

**NOM:** \_\_\_\_\_**(Poseu el nom i contesteu cada pregunta en el seu lloc reservat. Expliciteu i justifiqueu els passos en les respostes).****Problema B4.**

1) (2,5 punts) La implementació d'e-status en l'assignatura PE ha millorat el rendiment acadèmic dels alumnes que l'utilitzen. Se sap que la proporció d'execucions de problemes no aprovades és del 10%, tot i que s'espera disminuir aquest percentatge. Durant el Q1 del curs 2010-2011 es va prendre una mostra de 115 execucions i es va observar que en 10 execucions no es va aconseguir l'aprovat. Hi ha evidència de canvi en la proporció de problemes aprovats? Per respondre plantegeu una PS.

- (0,5 punts) Hipòtesi (i indicar si la prova és bilateral o unilateral):

- (0,5 punts) Càlcul del valor de l'estadístic:

- (0,5 punts) Càlcul del P-valor:

- (0,5 punts) Representeu gràficament el valor obtingut i el punt crític sobre la distribució de l'estadístic:

- (0,5 punts) Decisió de la PS i interpretació:

Per conèixer l'ús que fan els estudiants d'e-status i així avaluar la seva dedicació a l'assignatura es va decidir estudiar el nombre de vegades que els alumnes resolen els problemes del programa durant el quadrimestre. Es va prendre una mostra aleatòria de 20 estudiants i es va registrar el nombre d'execucions de problemes que van realitzar al llarg del curs, obtenint els següents resultats.

$$\sum x = 1228$$

$$\sum x^2 = 86984$$

Si es considera que el nombre d'execucions és una variable amb distribució normal:

2) (1 punt) Amb aquesta mostra calculeu una estimació puntual del nombre mitjà d'execucions realitzades durant el quadrimestre.

3) (1 punt) Doneu una estimació de l'error de la mitjana o error típic.

4) (1,5 punts) Estimeu el nombre mitjà d'execucions amb un IC al 90%.

5) (1,5 punts) Un altre aspecte a considerar per a una millor avaluació de l'ús d'e-status és estudiar les variacions en el nombre d'execucions. Obtingui un IC al 95% per a la variància del nombre d'execucions per a la mostra disponible.

6) (2,5 punts) Estudis realitzats en assignatures similars a PE d'altres estudis indiquen que el nombre mitjà d'execucions realitzades en un quadrimestre és de 80. Utilitzeu la mostra disponible per posar a prova amb un CH si els alumnes de PE formen part de la mateixa població que els d'altres estudis (en quant a la mitjana). Preneu  $\alpha = 5\%$ .

- (0,5 punts) Hipòtesis:

- (0,5 punts) Càlcul del valor de l'estadístic:

- (0,5 punts) Quin és el valor que utilitzarà per decidir si l'estadístic anterior permet rebutjar la hipòtesi nul·la?

- (0,5 punts) Representeu gràficament el valor obtingut i el punt crític sobre la distribució de l'estadístic.

- (0,5 punts) Decisió del CH i interpretació:

NOM: \_\_\_\_\_

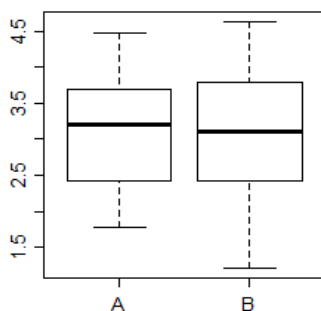
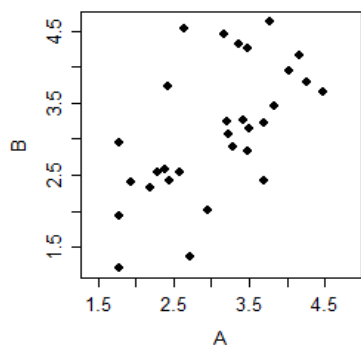
(Poseu el nom i contesteu cada pregunta en el seu lloc reservat. Expliciteu i justifiqueu els passos en les respostes).

