NOM:_____COGNOMS:_____

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

Problema 1 (B1-B2)

Tenim una moneda trucada de tal manera que la probabilitat de cara és ¼ i la de creu és 3/4, i la llencem 4 vegades seguides.

1.- (1 punt) Indiqueu l'arbre de l'experiència aleatòria, i el conjunt de resultats amb les seves probabilitats.

Resultats:

Ω = { w1 , w2 , w3 , w4 , w5 , w6 , w7 , w8 , w9 , w10 , w11 , w12 , w13 , w14 , w15 , w16 }

= { oo oo, oo o+, oo +o, oo ++, o+ oo, o+ o+, o+ +o, o+ ++, +o oo, +o o+, +o +o, +o ++, ++ oo, ++ o+, ++ +o, ++ ++ }

amb probabilitats:

$1/4^4$	$3/4^4$	$3/4^4$	$3^2/4^4$	$3/4^4$	$3^2/4^4$	$3^2/4^4$	$3^3/4^4$	$3/4^4$	$3^2/4^4$	$3^2/4^4$	$3^3/4^4$	$3^2/4^4$	$3^3/4^4$	$3^3/4^4$	$3^4/4^4$
1/256	3/256	3/256	9/256	3/256	9/256	9/256	27/256	3/256	9/256	9/256	27/256	9/256	27/256	27/256	81/256

2.- (1 punt) Calculeu i justifiqueu formalment quina és la probabilitat que surtin primer 2 cares seguides i després 2 creus. I que surtin primer 2 creus seguides i després 2 cares.

P(oo++) = 1/4*1/4*3/4*3/4 = $3^2/4^4$ **= 9 /256 = 0.035**

P(++oo) = P(oo++) = **0.035**

3.- (1 punt) Calculeu i justifiqueu formalment quina és la probabilitat que surtin 3 cares i una creu

P(ooo+) + P(oo+o) + P(o+oo) + P(+ooo) = 4 * 1/4*1/4*1/4*3/4 = **4 * $3/4^4$ = 12 / 256 = 3 / 64 = 0.047**

4.- (1 punt) Calculeu i justifiqueu formalment quina és la probabilitat que a la quarta tirada surti cara quan a la primera ja ha sortit cara

4o = "quarta tirada amb cara" i 1o = "primera tirada amb cara"

P(4o | 1o) = P(4o i 1o) / P(1o) =

= (P(**o**oo**o**) + P(**o**o+**o**) + P(**o**+o**o**) + P(**o**++**o**)) / (P(**o**oo**o**)+ P(**o**oo+)+P(**o**o+**o**)+ P(**o**o++)+P(**o**+o**o**)+ P(**o**+o+)+P(**o**++**o**)+P(**o**+++))

= ($1/4^4$ + $3/4^4$ + $3/4^4$ + $3^2/4^4$) / ($1/4^4$ + $3/4^4$ + $3/4^4$ + $3^2/4^4$ + $3/4^4$ + $3^2/4^4$ + $3^2/4^4$ + $3^3/4^4$))

= **($16/4^4$) / ($64/4^4$)) = 16 / 64 = 0.25**

* También es podria deduir sabent que els successos són independents:

P(4o | 1o) = P(4o) = P(o) = 1/4

5.- (1 punt) Calculeu i justifiqueu formalment quina és la probabilitat que surtin més cares que creus

P(oooo) + P(ooo+) + P(oo+o) + P(o+oo) + P(+ooo) = $1/4^4$ + 4* $3/4^4$ = **$13/4^4$ = 13 / 256 = 0.0508**

Seguint amb la moneda anterior que hem llençat 4 vegades seguides:

6.- (2 punts) Definiu les variables aleatòries nombre de cares i nombre de creus, indicant la funció de probabilitat i la funció de distribució de probabilitat i la seva representació gràfica

No = "nombre de cares en 4 tirades" N+="nombre de creus en 4 tirades"

