

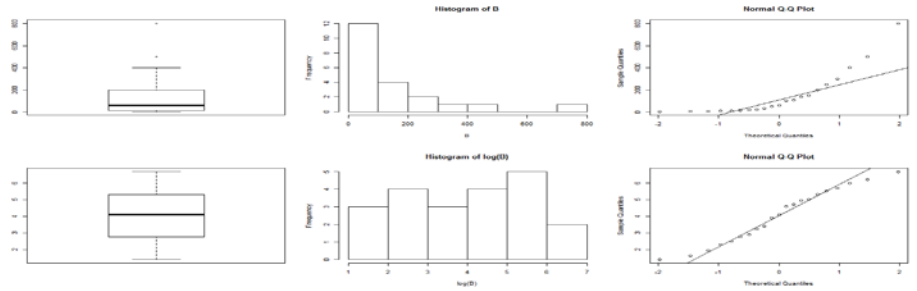
NOM: \_\_\_\_\_ COGNOM: \_\_\_\_\_

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

### Problema 1 (B4)

Els últims anys s'han popularitzat els serveis d'emmagatzematge al núvol. Hem recollit el temps de pujada a Dropbox(B) de 21 arxius de text de diverses mides, amb els següents resultats del temps de pujada (B) i dels logaritmes d'aquests temps (log\_B):

	mitjana	Desviació tipus
B	151.76	203.91
log_B	4.03	1.60



Indiqueu com argumentar amb els resultats de l'enunciat si es compleix o no la premisa de normalitat en B i log\_B (1 punt)

A) Primer suposarem coneguda la desviació poblacional del logaritme d'aquests temps (log\_B), igual a 1.5 i plantejarem la prova d'hipòtesis sobre si el valor esperat de la mitjana de log\_B és 5 o no amb un risc del 5%. Per això indiqueu:

- les hipòtesis (1 punt)

- l'estadístic amb la seva distribució sota la hipòtesis nul·la, i calculeu-ne el valor (1 punt)

- calculeu el p\_valor (1 punt)

- representeu gràficament l'estadístic amb els punts crítics, el p\_valor i les zones d'acceptació i rebuig (1 punt)

- en funció dels apartats anteriors, a quina conclusió arribeu sobre la prova d'hipòtesi. Interpreteu els resultats (1 punt)

B) Ara no suposarem coneguda la desviació poblacional del logaritme d'aquests temps i tornem a plantejar la prova d'hipòtesis sobre si el valor esperat de la mitjana del logaritme del temps és 5 o no amb un risc del 5%. Per això indiqueu:

- l'estadístic amb la seva distribució sota la hipòtesis nul·la, i calculeu-ne el valor (1 punt)

- l'error estàndard de l'estimador de la mitjana poblacional (1 punt)

- en funció dels apartats anteriors, a quina conclusió arribeu sobre la prova d'hipòtesi. Interpreteu els resultats (1 punt)

- proporcioneu un interval de confiança del 95% per a la  $\mu$  i interpreteu-lo (1 punt)

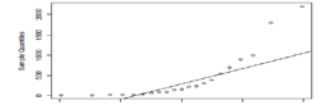
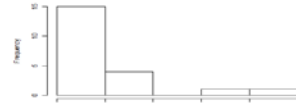
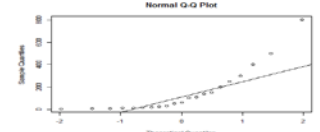
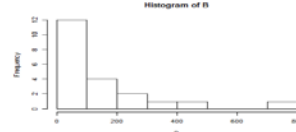
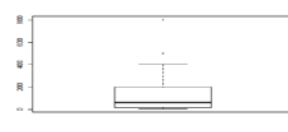
NOM: \_\_\_\_\_ COGNOM: \_\_\_\_\_

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

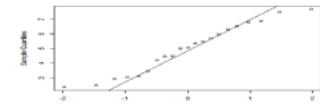
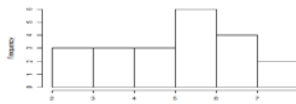
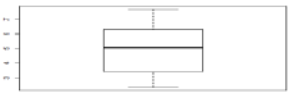
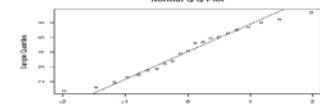
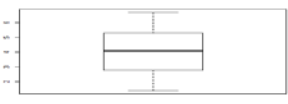
## Problema 2 (B5)