**Problema B5.**

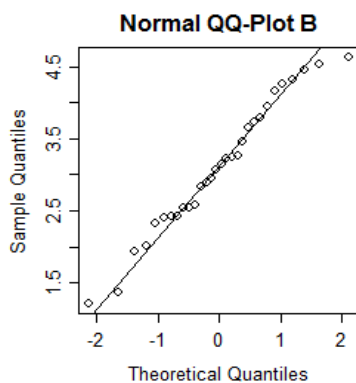
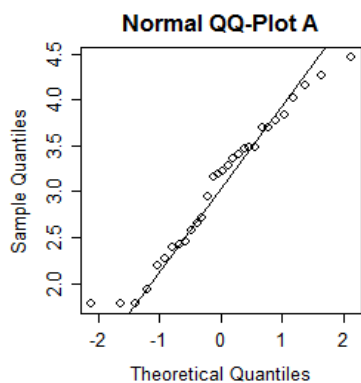
Un enginyer informàtic està interessat en la comparació de la eficiència de dos algorismes SortA i SortB per a ordenar vectors amb 10 milions de números reals. Per a tal propòsit, l'enginyer genera 10 milions de números aleatoris, i els ordena amb SortA i SortB dins l'entorn estadístic R, fent servir la funció `system.time` per mesurar el temps d'execució. Les mesures del temps d'execució es repeteixen 30 vegades pels dos algorismes, amb el codi:

```
n <- 30
m <- 10000000
A <- NULL
B <- NULL
for (i in 1:n) {
  z <- rnorm(m)
  a <- system.time(SortA(z))
  A <- c(A,a[3])
  b <- system.time(SortB(z))
  B <- c(B,b[3])
}
D <- B-A
```

La funció `system.time` retorna una estructura amb 3 elements, essent el tercer element el temps (en segons) gastat per la execució corresponent. Amb les dades obtingudes, l'enginyer fa estadística descriptiva gràfica i numèrica, obtenint els resultats a continuació.



Descriptiva gràfica



Descriptiva numèrica

	N	Mean	Variance
A	30	3.063	0.608
B	30	3.108	0.843
D	30	0.045	0.604

- (1p) El disseny emprat en aquest estudi és de dues mostres independents o de dades aparellades? Argumenteu la resposta.
- (3p) Feu el contrast de hipòtesi  $H_0: \mu_A = \mu_B$  versus  $H_1: \mu_A \neq \mu_B$ , detallant el càlcul de l'estadístic utilitzat, i indicant la seva distribució de referència, premisses del contrast, la regla de decisió, el P-valor (encara que sigui aproximadament) i la decisió.



**NOM:** \_\_\_\_\_

(Poseu el nom i contesteu cada pregunta en el seu lloc reservat. Expliciteu i justifiqueu els passos en les respostes).

**Problema B6**

Es vol estudiar si el temps de CPU (en milisegons) d'un programa ve afectat pel nombre d'operacions d'entrada i sortida (E/S) amb el disc. Es pren una mostra de 7 observacions i s'obtenen les següents dades:

nombre E/S (X)	CPU (Y)
14	2
16	5
27	7
42	9
39	10
50	13
83	20

Amb aquestes dades tenim que:

$$\sum X = 271,$$

$$\sum X^2 = 13855,$$

$$\sum Y = 66,$$

$$\sum Y^2 = 828,$$

$$\sum XY = 3375$$

a) Calculeu els estimadors de la constant i del pendent de la recta de regressió i representeu gràficament la recta estimada (2 punts)

b) Calculeu la taula de descomposició de la variància (ajut: la variància residual val 1.173) (2 punts)

c) Calculeu i interpreteu el coeficient de determinació  $R^2$  (1 punt)

d) Calculeu el interval de confiança al 95% pel pendent de la recta i poseu a prova si aquest és significatiu. Quines premisses s'han de complir per a que tingui validesa aquesta estimació? (2 punts)

e) Doneu una previsió, així com el seu interval de confiança al 95%, del valor promig del temps de CPU quan el nombre d'E/S és 30 (2 punts)

f) La següent figura ens mostra les gràfiques dels residus necessàries per poder validar aquest model de regressió. A partir de la inspecció gràfica d'aquestes, dieu quines premisses es verifiquen i si n'hi ha alguna que no. Justifiqueu la vostra resposta. (1 punt)

