

NOM: _____ COGNOM: _____

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

Problema 1 (Bloc C)

Volem estudiar unes dades sobre emissions de CO₂ (gCO₂/km) per models de vehicles dièsel (L), gasolina (G) i híbrids (H), tenint en compte tots els seus cicles de vida. Tenim unes mostres aleatòries simples de les emissions per a 24 models tant en versió dièsel com gasolina (L i G respectivament, i D per a la seva diferència) i per a uns altres 24 models híbrids (H), diferents tot i que comparables amb els anteriors.

$$\sum_{i=1}^{24} L_i = 3334.25 \quad \sum_{i=1}^{24} L_i^2 = 468018.8$$

$$\sum_{i=1}^{24} G_i = 3332.85 \quad \sum_{i=1}^{24} G_i^2 = 467669$$

$$\sum_{i=1}^{24} D_i = 1.4 \quad \sum_{i=1}^{24} D_i^2 = 1.315$$

$$\sum_{i=1}^{24} H_i = 2937$$

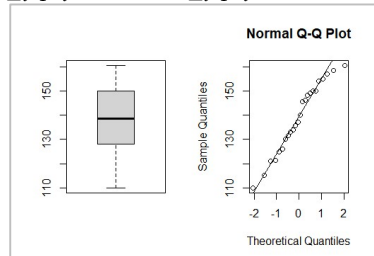


Figura 1

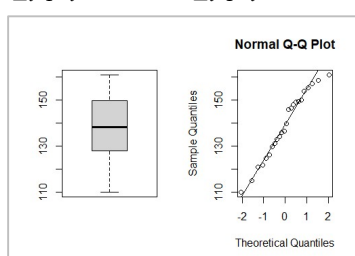


Figura 2

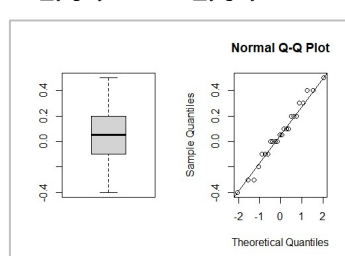


Figura 3

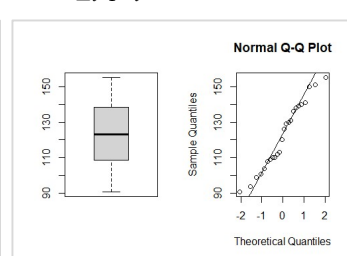


Figura 4

1.- (1 punt) Per una part volem comparar la mitjana d'emissions entre els dièsel i gasolina, i per altra entre els gasolina i híbrids. Indiqueu i justifiqueu si cadascuna d'aquestes dues proves de comparació de mitjanes serien de mostres aparellades o independents. Comenteu avantatges i inconvenients de les dues possibilitats.

2.- (1 punt) Calculeu una estimació puntual per a la diferència mitjana d'emissions entre dièsel i gasolina. Indiqueu la mitjana, la desviació, i el seu error tipus o estàndard i interpreteu-los

3.- (1.5 punt) Calculeu un IC al 95% de la diferència mitjana entre dièsel i gasolina. Interpreteu-lo i comenteu quina informació aporta de cara a concloure si les mitjanes d'emissions entre dièsel i gasolina són equivalents o no

Per comparar la mitjana entre gasolina i híbrids hem obtingut els següent resultat en R:

```
t.test(G,H,var.equal=T,conf.level=0.99)
t = 3.3904, df = 46, p-value = 0.001442
alternative hypothesis: true difference in means is not equal to 0
99 percent confidence interval:
 3.421888 29.565612
sample estimates:
mean of x mean of y
138.8688 122.3750
```

4.- (1 punt) Indiqueu una estimació puntual per a la diferència mitjana d'emissions entre gasolina i híbrids. Indiqueu la mitjana i el seu error tipus o estàndard

- 5.- (1.5 punts) Indiqueu un IC al **99%** de la diferència mitjana entre gasolina i híbrids. Interpreteu-lo i comenteu quina informació aporta de cara a concloure si les mitjanes d'emissions entre gasolina i híbrids són equivalents o no
- 6.- (1 punt) A partir dels resultats R i de les gràfiques, pel cas de comparació entre gasolina i híbrids, indiqueu si s'ha fet una comparació de mitjanes suposant homoscedasticitat o no i què vol dir
- 7.- (1 punt) Indiqueu pels intervals de les preguntes 3 i 5 les premisses que han de complir, i si es compleixen o no
- 8.- (2 punts) Ara volem comparar el percentatge de vehicles que superen un cert valor crític. En els híbrids, dels 24 estudiats, 10 el superen; i en els dièsel i gasolina, dels 48 estudiats, 38 superen el valor crític. Calculeu un IC al **90%** per a la diferència de proporcions per concloure si són percentatges equivalents o no

Valors que poden ser útils pels blocs C i D:

$qt(0.975,11)= 2.201$	$qt(0.975,21)= 2.079$	$qt(0.975,46)= 2.013$	$qchisq(0.025,46)= 15.308$	$qnorm(0,85)= 1,036$
$qt(0.975,12)= 2.179$	$qt(0.975,22)= 2.074$	$qt(0.975,47)= 2.012$	$qchisq(0.025,47)= 16.047$	$qnorm(0,9)= 1,282$
$qt(0.975,13)= 2.160$	$qt(0.975,23)= 2.069$	$qt(0.975,48)= 2.011$	$qchisq(0.025,48)= 16.791$	$qnorm(0,95)= 1,645$
$qt(0.995,11)= 3.106$	$qt(0.995,21)= 2.831$	$qt(0.995,46)= 2.687$	$qchisq(0.975,46)= 44.461$	$qnorm(0,975)= 1,960$
$qt(0.995,12)= 3.054$	$qt(0.995,22)= 2.819$	$qt(0.995,47)= 2.685$	$qchisq(0.975,47)= 45.722$	$qnorm(0,99)= 2,326$
$qt(0.995,13)= 3.012$	$qt(0.995,23)= 2.807$	$qt(0.995,48)= 2.682$	$qchisq(0.975,48)= 46.979$	$qnorm(0,995)= 2,576$

Problema 2 (Bloc D)

Un estudi sobre vehicles considera dades pels anys de 2011 a 2023: la variable **A** com a “any–2010” (pren valors des de 1 fins a 13); la variable **t**, tipus de vehicle (BEV, híbrid convencional, o PHEV, híbrid endollable); i les vendes anuals a tot el mon. Aquí teniu els gràfics de les variables amb les vendes, abans i després d’aplicar *logaritme natural* (**sales** i **LOG(sales)**).

Hem ajustat amb R dos models (a sota). Pel model (1) la variable de resposta és el logaritme natural del total de vendes anual (l’etiqueta “both”, al gràfic). Pel model (2) la variable de resposta és el logaritme natural de les vendes anuals per als dos tipus.

```
(1) summary(lm(log(total_sales) ~ A))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	11.67374	0.11956	97.64	<2e-16
A	0.44440	0.01506	29.50	8e-12

Residual standard error: 0.2032 on 11 deg. of fr.
Multiple R-squared: 0.9875, Adj. R-sq.: 0.9864
F-statistic: 870.4 on 1 and 11 DF, p-value: 7.998e-12

```
(2) summary(lm(log(sales) ~ A + t))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	11.19873	0.15581	71.875	< 2e-16
A	0.44738	0.01776	25.196	< 2e-16
tPHEV	-0.59866	0.13287	-4.505	0.00016

Residual standard error: 0.3388 on 23 deg. of fr.
Multiple R-squared: 0.9661, Adj. R-sq.: 0.9631
F-statistic: 327.6 on 2 and 23 DF, p-value: < 2.2e-16

- Expliqueu què és i com s’interpreta el valor 11.67374 del model (1).
- Pel valor equivalent al model (2), 11.19873, subratlleu les similituds i les diferències.
- Expliqueu què és i com s’interpreta el valor 0.44440 del model (1). *Pista: invertiu la transformació logarítmica i raoneu en conseqüència.*
- Què informació aporta el valor 0.01506?

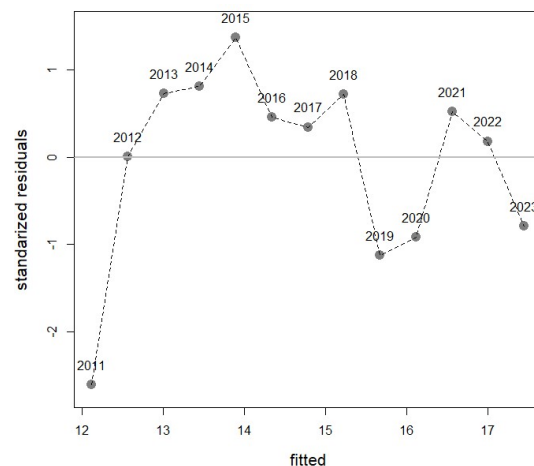
5. Interpreteu els tres primers valors de la línia t_{PHEV} , al model (2): quina informació proporciona aquesta estimació?

6. Calculeu un interval de confiança al 95% pel paràmetre associat a la variable t .

7. Valoreu la capacitat/qualitat del primer model per fer prediccions: quin indicador és l'adequat, i què ens diu?

8. Com ens retorna el model informació sobre les discrepàncies entre les observacions i les prediccions? Quantifiquem aquesta informació pel cas del segon model.

9. A la dreta teniu un gràfic obtingut pel model (1). Descriviu què està representant. Indiqueu les premisses del model i comenteu críticament i justificada quines es poden donar per assumibles.



10. Amb un dels models anteriors hem executat aquesta instrucció. Expliqueu quin resultat està proporcionant.

```
> predict(mod, data.frame(A=15,t='BEV'), int='prediction')
      fit      lwr      upr
1 17.90945 17.12508 18.69381
```