UNIVERSITE DE BATNA_2 **FACULTE DE MEDECINE**

SERIE DE TD N° 12 *en BIOSTATISTIQUE 2019/2020*

Exercices sur les tests de la régression linéaire

Formules rappelant le cours de la régression :

Remarque :

La droite de régression linéaire y = a + b $*$ x = A + B $*$ x où les symboles A = a et B = b sont les coefficients de la droite de régression linéaire de Y en X.

Dans la littérature on a plusieurs symboles, il suffit de connaître les définitions suivantes :

Somme des carrés des écarts totale : SCT = SCEt = SST = variabilité totale à expliquer.

Somme des carrés des écarts due au facteur ou au modèle ou à la régression : SCM = SCEfa = SSR (ici R désigne regression en anglais) = variabilité expliquée par le modèle en question.

Somme des carrés des écarts résiduelle = somme des carrés des résidus : SCR = SCEr = SSE = variabilité non expliquée par le modèle en question.

Formule de l'analyse de la variance : SCEt = SCEfa + SCEr

Le coefficient d'ajustement ou de détermination $r^2 = \frac{\text{SCEfa}}{\text{SCEt}}$ = (carré du coefficient de corrélation).

r^2 mesure la variabilité expliquée par le modèle en question.

Les seuils critiques sont fonction de test unilatéral gauche, droit et bilatéral :

- Test unilatéral gauche : le seuil critique est $-\mathbf{t}_{(\alpha\,;\,n-2)}$ distribution de Student
- Test unilatéral droit : le seuil critique est $+\mathbf{t}_{(1-\alpha\,;\,n-2)}$ distribution de Student
- Test bilatéral: la zone d'acceptation est $[-\mathbf{t}_{(\frac{\alpha}{2}\,;\,n-2)}\,;\,+\mathbf{t}_{(1-\frac{\alpha}{2}\,;\,n-2)}]$ distribution de Student

Hypothèse **: on donne ces égalités comme indiquées dans le cours.**

$S_{xx} = \sum(x_i - \bar{x})^2$ $S_{yy} = \sum(y_i - \bar{y})^2$ $S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y})$

$SSE = \sum(y_i - \hat{y}_i)^2$ $B = \frac{Sxy}{Sxx} = b$

Conclusion **: Retrouver et démontrer les relations suivantes.**

$S_{xx} = N\,\sigma_x^2$ $S_{yy} = N\,\sigma_y^2$ $S_{xy} = N\,\text{Cov}(x\,;\,y)$

$SSE = S_{yy} - b\,S_{xy} = N\,\sigma_y^2 - b\,(\sum x_i y_i - N\,\bar{x}\,\bar{y})$ $B = \frac{\text{Cov}(x\,;\,y)}{\text{Variance de x}} = b$

Solution

$S_{xx} = \sum(x_i - \bar{x})^2 = N * \frac{\sum(x_i - \bar{x})^2}{N} = N\,\sigma_x^2$ $S_{yy} = \sum(y_i - \bar{y})^2 = N * \frac{\sum(y_i - \bar{y})^2}{N} = N\,\sigma_y^2$

$S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = N * \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{N} = N\,\text{Cov}(x;\,y)$

$$B = \frac{Sxy}{Sxx} = \frac{[\sum(x_i - \bar{x})(y_i - \bar{y})]/n}{[\sum(x_i - \bar{x})^2]/n} = \frac{\text{Cov}(x\,;\,y)}{\text{Variance de x}} = b$$

$SSE = \sum(y_i - \hat{y}_i)^2 = \sum(y_i - (A + Bx_i))^2 = \sum(y_i - (\bar{y} - B\bar{x} + Bx_i))^2 =$

$\sum((y_i - \bar{y}) - B(x_i - \bar{x}))^2 = \sum(y_i - \bar{y})^2 + B^2\sum(x_i - \bar{x})^2 - 2B\sum(x_i - \bar{x})(y_i - \bar{y}) =$

