

NOM: _____

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

Problema 1 (B4). Es vol verificar la qualitat d'un sistema informàtic (SI) de vendes d'entrades de cinema. S'assumeix que aquest sistema funciona correctament si el temps de servei mitjà per client no supera els 2 minuts ja que si es superessin aquests 2 minuts, es generarien unes cues molt llargues amb el risc que els clients abandonin aquest sistema i vagin a comprar les entrades a un altre cinema, que tingui un temps de resposta més ràpid.

Per analitzar el temps de servei es recull una mostra de 25 observacions i s'observa que la mitjana mostral d'aquest temps de servei és de 2.17 minuts i la desviació típica mostral és de 0.4 minuts.

1. **(1 punt)** Doneu una estimació puntual de l'error tipus per estimar la esperança del temps de servei SI. **(1 punt)**

2. **(2 punts)** A continuació doneu un interval de confiança (IC) al 95% pel valor esperat (μ) del temps de servei SI, responnent als següents apartats:
 - a. Estadístic
 - b. Premisses
 - c. Distribució de probabilitat que ha de seguir aquest estadístic
 - d. Calculeu l'IC
 - e. Interpreteu els resultats

3. **(1 punt)** Quina distribució de referència podem usar per l'estadístic de l'apartat 2.a) si en comptes de tenir una mostra de 25 observacions tinguéssim una mostra de 100 observacions? Justifiqueu la vostra resposta.

4. **(1.5 punts)** Plantegeu i resoleu una prova d'hipòtesis bilateral (al 95%) que us permeti decidir si la mitjana de temps de servei és $\mu=2$.

5. **(2 punts)**

A) Doneu les estimacions per interval al 90, 95 i 99% de confiança amb $\sigma=1$ per estimar la esperança del temps de servei de **SI** ($n=25$).

B) Quina distribució de referència usem en aquest cas per l'estadístic? Per què?

C) Comenteu com i perquè canvia un IC quan augmentem la confiança de 90 a 95 i a 99%

6. **(2.5 punts)**

A) Doneu les estimacions per interval (al 90, 95 i 99% de confiança) per estimar la esperança del temps de servei de **SI** amb σ desconeguda ($n=25$)

B) Quina distribució de referència usem en aquest cas per l'estadístic? Per què?

C) Comenteu com i perquè canvia un IC si el calculem amb sigma coneguda o no coneguda.

NOM: _____

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

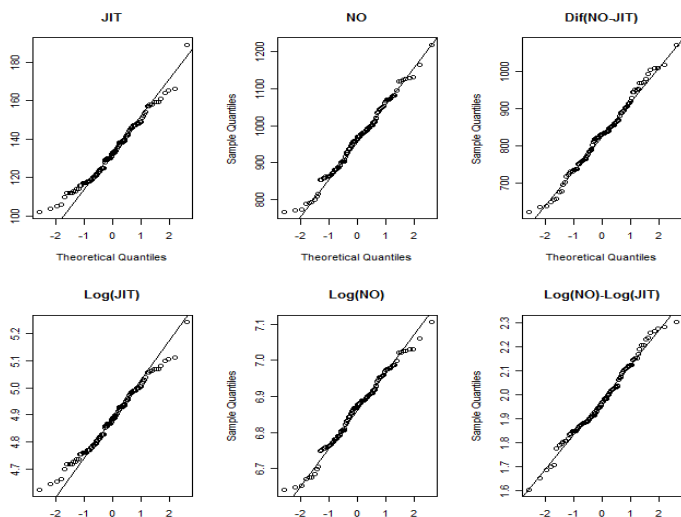
Problema 2 (B5). La màquina virtual de Java disposa d'una optimització anomenada JIT (*Just In Time*) que permet compilar les seccions de codi més utilitzades per reduir el temps d'execució. Estudis previs indiquen que els temps sense JIT són més de 7 vegades superiors als de amb JIT (anomenant els temps NO i JIT respectivament, indiquen que el rati $NO/JIT \approx 7$). Per confirmar-ho, uns companys de PE han generat de forma aleatòria 100 vectors de mida 200000 i per a cada vector han obtingut els temps d'ordenació de Mergesort amb i sense JIT.