Els valors possibles són 0, 1, 2, 3 o 4 i els nombre de resultats pels diferents valors són $\binom{4}{0}=1$ $\binom{4}{1}=4$ $\binom{4}{2}=6$ $\binom{4}{3}=4$ $\binom{4}{4}=1$

La v.a. "No" té per funcions de probabilitat i de distribució de probabilitat:

k	$p_{No}(k)$	$F_{No}(k)$
0	$1*3^4/4^4$ = **81/256**	81/256
1	$4*3^3/4^4$ = **108/256**	189/256
2	$6*3^2/4^4$ = **54/256**	243/256
3	$4*3/4^4$ = **12/256**	255/256
4	$1*1/4^4$ = **1/256**	256/256

La v.a. "N+" té per funcions de probabilitat i de distribució de probabilitat:

k	$p_{N+}(k)$	$F_{N+}(k)$
0	$1 * 1/4^4$ = **1/256**	1/256
1	$4 * 3/4^4$ = **12/256**	13/256
2	$6 * 3^2/4^4$ = **54/256**	67/256
3	$4 * 3^3/4^4$ = **108/256**	175/256
4	$1 * 3^4/4^4$ = **81/256**	256/256

7.- (1 punt) Calculeu l'esperança i la variància de les variables aleatòries anteriors

E(No) = 0*81/256 + 1*108/256 + 2*54/256 + 3*12/256 + 4*1/256 = (0+108+108+36+4)/256 = 256/256 = **1**

o bé "No" és B(4,1/4) amb esperança n*p = 4*1/4 = 1

V(No) = $(0-1)^2*81/256 + (1-1)^2*108/256 + (2-1)^2*54/256 + (3-1)^2*12/256 + (4-1)^2*1/256$ = (81+0+54+48+9)/256 = 192/256 = 3/4 = **0.75**

o bé "*No*" és B(4,1/4) amb variància n*p*q = 4*1/4*3/4 = 0.75

E(N+) = 0*1/256 + 1*12/256 + 2*54/256 + 3*108/256 + 4*81/256 = (0+12+108+324+324)/256 = 768/256 = **3**

o bé "N+" és B(4,3/4) amb esperança n*p = 4*3/4 = 3 o bé N+=4-No i E(N+) = E(4-No) = 4 – E(No) = 4-1=3

V(N+) = $(0-3)^2*1/256 + (1-3)^2*12/256 + (2-3)^2*54/256 + (3-3)^2*108/256 + (4-3)^2*81/256$ = (9+48+54+0+81)/256 = 192/256 = 3/4 = **0.75**

o bé "*N+*" és B(4,3/4) amb variància n*p*q = 4*3/4*1/4 = 0.75 o bé N+=4-No i V(N+) = V(4-No) = V(No) = 0.75

8.- (2 punts) Indiqueu i justifiqueu la taula de probabilitat conjunta de les dues variables anteriors. Calculeu la seva covariància i correlació i relacioneu els resultats amb el fet de ser dependents o independents

	No = 0	No = 1	No = 2	No = 3	No = 4
N+ = 0	0	0	0	0	1/256
N+ = 1	0	0	0	12/256	0
N+ = 2	0	0	54/256	0	0
N+ = 3	0	108/256	0	0	0
N+ = 4	81/256	0	0	0	0

Cov(No,N+) = (0-1)(4-3)*81/256 + (1-1)(3-3)*108/256 + (2-1)(2-3)*54/256 + (3-1)(1-3)*12/256 + (4-1)(0-3)*1/256 =

= -81/256 + 0 - 54/256 - 48/256 - 9/256 =

= - 192 / 256 = - 3/4

Corr(No,N+) = Cov(No,N+) / $(\sqrt{V(N_o)}\sqrt{V(N_+)})$ = (-3/4) / $(\sqrt{3/4}\sqrt{3/4})$ = (-3/4) / (3/4) = -1