$S_{yy} + B^2 S_{xx} - 2 B S_{xy} = S_{yy} + B \frac{Sxy}{Sxx} S_{xx} - 2 B S_{xy} = S_{yy} + B S_{xy} - 2 B S_{xy} = S_{yy} - B S_{xy} \Rightarrow$

$SSE = N \sigma_y^2 - B (\sum x_i y_i - N \bar{x}\ \bar{y}) = \sum y_i^2 - A \sum y_i - B \sum x_i\, y_i = SST - B S_{xy}$

$SSR = A \sum y_i + B \sum x_i\, y_i - \frac{(\sum y_i)^2}{n} = B S_{xy} = B\, n\, Cov\,(x\,;y)$

$SST = N \sigma_y^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = \sum y_i^2 - \bar{y} \sum y_i$

Exercice n° 01 :

Les données suivantes pour la variable dépendante y et la variable indépendante x, ont été recueillies à l'aide d'un échantillonnage aléatoire simple :

x	10	14	16	12	20	18	16	14	16	18
y	120	130	170	150	200	180	190	150	160	200

I)

a) Trouvez une équation de régression linéaire simple ($y = a + b * x$) de la série statistique double (X ; Y). Déterminez une estimation de y pour la valeur de $x = 17$.

b) Calculez la somme des carrés des résidus (SSE), la somme des carrés des écarts totale (SST) et le coefficient de détermination r^2.

c) Calculez l'erreur type de l'estimation : $\hat{\sigma}$.

II) On considère la droite de régression $E\,(Y \mid x) = \alpha + \beta * x$ de la série aléatoire double.

a) Tester l'hypothèse $H_0 = \{\beta = 0\}$ contre $H_1 = \{\beta > 0\}$ en utilisant un risque de 5 %.

b) Estimer la valeur moyenne de y pour $x = 17$ en prenant un risque de 5 %.

Solution :

I) En considérant la série statistique double (X ; Y), la calculatrice donne les résultats suivants :

$\bar{x} = 15.4$ $\bar{y} = 165$ $\sigma_x^2 = 8.04$ $\sigma_x = 2.835489376$ $\sigma_y = 26.55183609$

$\sigma_y^2 = 705$ $\sum x_i = 154$ $\sum y_i = 1650$ $\sum x_i\, y_i = 26\,080$ $\sum y_i^2 = 279\,300$

$A = 36.66666667$ $B = 8.333333333$ $r = 0.889922566$

a) La droite de régression linéaire $y = a + b * x = 36.66667 + 8.33333 * x \approx \mathbf{36.67 + 8.33 * x}$

L'estimation de y pour $x = 17$: $\hat{y}\,(17) = 36.66667 + 8.33333 * 17 = \mathbf{178.3333333}$

b) Calcul de la somme des carrés des résidus : $SSE = \sum y_i^2 - A\sum y_i - B\sum x_i\, y_i = \mathbf{1466.666667}$

Calcul de la somme des carrés des écarts totale : $SST = n\, \sigma_y^2 = \mathbf{7\,050}$

Calcul du coefficient de détermination $\Rightarrow r^2 = 0.889922566^2 = \mathbf{0.791962174}$

c) Calcul de $\hat{\sigma} = \sqrt{\frac{SSE}{n-2}} = \sqrt{183.3333333} = \mathbf{13.540064}$

II) En utilisant la droite $E\,(Y \mid x) = \alpha + \beta * x$ de la série aléatoire double

a) Tester l'hypothèse $H_0 = \{\beta = 0\}$ contre $H_1 = \{\beta > 0\}$; c'est un test unilatéral (égal contre supérieur)

Le risque étant de 5 % $\Rightarrow$ le seuil critique est $t_{(0.05\,;\,8)} = 1.86$ si $To > 1.86$ on rejette H_0.

