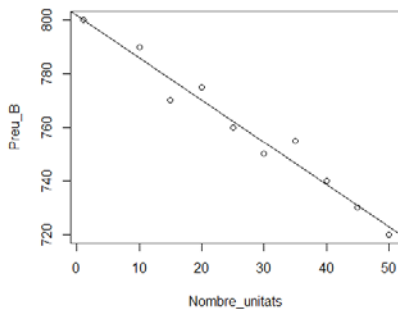


- Calculeu els coeficients de la recta de regressió del Preu_A en funció del Nombre d'unitats. Indiqueu i interpreteu l'equació de la recta ajustada. Calculeu i interpreteu el coeficient de determinació i la desviació residual (2 punts)

Un altre equip de l'empresa analitza les dades del proveïdor B i presenta els següents resultats:

Nombre d'unitats comprades en 10 comandes del darrer any: `Nombre_unitats <- c(1, 10, 15, 20, 25, 30, 35, 40, 45, 50)`

Preu unitari aplicat en cadascuna de les 10 comandes indicades: `Preu_B <- c(800,790,770,775,760,750,755,740,730,720)`

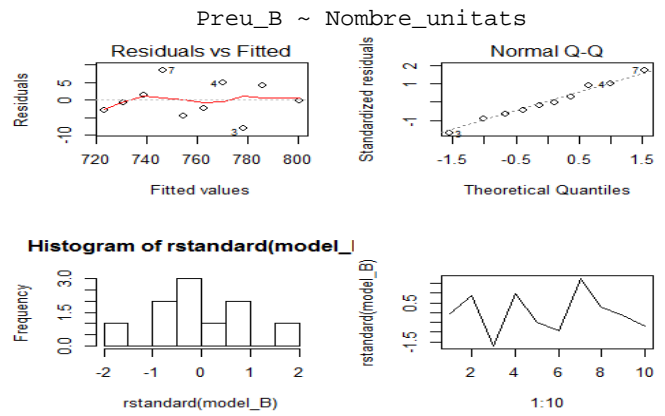
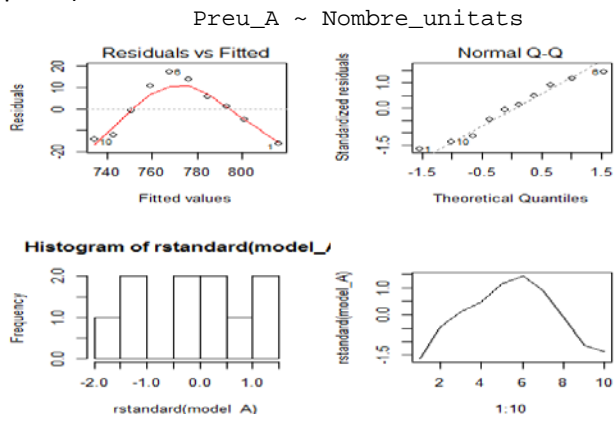


```
lm(formula = Preu_B ~ Nombre_unitats)
Coefficients:
              Estimate      Std. Error    t value      Pr(>|t|)
(Intercept)      _____    3.3566      238.87    < 2e-16 ***
Nombre_unitats  -1.5792      0.1083      _____    4.81e-07 ***
Residual standard error: 5.146 on 8 degrees of freedom
Multiple R-squared:  0.9637
```

- Indiqueu i interpreteu l'equació de la recta ajustada del Preu_B en funció del Nombre d'unitats. Compareu les dues regressions en quant a la recta ajustada, al coeficient de determinació i a la desviació residual (2 punts)

- En el cas de la segona regressió amb el Preu_B, calculeu i interpreteu un IC 95% del pendent de la recta, i plantegeu la prova d'hipòtesis de si la recta és plana o no indicant el valor de l'estadístic i la conclusió (2 punts)

- Enuncieu les premisses de la regressió lineal i compareu-les en els dos models (indiqueu els gràfics d'on es dedueixen) (2 punts)



- Feu una valoració global dels dos models de regressió en quant a quins resultats són semblants i quins no entre els dos proveïdors, i en quant a quin preferirieu i perquè (2 punts)