Les dues variables tenen una dependència totalment inversa, la correlació és -1 i de fet es construeix la taula amb el fet que quan una variable pren un valor (per exemple 0 cares) ja determina totalment que l'altra pren el valor oposat (4 creus)

NOM:_____COGNOMS:_____

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

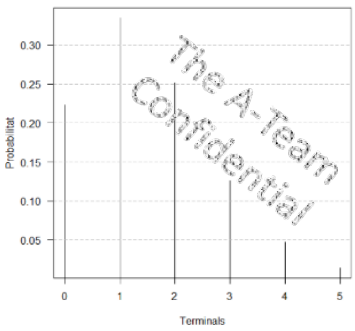
Problema 2 (B3-B4)

Els sistemes de detecció de possibles atacs informàtics de la UPC han previst un assalt imminent d'un equip d'hackers (*Equip X*) als sistemes de la Universitat. Només un equip format per 4 ex-alumnes de la FIB (*Equip A*) té els coneixements suficients per evitar aquest atac. Malgrat que, actualment, es troben en parador desconegut, la UPC els ha aconseguit trobar i els contractarà per aquesta tasca. **(Totes les qüestions valen 1 punt excepte les 2 últimes que valen 2 punts cadascuna)**

1. Hi ha 10 servidors a la UPC que proporcionen servei a les unitats fonamentals. Si fallen més de 5 d'aquests servidors, la UPC haurà de tancar. L'Equip A creu que la probabilitat de que el Equip X deixi inoperatius cadascun d'aquests servidors és de 0.6. Quina és la probabilitat que la UPC hagi de tancar?
   X="Número servidors inoperatius" ~ B(n=10, p=0.6)
   Y="Número servidors operatius" ~ B(n=10, p=0.4)
   $$P(X > 5) = P(Y \leq 4) = (Taules) = \mathbf{0.6331}$$

2. L'Equip A creu que pot introduir una millora en els *Firewalls* dels servidors per a que la probabilitat de que tanqui la UPC sigui de 0.102. Quina és ara la probabilitat de que quedi inoperatiu cadascun dels servidors?
   $$P(X > 5) = 1 - P(X \leq 5) = 0.102 \rightarrow P(X \leq 5) = 0.898 \rightarrow (Taules) \rightarrow p = \mathbf{0.45}$$

3. Entre els informes de l'Equip A, tenen el gràfic de la funció de probabilitat del número de terminals infectats per segon per un virus informàtic en els atacs de l'Equip X i saben que es distribueix de forma poissoniana. Digues a partir d'aquesta informació, quin és el temps esperat entre les infeccions consecutives de dos terminals.



   W="Nombre de terminals infectats per segon" ~ P(λ)
   T="Segons entre infeccions" ~ Exp(λ)
   Es veu en el gràfic que $P(W = 0) \approx 0.22$. Per tant:
   $$P(W = 0) \approx 0.22 \rightarrow e^{-\lambda} \approx 0.22 \rightarrow \lambda \approx -\ln(0.22) \approx 1.5$$
   Per tant, el temps esperat entre 2 infeccions consecutives serà:
   $$E(T) = 1/\lambda = 1/1.5 = \mathbf{0.67\ segons}$$

4. L'Equip A està capacitat per desactivar l'atac de l'equip X, un cop es comenci, en 3 minuts. Quina és la probabilitat que en aquest temps NO s'hagin infectat més de 250 terminals de la universitat?
   N="Nombre de terminals infectats en 3 minuts" ~ P(λ=180·1.5=270) ~ N(μ=270, σ^2=270)
   Z~N(0,1)
   $$P(N < 250) = P\left(\frac{N-\mu}{\sigma} < \frac{250-270}{\sqrt{270}}\right) = P(Z < -1.21) = (Taules) = 1 - 0.8869 = \mathbf{0.1131}$$