Calcul de $Se\,(B) = \frac{\hat{\sigma}}{\sqrt{S_{xx}}} = \frac{13.540064}{\sqrt{80.4}} = 1.510054747$

Calcul de $To = \frac{B}{\sqrt{\frac{\hat{\sigma}^2}{Sxx}}} = \frac{8.333333}{1.510055} = 5.51856$

Décision : Comme $To > t_{(0.05\,;\,8)}$ on rejette H_0 et **on conclut que la pente $\beta > 0$.**

b) Estimation de la valeur moyenne de y lorsque x prend la valeur 17 avec un risque de 5 %.

$E\,(Y| x) = \hat{y}\,(17) \pm t_{(0.05\,;\,8)} \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}\right]} = [166.9961563\,;\,189.6705104] \approx \mathbf{[167\,;\,189.7]}$

Exercice n° 02 :

Une analyse de régression linéaire à partir d'un échantillon de taille n = 15 a produit ce qui suit :

$\sum(x_i - \bar{x})(y_i - \bar{y}) = 156.4$ $\qquad$ $\sum(x_i - \bar{x})^2 = 173.5$

$\sum(y_i - \bar{y})^2 = 181.6$ $\qquad$ $\sum(y_i - \hat{y})^2 = 40.621$

$\bar{x} = 13.4$ and $\bar{y} = 56.4$

I)

a) Déterminez la droite de régression linéaire de Y en X (y = a + b ∗ x) de la série statistique double (X ; Y).

b) Déterminer le coefficient de corrélation linéaire r.

II) En considérant la droite de régression E (Y| x) = α + β ∗ x de la série aléatoire double,

a) Tester l'existence d'une corrélation linéaire entre la variable aléatoire Y et la variable aléatoire X, par deux méthodes :

i. On teste les hypothèses $H_0 : \{\rho = 0\}$ contre $H_1 : \{\rho \neq 0\}$. Le risque est de 5 %.

ii. On teste les hypothèses $H_0 : \{\beta = 0\}$ contre $H_1 : \{\beta \neq 0\}$. Le risque est de 5 %.

b) Estimer les valeurs moyennes de Y pour X = 8.3 et X = 18.5 en prenant un risque de 5 %.

Solution :

I) A partir des données on aura la droite de régression linéaire et le coefficient de corrélation

a) *La droite de régression* : y = A + B ∗ x

Où $B = \frac{Sxy}{Sxx} = \frac{[\sum(x_i - \bar{x})(y_i - \bar{y})]}{[\sum(x_i - \bar{x})^2]} = \frac{156.4}{173.5} = 0.901440922$

Et $A = \bar{y} - B\,\bar{x} = 56.4 - \frac{156.4}{173.5} * 13.4 = 44.32069164$

y = 44.32069164 + 0.901440922 ∗ x $\qquad \Rightarrow \qquad$ **y = 44.32 + 0.9 ∗ x**

b) *Le coefficient de corrélation* :

Détermination de $\sigma_x^2 = \frac{\sum(x_i - \bar{x})^2}{15} = 11.56666667 \Rightarrow \sigma_x = 3.400980251$

$\sigma_y^2 = \frac{\sum(y_i - \bar{y})^2}{15} = 12.10666667 \Rightarrow \sigma y = 3.47946356$

$Cov\,(x; y) = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{15} = 10.42666667$

$r = \frac{Cov\,(x\,;y)}{\sigma x\ \ \sigma y} = \frac{10.4266667}{3.40098 * 3.47946} = 0.881107883 \Rightarrow$ **r ≈ 88.11 %**

II) A partir de la droite E (Y| x) = α + β ∗ x de la série aléatoire double, on a

a) Existence d'une liaison entre les deux variables aléatoires X et Y.

i) Test d'hypothèse sur la corrélation linéaire ρ

$H_0 = \{\rho = 0\}$; contre $H_1 = \{\rho \neq 0\}$, avec un risque de 5 %

C'est un test bilatéral (égal contre inégal) et le seuil critique $t_{(0.05\,;\,13)} = 2.16$

Calcul de $To = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.881107883}{\sqrt{\frac{1 - 0.881107883^2}{15-2}}} = 6.717650152 \Rightarrow$ **To ≈ 6.72**

Décision : comme **|To| = 6.72 > 2.16** on rejette H_0 et on conclut que $\rho \neq 0$.