La taula mostra la seva descriptiva (JIT i NO) i la del seus logaritmes naturals ($\text{Log}(JIT)$ i $\text{Log}(NO)$), així com de les diferències respectives ($NO-JIT$, i $\text{Log}(NO)-\text{Log}(JIT)=\text{Log}(NO/JIT)$ que anomenarem D).

	N=100	Mitjana	Desv. est	Corr
JIT		133.7	16.59	0.170
NO		962.8	96.56	
NO-JIT		829.1	95.15	
$\text{Log}(JIT)$		4.887	0.1228	0.154
$\text{Log}(NO)$		6.874	0.1009	
$\text{Log}(NO)-\text{Log}(JIT)=D$		1.976	0.1464	

1) ^{1 punt} D'acord amb la recollida i amb la descriptiva de les dades, es tracta de dues mostres independents o aparellades? Raoneu la resposta.

2) ^{1 punt} Té avantatges haver recollit les dades de forma aparellada? Compareu els errors estàndard.



3) ^{1 punt} Veient aquests Q-Q plots, digueu si les variables s'acosten al model Normal i quines s'acosten més.

4) ^{1 punt} Comenteu avantatges en interpretació de resultats i premisses per tractar les dades logotransformades en aquest exemple.

5) ^{4 punts} Amb les dades de les diferències de logaritmes, contrasteu la H_0 que l'esperança poblacional (μ_D) de D (on $D=\log(\text{NO})-\log(\text{JIT})=\log(\text{NO}/\text{JIT})$) sigui 1.946 ($\approx\log(7)$) enfront de l'alternativa que sense JIT trigui més (més de 7 vegades abans de fer logaritmes).

Indiqueu: hipòtesi, premisses, estadístic, distribució i punt crític, càlculs, resultats i interpretació.

6) ^{2 punts} Calculeu un IC per a μ_D amb una confiança del 90%, i interpreteu com ens informa del rati NO/JIT. Indiqueu i calculeu com trobaríeu un valor k del qual podem afirmar, amb una confiança del 95%, que JIT és, com a mínim, k vegades més ràpid que sense JIT ($\text{NO}/\text{JIT} \approx k$)

>qt(0.975,99) = 1.984217
>qt(0.95,99) = 1.660391
>qt(0.90,99) =] 1.290161

>qt(0.975,198) = 1.972017
>qt(0.95,198) = 1.652586
>qt(0.90,198) = 1.285842

>pt(0.975,99) = 0.8340317
>pt(0.95,99) = 0.8277872
>pt(0.90,99) = 0.8148479

>pt(0.975,198) = 0.834625
>pt(0.95,198) = 0.8283648
>pt(0.90,198) = 0.8153932

NOM:

PROBLEMA 3 BLOC B6

S'ha mesurat, en segons, el temps de processament de 7 programes en funció del nombre de milions d'accessos Input/Output al disc. Els resultats apareixen a la següent taula:

Accessos I/O	Temps de CPU
14	2
16	5
27	7
42	9
39	10
50	13
83	20

1) Especifiqueu quina és la variable explicativa o independent (X) i quina la variable resposta o dependent (Y). De quin tipus són aquestes variables? (**1 punt**).

2) Calculeu la recta de regressió que relaciona les dues variables anteriors. Per tal d'alleugerir els càlculs, sapiguen que: $\sum_i x_i = 271$, $\sum_i x_i^2 = 13855$, $\sum_i y_i = 66$, $\sum_i y_i^2 = 828$ i el coeficient de correlació és $r = 0.9867$ (**3 punts**).

3) Què passarà amb la variable resposta si incrementem en dos milions d'unitats el nombre d'accessos I/O? I si l'increment és de 10 milions? **(2 punts)**.

4) Calculeu l'estimació puntual de la variància residual **(1 punt)**.

5) Calculeu la predicció puntual corresponent a 43 milions d'accessos I/O **(1 punt)**.

6) Calculeu un interval de confiança al 95% per el pendent de la recta de regressió. **(2 punts)**.