5. La UPC ha demanat un pressupost a l'Equip A per dur a terme la missió en funció del número d'hores invertides. Com que aquest número d'hores és variable i desconegut a priori, la UPC es conforma amb que se li proporcioni un interval dels euros pels qual es tingui una probabilitat de 0.90 de que l'import final estigui dins de l'interval. Es sap que la distribució de les hores invertides en missions similars ha estat N(μ=40, σ=10) i que el preu/hora és de 50€. Proporciona aquest interval.
   H="Nombre d'hores invertides" ~ N(μ=40,σ=10)
   Z~N(0,1)

   $$P(a \leq W \leq b) = 0.90 \rightarrow P\left(\frac{a-40}{10} \leq \frac{W-\mu}{\sigma} \leq \frac{b-40}{10}\right) = 0.90 \rightarrow P\left(\frac{a-40}{10} \leq Z \leq \frac{b-40}{10}\right) = 0.90$$
   $$\frac{b-40}{10} = (Taules) = 1.645 \rightarrow b = 40 + 1.645 \cdot 10 = 56.45 \quad \rightarrow \quad a = 40 - 16.45 = 29.55$$

   El Pressupost estarà entre **1477.5 €** (=29.55 · 50) i **2822.5 €** (=56.45 · 50) amb una probabilitat de 0.90.

La UPC vol tenir informació a priori de la magnitud dels atacs informàtics perpetrats per l'Equip X. Per aquest motiu, ha demanat a l'Equip A que li passi algunes dades sobre el percentatge (**P**) d'ordinadors afectats resultat d'atacs previs en altres institucions. L'Equip A proporciona la següent informació sobre aquest percentatge d'ordinadors afectats en els darrers 30 atacs pels quals disposen informació

Report Equip X — *18 de juny de 2021*

$$\sum_{i=1}^{30} P_i = 1575.7$$

$$\sum_{i=1}^{30} P_i^2 = 85609.77$$

Taula 1. Quantils del percentatge d'afectació en els darrers atacs perpetrats per l'equip X.

0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
35.8%	42.7%	44.3%	46.5%	50.0%	51.2%	53.5%	55.8%	60.3%	63.0%	76.9%

6. Calcula una estimació puntual de la mitjana d'aquest percentatge i del seu error estàndard.

$$\bar{P} = \frac{\sum_{i=1}^{30} P_i}{n} = \frac{1575.7}{30} = \mathbf{52.5}$$

$$s^2 = \frac{\sum_{i=1}^{30} P_i^2 - n \cdot (\bar{P})^2}{n-1} = \frac{85609.77 - 30 \cdot 52.5^2}{29} = 98.23 \rightarrow s = 9.9$$

$$SE = \frac{s}{\sqrt{n}} = \frac{9.9}{\sqrt{30}} = \mathbf{1.81}$$

7. Fes un contrast d'hipòtesi per veure si la mitjana poblacional és del 50% o no amb un α=5%. 1) Planteja el test; 2) Digues les premisses; 3) Calcula l'estadístic; i 4) Conclou sobre la prova

$$\begin{cases} \boldsymbol{H_0: \mu_P = 50} \\ \boldsymbol{H_1: \mu_P \neq 50} \end{cases}$$

Premissa: X Normal

$$\hat{t} = \frac{\bar{P} - \mu_P}{\sqrt{\frac{s}{n}}} = \frac{52.5 - 50}{1.81} = \mathbf{1.39}$$

Amb una t_{29}, no rebutjaríem la H_0 ja que el punt crític és 2.045>1.39. Per tant no descartem que la mitjana del percentatge d'ordinadors afectats sigui del 50%

8. Fes un contrast d'hipòtesi per veure si la proporció poblacional d'atacs amb més del 50% d'ordinadors afectats és de 0.5 o superior amb un α=1%. 1) Planteja el test; 2) Digues les premisses; 3) Calcula l'estadístic; i 4) Conclou sobre la prova

$$\begin{cases} \boldsymbol{H_0: \pi_P = 0.5} \\ \boldsymbol{H_1: \pi_P > 0.5} \end{cases}$$