ii) Test d'hypothèse sur la pente β

$H_0 = \{\beta = 0\}$; contre $H_1 = \{\beta \neq 0\}$, avec un risque de 5 %

C'est un test bilatéral (égal contre inégal) et le seuil critique t $_{(0.05\ ;\ 13)}$ = 2.16

$$Se(B) = \frac{\hat{\sigma}}{\sqrt{S_{xx}}} = \frac{\sqrt{\frac{SSE}{n-2}}}{\sqrt{S_{xx}}} = \frac{\sqrt{\frac{40.621}{15-2}}}{\sqrt{173.5}} = 0.134200424 \qquad B = 0.901440922$$

$$\text{Calcul de } T_o = \frac{B-\beta}{Se_{(B)}} = \frac{0.901440922-0}{0.134200424} = 6.717124222 \Rightarrow \mathbf{T_o \approx 6.72}$$

Décision : comme $|\mathbf{T_o}|$ **= 6.72 > 2.16** on rejette H_0 et **on conclut que** $\boldsymbol{\beta \neq 0}$.

b) Estimation des valeurs moyennes de Y pour X = 8.3 et X = 18.5 avec un risque de 5 %

$\hat{y}$ (8.3) = 44.32069164 + 0.901440922 $*$ 8.3 = 51.80265129

$\hat{\sigma} = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{40.621}{15-2}} = 1.767679922 \qquad \bar{x} = 13.4 \qquad S_{xx} = 173.5$

$E(Y|x) = \hat{y}(8.3) \pm t_{(0.05\ ;\ 8)} \sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}\right]}$ = [51.80265129 $\pm$ 1.776915463] =

[50.02573583 ; 53.5796675] ≈ **[50.03 ; 53.58]**

$\hat{y}$ (18.5) = 44.32069164 + 0.901440922 $*$ 18.5 = 60.9973487

$E(Y|x) = \hat{y}(18.5) \pm t_{(0.05\ ;\ 8)} \sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}\right]}$ = [60.9973487 $\pm$ 1.776915463] =

[59.22043324 ; 62.77426416] ≈ **[59.22 ; 62.77]**

Remarque : ***ces deux intervalles ont les mêmes amplitudes.***

Exercice n° 03 :

Dans le but d'étudier le lien entre la quantité de gaz de ville (Y en m^3) consommée par mois dans un hôpital et la température ambiante moyenne (X en degré F) pour le mois considéré, on a recueilli les données dans le tableau suivant :

Mois	Temp. x_i °F	y_i en m^3	Mois	Temp. x_i °F	y_i en m^3
Janvier	21	185.79	Juillet	68	621.55
Février	24	214.47	Août	74	675.06
Mars	32	288.03	Septembre	62	562.03
Avril	47	424.84	Octobre	50	452.93
Mai	50	454.58	Novembre	41	369.95
Juin	59	539.03	Décembre	30	273.98

I)

a) Calculer le coefficient de corrélation r entre la quantité de gaz de ville consommée par mois et la température ambiante moyenne pour le même mois.

b) Déterminez l'équation de la droite de régression linéaire de Y en X (y = a + b $*$ x). Déterminez la valeur de $\hat{\sigma}^2$.

c) Calculer la variation de Y correspondant à une variation de X de 1 °F.

d) Quelle quantité de gaz de ville doit-on prévoir lorsque la température ambiante moyenne d'un mois serait de 55 °F ?

II)

a) Testez la signification de la régression par deux méthodes différentes en considérant un risque de 5 %.

b) Testez l'hypothèse $H_0 = \{\alpha = 0\}$ contre $H_1 = \{\alpha \neq 0\}$ en prenant un risque de 10 %.

