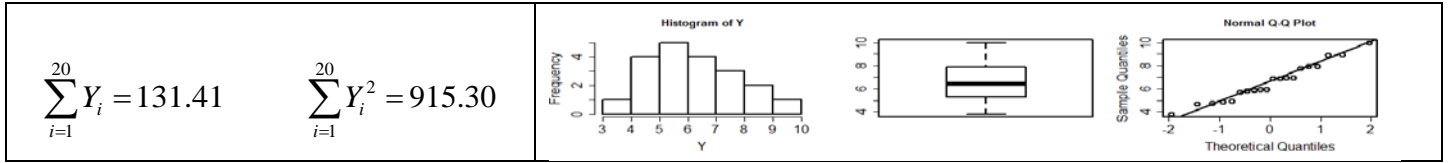


NOM: _____ COGNOM: _____

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

Problema 1 (B4)

En l'estudi del seguiment d'una assignatura al llarg dels quadrimestres i després de canvis aplicats a l'assignatura, els professors volen comprovar si la nota mitjana esperada segueix sent 6 o si hi ha evidència de que ha augmentat. Amb 20 notes obtenen:



(1 punt) **1.-** Primer calculen les estimacions puntuals de la nota mitjana i de la desviació tipus:

(3 punts) **2.-** Llavors posen a prova (amb un risc del 5%) si l'esperança poblacional de la nota és 6 o superior, suposant que la variabilitat poblacional s'ha mantingut i és coneguda (assumeixen desviació poblacional igual a 1.5). Indiqueu:

- (0.5) les hipòtesis, premisses, la fórmula de l'estadístic i dir quina distribució segueix sota la hipòtesis nul·la

- (0.5) càlcul de l'error tipus (o estàndard error) i del valor de l'estadístic

- (0.5) representació gràfica de l'estadístic, el/s punt/s crític/s i les zones d'acceptació i rebuig

- (0.5) càlcul del p-valor

- (1) segons els dos apartats anteriors, a quina conclusió arribeu sobre la prova d'hipòtesi. Interpreteu els resultats

(2 punts) **3.-** Des de la Facultat també es fa la mateixa prova (si l'esperança poblacional de la nota és 6 o superior, amb un risc del 5%) usant les mateixes dades però sense suposar la variabilitat coneguda. Indiqueu:

- (0.5) les hipòtesis, premisses, la fórmula de l'estadístic i dir quina distribució segueix sota la hipòtesis nul·la

- (0.5) càlcul de l'error tipus (o estàndard error) i del valor de l'estadístic

- (0.5) representació gràfica de l'estadístic, el/s punt/s crític/s i les zones d'acceptació i rebuig

- (0.5) en funció de l'apartat anterior, a quina conclusió arribeu sobre la prova d'hipòtesi. Interpreteu els resultats

(1 punt) **4.-** Compareu els apartats **2** i **3**

(3 punts) **5.-** Finalment el que tenim és una variable dicotòmica (A aprovat, o S suspès) amb 6 suspesos i 14 aprovats. Poseu a prova si el valor esperat de la proporció d'aprovats és del 75% o no. Amb un risc del 5% indiqueu:

- (0.5) les hipòtesis i l'estimació puntual de la proporció d'aprovats

- (0.5) la fórmula i el càlcul de l'estadístic

- (0.5) representació gràfica de l'estadístic, el/s punt/s crític/s i les zones d'acceptació i rebuig

- (0.5) càlcul d'un interval de confiança pel valor esperat de la proporció d'aprovats, i interpretació

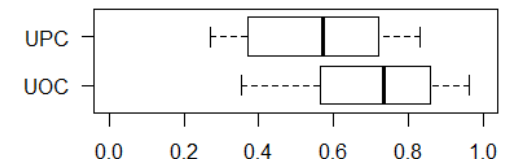
- (1) segons els dos apartats anteriors, a quina conclusió arribeu sobre la prova d'hipòtesi. Interpreteu els resultats

NOM: _____ COGNOM: _____

Problema 2 (B5)

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

1. Les universitats UPC i UOC estan preocupades pel plagi dels treballs acadèmics dels alumnes. Ambdues universitats tenen algoritmes per determinar la probabilitat que un treball hagi estat plagiat en base a fer cerques de fragments dels documents a internet a través de diferents cercadors. La UPC sospita que l'algoritme de la UOC dona millor resultats que el seu propi i decideix comparar les probabilitats que retornen ambdós algoritmes en treballs que es sap amb tota certesa que han estat plagiats. Per fer l'estudi, s'han recollit 10 treballs d'alumnes de la UPC i 21 treballs d'alumnes de la UOC, obtenint la següent descriptiva de resultats respecte a les probabilitats.