Premissa: $n\pi > 5$ i $n(1-\pi) > 5$

p és la proporció d'atacs a la mostra que han superat el 50% d'afectació. Mirant la taula 1, veiem que el quantil 0.4 és 50.0% Per tant hi ha una proporció de 0.6 atacs que han superat aquest llindar.

$$\hat{Z} = \frac{p - \pi_P}{\sqrt{\frac{\pi_P \cdot (1-\pi_P)}{n}}} = \frac{0.6 - 0.5}{\sqrt{\frac{0.5 \cdot (1-0.5)}{30}}} = \mathbf{1.09}$$

Amb una N(0,1), no rebutjaríem la H_0 ja que el punt crític unilateral amb α=1% és 2.33>1.09. Per tant no descartem que la proporció d'atacs amb un percentatge d'ordinadors afectats sigui del 50%

NOM:_____COGNOMS:_____

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

Problema 3 (B5-B6)

Per tal de mesurar distàncies en xarxes de comunicacions de forma indirecta, uns investigadors han publicat un treball on descriuen que utilitzen el nombre de *hops* (nombre de salts entre nodes) emprats per els missatges viatjant entre clients i servidors. Els investigadors han distingit entre xarxes als Estats Units (US) i resta del món (no-US), segons la localització dels extrems. Llavors, tenen 364 observacions a US i 417 a no-US: la mitjana dels *hops* a US és de 18.64 (20.16 a no-US), i la desviació tipus és 13.6 (14.4 a no-US)

1.- (1.5 punts) Desenvolupeu una prova d'hipòtesis formal per demostrar si hi ha diferències en el nombre mitjà de *hops* entre ambdues poblacions. Recordeu definir correctament les hipòtesis, premisses, estadístic de la prova amb la seva distribució de probabilitat, zones d'acceptació i rebuig per a un risc del 5%, resolució i interpretació de la conclusió.

$H_0: \mu_{US} = \mu_{noUS}$

$H_1: \mu_{US} \neq \mu_{noUS}$

La premissa de la normalitat de les variables (nombre de *hops*) no és necessària, donat el elevat nombre d'observacions. Les variàncies poblacionals desconegudes les considerarem iguals. Les dues mostres són independents, i suposarem que cada mostra és una mostra aleatòria simple.

Estadístic de la prova: $\hat{t} = \frac{\overline{US} - \overline{noUS}}{S\sqrt{1/n1 + 1/n2}}$, si la hipòtesi nul·la és certa, $\hat{t}$ segueix una distribución t-Student amb n1+n2-2 graus de llibertat. Com que les mides mostrals seran grans, també podem aproximar la distribució de $\hat{t}$ a una $N(0,1)$.

Zona d'acceptació: H_0 no es rebutjarà si $|\hat{t}| < 1.96$; Resolució: càlcul de la variància comuna

$s^2 = \frac{(364-1)S_1^2 + (417-1)S_2^2}{364+417-2} = \frac{67140.48 + 86261.76}{779} = \frac{153402.2}{779} = 196.922$

$\hat{t} = \frac{20.16 - 18.64}{\sqrt{196.9}\sqrt{\frac{1}{364} + \frac{1}{417}}} = \frac{1.52}{1.01} = 1.51$

L'estadístic cau en zona d'acceptació. Això vol dir que la diferència entre mitjanes (1.52) no és molt més gran que l'error tipus (1.01) i, per tant, inclús si les mitjanes poblacionals fossin iguals (és a dir, el nombre esperat de *hops* fos el mateix en US que fora dels US), la diferència observada pot ser conseqüència de fluctuació aleatòria.