Solution :

I) En utilisant la calculatrice avec les données de la série statistique double (X ; Y), on a :

$\bar{x} = 46.5$ $\bar{y} = 421.8533333$ $\sigma_x = 16.60572191$ $\sigma_y = 152.9217843$ $\sum x_i\, y_i = 265.86463$

Cov (x ; y) = 2 539.205833 $\sum x_i = 558$ $\sum y_i = 5\,062.24$ $\sum x_i^2 = 29\,256$

$\sum y_i^2 = 2\,416\,143.684$ A = – 6.335501662 B = 9.208362043 r = 0.999932743

a) Le coefficient de corrélation r = 0.999932743

b) La droite de régression linéaire est y = A + B ∗ x = – 6.335501662 + 9.208362043 ∗ x

y ≈ – 6.336 + 9.208 ∗ x

SSE = $\sum y_i^2 - A \sum y_i - B \sum x_i\, y_i = 37.74608894$ n – 2 = 10

$\hat{\sigma}_y^2 = \frac{SSE}{n-2} = \frac{37.7460889}{12-2} = 3.774608894 \Rightarrow \boldsymbol{\hat{\sigma}_y = 1.942835272}$

c) Calcul de la variation de Y correspondant à une variation de X de 1 °F

Si $x = x_1 \Rightarrow$ la valeur correspondante de y est $\hat{y}(x_1) = a + b * x_1$

Une variation de x d'une unité change la valeur de x de x_1 à $x_2 = x_1 + 1$ et ainsi la valeur de y varie de $\hat{y}(x_1)$ à $\hat{y}(x_2) = \hat{y}(x_1 + 1) = a + b * (x_1 + 1) = a + b * x_1 + b = \hat{y}(x_1) + b$

La variation de y est donc = $\hat{y}(x_2) - \hat{y}(x_1)$ = **b = 9.208362043 ≈ 9.21**

⇒ On remarque qu'à chaque changement de température de 1°F, il y a une augmentation dans le gaz d'une quantité égale à la valeur de la pente b.

d) L'estimation de y quand x vaut 55 °F ⇒ $\hat{y}(55) = 500.1244107 = A + 55 * B \approx 500$.

II) Tests d'hypothèses.

a) Test sur la signification de la régression linéaire par deux méthodes

i) Test d'hypothèse sur la corrélation linéaire ρ

$H_0 = \{\rho = 0\}$; contre $H_1 = \{\rho \neq 0\}$, avec un risque de 5 %

C'est un test bilatéral (égal contre inégal) et le seuil critique $t_{(0.05\,;\,10)} = 2.228$

Calcul de To = $\frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.999932743}{\sqrt{\frac{1-0.999932743^2}{12-2}}} = 272.6432752 \approx 273$

Décision : comme |To| ≈ 273 > 2.16 on rejette H_0 et **on conclut que ρ ≠ 0**

ii) Test d'hypothèse sur la pente β

$H_0 = \{\beta = 0\}$; contre $H_1 = \{\beta \neq 0\}$, avec un risque de 5 %

C'est un test bilatéral (égal contre inégal) et le seuil critique $t_{(0.05\,;\,10)} = 2.228$

$Se(B) = \frac{\hat{\sigma}}{\sqrt{S_{xx}}} = \frac{\sqrt{\frac{SSE}{n-2}}}{\sqrt{S_{xx}}} = \frac{\mathbf{1.942835272}}{\sqrt{3\,309}} = 0.033774396$ B = 9.208362043

Calcul de To = $\frac{B-\beta}{Se_{(B)}} = \frac{9.208362043 - 0}{0.033774396} = 272.6432752 \Rightarrow$ **To ≈ 273**

Décision : comme **|To| = 273 > 2.228** on rejette H_0 et **on conclut que β ≠ 0.**

b) Testez l'hypothèse $H_0 = \{\alpha = 0\}$ contre $H_1 = \{\alpha \neq 0\}$ en prenant un risque de 10 %.