<pre>> summary(UPC) Min. 1st Qu. Median Mean 3rd Qu. Max. 0.2696 0.4015 0.5709 0.5548 0.6867 0.8311</pre>	<pre>> sd(UPC) [1] 0.1888975</pre>	
<pre>> summary(UOC) Min. 1st Qu. Median Mean 3rd Qu. Max. 0.3542 0.5654 0.7350 0.7110 0.8583 0.9614</pre>	<pre>> sd(UOC) [1] 0.1796934</pre>	

a) Justifiqueu si es tracta de dades aparellades o independents (1 punt)

b) Plantegeu una prova d'hipòtesis per decidir si la mitjana de les probabilitats de l'algoritme de la UOC és superior o no a la de l'algoritme de la UPC en els treballs plagiats. Digueu explícitament si és un contrast bilateral o unilateral. (1 punt)

c) Indiqueu l'expressió de l'estadístic, la seva distribució sota H_0 i les premisses necessàries (1 punt)

d) Només mirant la descriptiva de l'enunciat, comenteu si us semblen raonables les premisses (1 punt)

e) Calculeu el valor de l'estadístic i el/s punt/s crític/s suposant un $\alpha = 0.05$ (1 punt)

f) Indiqueu, justifiqueu i interpreteu quina és la conclusió de la prova d'hipòtesis (1 punt)

2. Evidentment, a les universitats els interessa un algoritme que classifiqui els treballs plagiats com a tal, però també que classifiqui els treballs no plagiats com a originals. Ambdues universitats determinen que hi ha plagi en un treball quan la probabilitat d'alguns dels seus algorismes respectius és superior a 0.5. A més dels 31 treballs previs, s'escullen dues mostres aleatòries de treballs addicionals per cada universitat dels quals es té la certesa que no han estat plagiats. Es recullen un total de 100 treballs per a cada universitat. Es vol saber si la proporció d'encert global de l'algoritme (en tots els treballs, tant en els plagiats com en els no plagiats) és la mateixa en les dues universitats o no. Els resultats estan a les taules següents:

UPC	Prob≤0.5	Prob>0.5	Total
NO Plagiats	76	14	90
Plagiats	6	4	10
Total	82	18	100

UOC	Prob≤0.5	Prob>0.5	Total
NO Plagiats	53	26	79
Plagiats	4	17	21
Total	57	43	100

a) Indiqueu les hipòtesis i si el contrast ha de ser unilateral o bilateral (1 punt)

b) Indiqueu l'expressió de l'estadístic, la seva distribució sota H_0 i les premisses assumides. (1 punt)

c) Calculeu el valor de l'estadístic i el punt/s crític/s (1 punt)

d) Indiqueu, justifiqueu i interpreteu quina és la conclusió de la prova d'hipòtesis (1 punt)

NOM: _____ COGNOM: _____
 (Contesteu cada pregunta en el seu lloc. Expliqueu i justifiqueu els càlculs.)

Problema 3 (B6)

Es vol estudiar la relació lineal entre els resultats de dos exàmens per aprovar unes oposicions. S'han recollit les notes del primer examen (X) i del segon examen (Y) de 50 aspirants a la plaça, i obtenim els següents resultats:

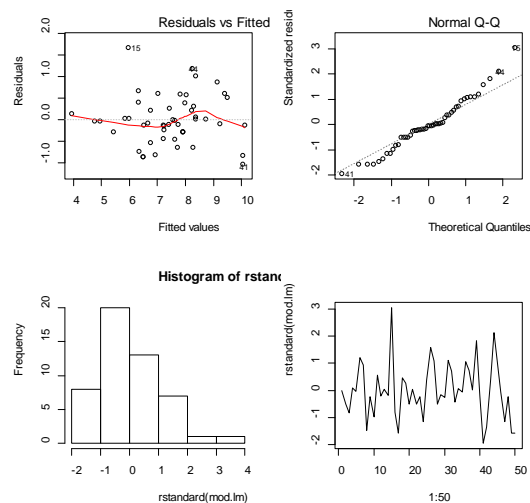
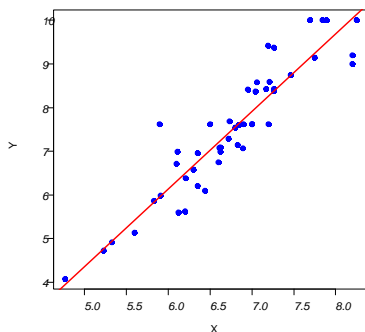
$$\sum x_i = 337.49$$

$$\sum y_i = 373.12$$

$$\sum x_i^2 = 2306.08$$

$$\sum y_i^2 = 2888.158$$

$$\sum x_i y_i = 2568.386$$



Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	---	0.7197	---	---
X	---	0.1060	16.764	< 2e-16

Residual standard error: 0.5616 on 48 degrees of freedom
 Multiple R-squared: ---, Adjusted R-squared: ---
 F-statistic: 281 on 1 and 48 DF, p-value: < 2.2e-16

1.- Calculeu la covariància i la correlació entre X i Y (2 punts).

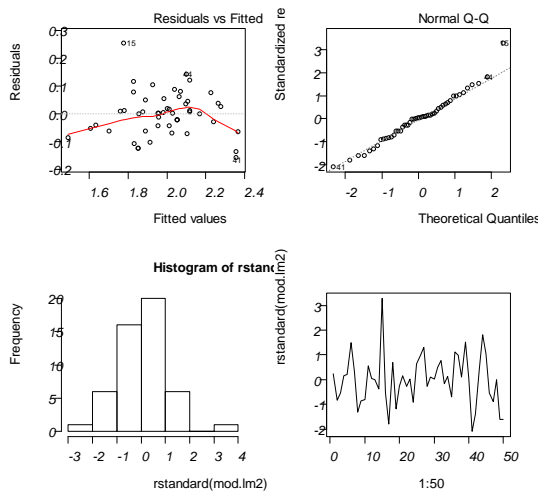
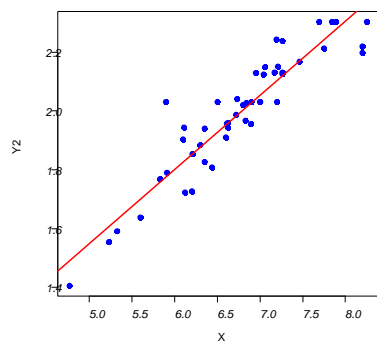
2.- Calculeu la recta de regressió tot estimant els valors de β_0 i β_1 . Com interpreteu el valor de β_1 ? (2 punts).

3.- Poseu a prova si la recta de regressió passa per l'origen de coordenades amb un risc $\alpha = 0.05$ (1 punt).

4.- Calculeu un interval del 95% de confiança pel paràmetre β_1 i interpreteu els resultats (1 punt).

5.- Interpreteu els valors del coeficient de determinació i de correlació (1 punt).

Hem aplicat una transformació logarítmica als resultats del segon examen (Y) obtenint unes noves dades (Y2) pels 50 aspirants i n'estudiem la relació lineal amb X:



Call:
lm(formula = Y2 ~ X, data = datos)

Residuals:

Min	1Q	Median	3Q	Max
-0.157903	-0.050662	0.002815	0.043129	0.254894

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	0.2900	0.1019	2.847	0.00648	**
X	0.2519	0.0150	16.790	< 2e-16	***

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

Residual standard error: 0.0795 on 48 degrees of freedom
Multiple R-squared: 0.8545, Adjusted R-squared: 0.8515
F-statistic: 281.9 on 1 and 48 DF, p-value: < 2.2e-16

6.- Indiqueu el nou model, calculeu una predicció puntual per la nota (Y) pel cas de X=7 i calculeu un interval de confiança per valors individuals per la predicció (2 punts).

7.- Compareu els dos models en quant al coeficient de determinació i les premisses. Creieu que ha valgut la pena aplicar el logaritme a la Y en aquest sentit? (1 punt).