2.- (1 punt) Trobeu també el p-valor del resultat anterior i expliqueu formalment què significa (*no feu una interpretació pràctica, del tipus "si és major que ... llavors ..."*)

p-valor: $P(|Z| > |\hat{t}|) = 2\,P(Z > 1.51) = 2(1 - P(Z < 1.51)) = 0.1310$

El p-valor és la probabilitat amb que, suposant que les mitjanes poblacionals són iguals (és a dir, que les diferències entre les mitjanes mostrals es deuen únicament a l'atzar), es podria obtindré d'un estudi de similars dimensions un resultat tan o més extrem que l'obtingut. Un resultat més extrem és un resultat "suggestiu" de que la hipòtesi nul·la no és certa, com por exemple major distància entre les mitjanes observades. Si aquesta probabilitat és relativament gran significa que la hipòtesi H_0 és creïble perquè pot explicar les observacions, en canvi si és molt petita significa que és molt difícil poder observar un resultat com el obtingut (o més extrem) sota la hipòtesi de igualtat de mitjanes.

3.- (1 punt) Amb quins arguments es podria haver plantejat una prova unilateral, i amb quin objectiu?

NO es pot plantejar en base a les mitjanes mostrals. Les hipòtesis són prèvies a les dades, sempre. Ens podríem basar, per exemple, en el nostre coneixement respecte les xarxes del territori US i del de fora US. Com que les distàncies són més curtes als US que no a la resta del món i, previsiblement, hi hauria menys nodes entre dues extrems de la connexió, podríem hipotetitzar que la mitjana dels *hops* als US és inferior que fora US, donant joc a aquesta hipòtesi alternativa:

$H_1: \mu_{US} < \mu_{noUS}$

4.- (1.5 punts) En les dades de US s'han comptat 23 casos en els que el missatge ha estat rebutjat i va haver de ser reenviat de nou. En l'altre conjunt, això va ocórrer 40 vegades. Es pot dir que el rebuig és més freqüent a les xarxes de la resta del món? Estructureu la resposta d'acord al procediment de la prova d'hipòtesis.

$H_0: \pi_{US} = \pi_{noUS}$

$H_1: \pi_{US} < \pi_{noUS}$

On π és un paràmetre que s'interpreta com la proporció de missatges que han estat rebutjats i ser reenviats (segons es tracti dels US o de fora US). Com la pregunta diu literalment si "es pot dir <u>més</u> freqüent a ...", es planteja la hipòtesi alternativa com a unilateral. Estadístic de la prova:

$\hat{z} = \frac{P_1 - P_2}{\sqrt{P \cdot \frac{1-P}{n_1} + P \cdot \frac{1-P}{n_2}}}$, $\hat{z} \sim N(0,1)$

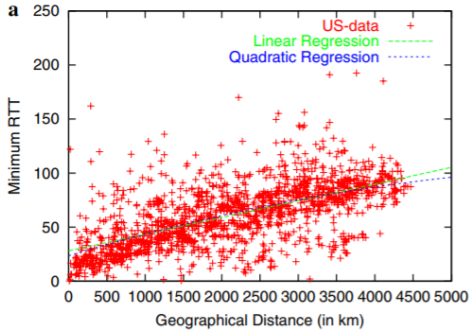
Sota la premissa de que tenim dues mostres independents. Es rebutjarà la hipòtesi nul·la si $\hat{z} > 1.645$ (punt que deixa un 5% de probabilitat per amunt a la distribució N(0,1)).

$P = \frac{23+40}{364+417} = 0{,}080666$, estimació de la proporció comuna, suposant que ni hi ha diferències entre els dos grups.

$\hat{z} = \frac{23/364 - 40/417}{\sqrt{0.08 \cdot \frac{1-0.08}{364} + 0.08 \cdot (1-0.08)/417}} = 0{,}0327/0{,}0195 = 1.676$

El valor de l'estadístic supera el punt crític, per tant es podria concloure que hi ha certa evidència que apunta a que fora dels US es rebutgen més missatges.