C'est un test bilatéral (égal contre inégal) ; le seuil critique $t_{(0.05\,;\,10)} = 2.228 \Rightarrow$

Si To > 2.228 ou To < – 2.228 alors il y a Rejet de H_0

Calcul de Se(A) = $\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}} = 1.942835272\sqrt{\frac{1}{12} + \frac{46.5^2}{3309}} = 1.667648229$

Calcul de la statistique de test To = $\frac{A}{Se(A)} = \frac{-6.33501662}{1.667648229} = -3.799063587$

Comme **To = – 3.799 < – 2.228** alors on doit rejeter H_0. **On conclut que α ≠ 0.**

Exercice n° 04 :

Soient les données de l'échantillon suivant donnant les valeurs y et x pour 28 sujets :

Sujets	y	x	Sujets	y	x
1	10	2205	15	6	1901
2	11	2096	16	5	2288
3	11	1847	17	5	2072
4	13	1903	18	5	2861
5	10	1457	19	6	2411
6	11	1848	20	4	2289
7	10	1564	21	3	2203
8	11	1821	22	3	2592
9	4	2577	23	4	2053
10	2	2476	24	10	1979
11	7	1984	25	6	2048
12	10	1917	26	8	1786
13	9	1761	27	2	2876
14	9	1709	28	0	2560

I)

a) Trouvez la droite de régression linéaire de Y en X par la méthode des moindres carrés ordinaires. Déterminer la valeur de $\hat{\sigma}$.

b) Trouvez la prédiction ($\hat{y}$) si x prend la valeur de 1800.

c) Quelle est la variation attendue de y associée à une diminution de 100 unités de x ?

d) Pour augmenter de 1 la valeur de y, quelle diminution doit être générée par x ?

e) Étant donné que x = 1917, trouver la valeur estimée de y et le résidu correspondant.

II) Dans les questions suivantes on prend un risque de 5 % pour a – c – d et 1 % pour b.

a) Testez les hypothèses $H_0 = \{\rho = 0\}$ contre $H_1 = \{\rho \neq 0\}$.

b) Testez les hypothèses $H_0 = \{\rho = -0.7\}$ contre $H_1 = \{\rho \neq -0.7\}$.

c) Testez les hypothèses $H_0 = \{\beta = 0\}$ contre $H_1 = \{\beta < 0\}$.

d) Testez les hypothèses $H_0 = \{\alpha = 22\}$ contre $H_1 = \{\alpha \neq 22\}$.

Solution :

I) En utilisant la calculatrice avec les données de la série statistique double (X ; Y), on a :

$\bar{x} = 2110.142857$ $\bar{y} = 6.964285714$ $\sigma_x = 128878.9796$ $\sigma_y = 11.67729592$

$\sum x_i y_i = 386\,127$ $A = 21.78825088$ $B = -0.0070251$ $r = -0.738027303$

$\sum x_i = 59\,084$ $\sum y_i = 195$ $\sum x_i^2 =$ $\sum y_i^2 = 1\,685$

a) La droite de régression linéaire est $y = A + B * x = 21.78825088 - 0.0070251 * x$

$$\mathbf{y \approx 21.7883 - 0.00703 * x}$$

b) $\hat{y}\,(1800) = A + 1800 * B = 9.1430709$ (1)

c) Une diminution de x de 100 unités change la valeur de x de x_1 à $x_2 = x_1 - 100$.

Par conséquent la valeur de y change de $\hat{y}\,(x_1)$ à $\hat{y}\,(x_2) = \hat{y}\,(x_1 - 100) = a + b * (x_1 - 100)$

$= a + b * x_1 - 100b = \hat{y}\,(x_1) - 100b \Rightarrow$

La variation de y est donc $\hat{y}\,(x_2) - \hat{y}\,(x_1) = \mathbf{-100b = 0.70251 \approx 0.7}$

d) Les valeurs $\hat{y}\,(x_1)$ et $\hat{y}\,(x_2)$ correspondant à deux valeurs x_1 et x_2 sont données par les expressions$\hat{y}\,(x_1) = a + b * x_1$ et $\hat{y}\,(x_2) = a + b * x_2$.