Continuant el problema 1, ara volem avaluar la diferència de temps de pujada entre Dropbox (B) i Google Drive (G). Contractem dos comptes amb espai il·limitat i pugem 21 arxius de diverses mides a Dropbox i **uns altres 21 arxius** de mides similars a Google Drive. Els arxius són pujats en ordre aleatori a cadascun dels 2 sistemes. Els resultats del temps de pujada mesurats (B i G) i dels logaritmes d'aquests temps ( $\log_B$  i  $\log_G$ ) són:

	mitjana	Desviació tipus
B	151.76	203.91
G	424.94	602.05



	mitjana	Desviació tipus
$\log_B$	4.03	1.60
$\log_G$	4.99	1.65



- Indiqueu i justifiqueu d'acord amb l'enunciat si el disseny emprat és independent o aparellat (1 punt)

- Expressu les hipòtesis de la prova de igualtat de mitjanes (bilateral) dels logaritmes dels temps (1 punt)

- Indiqueu com podríem argumentar igualtat de variàncies amb els gràfics de l'enunciat (1 punt)

- Sota la hipòtesi d'igualtat, quin seria l'error tipus estimat per a la diferència de mitjanes mostrals? (1 punt)

- Indiqueu quin és l'estadístic de la prova i calculeu-lo (1 punt)

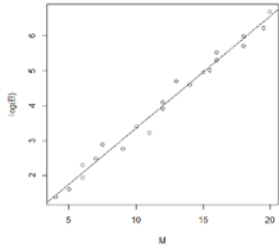
- Si no hi hagués cap diferència en la velocitat mitjana dels dos fabricants, com es distribuiria l'estadístic de la prova? Amb un risc  $\alpha=5\%$ , feu un gràfic per il·lustrar els punts crítics i situar les àrees d'acceptació i de rebuig de la hipòtesi nul·la. (1 punt)
  
- Doneu la conclusió de la prova i interpreteu els resultats. (1 punt)
  
- Calculeu l'interval de confiança del 95% per a la diferència de esperances ( $\mu_{\log\_B} - \mu_{\log\_G}$ ) (1 punt)
  
- Interpreteu el IC anterior i comenteu què aporta a la conclusió de la prova. (1 punt)
  
- Si els 21 arxius pujats a Dropbox i Google Drive fossin els mateixos comenteu què canviaria en el disseny emprat (1 punt)

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

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

### Problema 3 (B6)

Finalment volem explicar el logaritme del temps de pujada a Dropbox ( $\log_B$ ) en funció de la mida  $M$  (entre i 20 GB) dels 21 arxius. Els resultats de la descriptiva de  $\log_B$  i  $M$ , de la covariància entre  $\log_B$  i  $M$ , i de la regressió lineal corresponents són:



	Mitjana	Desviació tipus
Log B	4.03	1.60
M	12.1	4.97

$$\text{cov}(\log B, M) = S_{M \log B} = 7.90$$

```
summary(lm(log_B ~ M)):  Coefficients:      Estimate      Std. Error      t value      Pr(>|t|)
      (Intercept)      -----      0.122702      1.284      0.214
           M           -----      0.009399     34.020     <2e-16
Residual standard error: 0.209 on 19 degrees of freedom
Multiple R-squared:  -----
```

- Calculeu la correlació entre  $M$  i  $\log_B$  i el coeficient de determinació ( $R^2$ ) (1 punt)

- Interpreteu els valors anteriors (1 punt)

- Calculeu els coeficients de la recta de regressió del logaritme del temps ( $\log_B$ ) en funció de la mida  $M$  (1 punt)

- Interpreteu els coeficients de la recta de regressió del logaritme del temps ( $\log_B$ ) en funció de la mida  $M$  (1 punt)

- Contrasteu si la recta passa per l'origen (risc del 5%). Indiqueu l'estadístic i la conclusió i interpretació de la prova (1 punt)

- Contrasteu si la recta de regressió és plana (risc del 5%). Indiqueu l'estadístic i la conclusió i interpretació de la prova (1 punt)

- Calculeu un IC95% del pendent de la recta. (1 punt)

- Quin és el valor de la variabilitat residual (o del terme d'error) i quina informació aporta ? (1 punt)

- Enuncieu les premisses o hipòtesis de la regressió lineal i comenteu si es compleixen o no per aquest cas concret. Especifiqueu de quins resultats i/o gràfics es dedueixen els vostres comentaris. (2 punts)