El treball d'aquells investigadors se centra en la relació entre la distància geogràfica entre origen i destí, i RTT (round trip time, ms), per tal de dirigir a l'usuari al servidor amb la menor latència. La correlació entre les dues variables a les dades US s'indica a l'article (0,6113), però no l'equació de la recta de regressió. No obstant això, hem pogut determinar que el centre del núvol de punts està en (2287km, 62.677ms). D'altra banda, visualment s'ha estimat que les desviacions respectives serien aproximadament 1300km i 33ms



5.- (1.5 punts) Amb aquesta informació, calculeu l'equació de la recta de regressió i estimi el temps RTT que trigaria una connexió entre dos punts separats 1500km.

$b_1 = r\frac{s_{RTT}}{s_D} = 0.6113\,\frac{33}{1300} = 0.0155$; $b_0 = \overline{RTT} - b_1\overline{D} = 62.677 - 0.0155\ 2287 = 27.19$

$RTT = 27.19 + 0.0155 \times D$.

Per D=1500, el temps estimat puntualment seria: 27.19+ 0.0155 ×1500 = 50.46 ms

6.- (1 punt) Calculeu també el valor de la desviació residual i el coeficient de determinació, explicant quina informació dóna cada un dels dos indicadors.

Desviació residual s: $s^2 = \frac{\sum e_i^2}{n-2} = \frac{(n-1)s_{RTT}^2(1-r^2)}{n-2} = \frac{(363)33^2(1-0.6113^2)}{362} = 683.94$; $s = 26.15\ ms$

Aquest valor ens quantifica la variabilitat de les observacions respecte la recta de regressió, o la variabilitat dels errors quan s'utilitza la recta per a predir el RTT a partir de la distància.

Coeficient de determinació: R^2és el valor de la correlació elevat al quadrat, 0.3737, o 37.37%

Aquest valor ens quantifica què proporció de la variabilitat de la resposta es pot explicar per la variable explicativa D. És una mesura de la qualitat de l'ajust en el model lineal. En aquest sentit, quan més a prop d'1 sigui el R^2, millor qualitat de l'ajust, en el sentit que els errors (diferència entre valor observat i el valor obtingut a la recta de regressió) seran menors.

7.- (1.5 punts) Interpreteu el valor del pendent i obtingui un interval de confiança a l'95% per al pendent poblacional.

El valor 0.0155 s'interpreta com: afegir una distància d'1 km implica un increment de 0.0155 ms al temps RTT

$$IC(\beta_1, 95\%) = b_1 \mp z_{0.975} \cdot \sqrt{\frac{s^2}{(n-1)\cdot S_D^2}} = 0.01552 \mp 1.96 \cdot \sqrt{\frac{683.94}{363 \cdot 1300^2}} = [0.01345, 0.01759]$$

8.- (1 punt) Indiqueu les premisses del model lineal, què impliquen i com es comproven. A la vista del gràfic indicat plantegeu per cadascuna de les premisses si seria assumible o no i per què.

L'ajust lineal és relativament bo però s'aprecia lleugerament una certa curvatura: a l'inici els punts es situen més per sota de la recta: de fet els autors han intentat ajustar també un polinomi quadràtic. Per tant, la premissa de linealitat podria no complir-se.

La premissa de normalitat també és dubtosa: sembla que els temps tenen una tendència a anar-se per la cua de la dreta (valors alts), diferent que per la cua de la esquerra (valors baixos), el que es traduiria en una distribució dels residus asimètrica.

En canvi, la variabilitat sembla constant i estable per a qualsevol distància, per tant la premissa d'homocedasticitat seria acceptable.

No podem dir gran cosa de la premissa d'independència, no es coneixen els detalls de com s'han obtingut les dades ni cap gràfic que pugui ajudar (com el de la seqüència temporal).