Une augmentation de la valeur de y d'une unité résultant d'un changement de la valeur de x de x_1 à x_2 équivaut à $\hat{y}$ (x_2) – $\hat{y}$ (x_1) = 1 ⇒ [a + b ∗ x_2] – [a + b ∗ x_1] = b ∗ (x_2 – x_1)

Ainsi on a : **$x_2 - x_1 = b^{-1} = -0.0070251^{-1} = -142.3467282 \approx -142.35$** donc pour augmenter la valeur de y d'une unité, x doit générer une diminution de 142.35 unités.

e) $\hat{y}$ (1917) = A + 1917 ∗ B = 8.3211342 ⇒

Le résidu = y(1917) – $\hat{y}$ (1917) = 10 – 8.3211342 = **1.6788658**

II) La calculatrice a donné r = – 0.738027303

a) Test des hypothèses : $H_0 = \{\rho = 0\}$ contre $H_1 = \{\rho \neq 0\}$ avec un risque de 5 % ⇒
C'est un test bilatéral (égal contre inégal) ; $t_{(0.05\,;\,26)} = 2.056$ si $|To| > 2.056$ on rejette H_0.

Calcul de la statistique de test To = $r\sqrt{\frac{n-2}{1-r^2}}$ = 5.577027526 ⇒ **To ≈ 5.577**

Comme **$|To|$ = 5.577 > 2.056** on rejette H_0 et **les variables (x et y) sont liées**.

b) Test des hypothèses $H_0 = \{\rho = -0.7\}$ contre $H_1 = \{\rho < -0.7\}$ avec un risque de 1 %.
C'est un test unilatéral gauche : (égal conte inférieur) ; $t_{(0.01\,;\,26)} = -2.479$

Si To < – 2.479 on rejette H_0.

To = $[\frac{1}{2}\,\text{Ln}\,\frac{1+r}{1-r} - \frac{1}{2}\,\text{Ln}\,\frac{1+\rho}{1-\rho}]\sqrt{n-3}$ = – 0.078832337 ≈ **– 0.0788**

__**Décision**__ : comme **To = – 0.0788 > $t_{(0.01\,;\,26)}$ = – 2.479** alors on ne peut pas rejeter H_0 et on conclut que le coefficient de corrélation ρ est égal à – 0.7

c) Test des hypothèses $H_0 = \{\beta = 0\}$ contre $H_1 = \{\beta < 0\}$; avec un risque de 5 %
Test unilatéral (égal contre inférieur) ; $t_{(0.05\,;\,26)} = -1.706$ si To < – 1.706 on rejette H_0.
SSE = $\sum y_i^2 - A\sum y_i - B\sum x_i\, y_i$ = 148.8719722 ; S_{xx} = 3 608 611.429 ; n = 28

$\hat{\sigma} = \sqrt{\frac{SSE}{n-2}} = \sqrt{5.725845085} = 2.392873813$

Se (B) = $\frac{\hat{\sigma}}{\sqrt{Sxx}} = \frac{2.392873813}{\sqrt{3\,608\,611.429}}$ = 0.001259649553

Calcul de la statistique de test To = $\frac{B}{S(B)} = \frac{-0.0070251}{0.0012599649553}$ = – 5.577027564 ≈ **– 5.577**

__**Décision**__ : Comme la statistique de test **To = – 5.577 < $t_{(0.05\,;\,26)}$ = – 1.706 ⇒ rejeter H_0**.

d) Test des hypothèses $H_0 = \{\alpha = 22\}$ contre $H_1 = \{\alpha \neq 22\}$; avec un risque de 5 %

Test unilatéral (égal contre différent) ; $t_{(0.05\,;\,26)} = 2.056$ si $|To| < 2.056$ on rejette H_0.

Se(A) = $\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{Sxx}} = 2.392873813\sqrt{\frac{1}{28} + \frac{2110.142857^2}{3608611.429}}$ = 2.696233261

Calcul de la statistique de test To :

To = $\frac{A-\alpha}{Se(A)} = \frac{21.788251-22}{2.696233261}$ = – 0.078535162 ⇒ **To ≈ – 0.0785**

__**Décision**__ : Comme **To = – 0.0785 ∈ [– 2.056 ; + 2.056] on ne peut pas rejeter H_0**.