Université A. Mira -Bejaia
Faculté des Sciences de Gestions, commerciales et économiques
Département des Sciences économiques

Niveau : $L3$ EMB

BERKAI née MECHERI Kheira

Chapitre 2 : Régression Linéaire multiple

Le modèle de régression linéaire multiple est une extension du modèle de régression simple lorsque les variables explicatives sont en nombre quelconque (c-à-d la relation entre la variable dépendante Y avec plusieurs variables indépendantes X_i, ce qui va nous permettre de faire ensuite des prévisions de Y lorsque $X_i$est mesurée.

0.1 Présentation du modèle

Nous supposons donc que les données collectées suivent le modèle défini par :

$$Y_t = B_0 + B_1 X_{1t} + B_2 X_{2t} + B_3 X_{3t} + ... + B_k X_{kt} + \varepsilon_t, \quad t = \overline{1, n} \tag{1}$$

où

- Y_t est la variable dépendante ou endogène (une variable à expliquer) au temps t.
- B_0, ..., B_k sont les coefficients de la régression.
- X_{it} est la variable indépendante ou exogène i (variable explicative) au temps t.
- ε_t est une erreur aléatoire.
- n est le nombre de variables explicatives non aléatoires.
- n est le nombre d'observations.

0.2 la forme matricielle

En utilisant l'écriture précédente et en écrivant le modèle observation par observation, nous obtenons :

$$Y_1 = B_0 + B_1X_{11} + B_2X_{21} + B_3X_{31} + ... + B_kX_{k1} + \varepsilon_1$$
$$Y_2 = B_0 + B_1X_{12} + B_2X_{22} + B_3X_{32} + ... + B_kX_{k2} + \varepsilon_2$$
$$...$$
$$...$$
$$...$$
$$Y_n = B_0 + B_1X_{1n} + B_2X_{2n} + B_3X_{3n} + ... + B_kX_{kn} + \varepsilon_n$$

Ce qui nous permis d'obtenir la forme matricielle suivante :

$$\underset{(n,1)}{Y} = \underset{(n,k+1)}{X}\ \underset{(k+1,1)}{B} + \underset{(n,1)}{\varepsilon} \tag{2}$$

où

- Y est un vecteur aléatoire de dimension n,
- X est une matrice de taille $n \times (k+1)$ connue, appelée matrice du plan d'expérience,
- B est le vecteur des paramètres inconnus du modèle,
- ε est le vecteur des erreurs.

Avec

$$\underset{(n,1)}{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{pmatrix} ; \underset{(n,k+1)}{X} = \begin{pmatrix} 1 & X_{11} & X_{21} & . & . & . & X_{k1} \\ 1 & X_{12} & X_{22} & . & . & . & X_{k2} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & X_{1n} & X_{2n} & . & . & . & X_{kn} \end{pmatrix} ; \underset{(k+1,1)}{B} = \begin{pmatrix} B_0 \\ B_1 \\ . \\ . \\ . \\ B_k \end{pmatrix} ; \underset{(n,1)}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}.$$

Ce modèle doit satisfaire les hypothèses suivantes :

H_1 : les valeurs X_{it} sont observées sans erreurs.
H_2 : La moyenne des erreurs s'annulent, le modèle est bien spécifié, $E(\varepsilon_t) = 0$.
H_3 : La variance de l'erreur est constante et ne dépend pas des observations : homoscédasticité, $Var(\varepsilon_t) = \sigma_\varepsilon^2$.
H_4 : L'erreur est indépendante des variables exogènes, $COV(\varepsilon_t, X_{it}) = 0$.
H_5 : Indépendance des erreurs, les erreurs relatives à 2 observations sont indépendantes

(on dit aussi que les erreurs sont non corrélées), $COV(\varepsilon_t, \varepsilon_s) = 0$.
H_6 : $(X^{'}X)$ est régulière et inversible ($(X^{'}X)^{-1}$ existe).
H_7 : $\frac{(X^{'}X)}{n}$ tend vers une matrice finie non singulière
H_8 : $n > k + 1$ Nombre d'observations est supérieur aux nombre des séries explicatives.

0.3 Estimation des paramètres par la méthode (MCO)

Comme pour la régression linéaire simple, nous allons estimer le vecteur des paramètres B par la méthode Moindres carrés ordinaires (MCO), en supposant que : $Y = X\ B\ +\ \varepsilon$. Pour cela nous minimisons la somme des carrés des erreurs.

$$\min \sum_{t=1}^{n} \varepsilon^2 = \min \left(\varepsilon^{'} \varepsilon\right) = \min \left(Y - X\ \widehat{B}\right)^{'} \left(Y - X\ \widehat{B}\right) = \min S \tag{3}$$

Et

$$S = \left(Y - X\ \widehat{B}\right)^{'} \left(Y - X\ \widehat{B}\right) = \left(Y^{'}Y - Y^{'}X\widehat{B} - \widehat{B}^{'}X^{'}Y + \widehat{B}^{'}X^{'}X\widehat{B}\right) = \left(Y^{'}Y - 2\widehat{B}^{'}X^{'}Y + \widehat{B}^{'}X^{'}X\widehat{B}\right)$$

avec $\varepsilon^{'}$ est le transposé du vecteur ε. Pour minimiser S, il suffit de différencier S par rapport au vecteur B (résoudre cette équation $\frac{\partial S}{\partial B} = 0$) et on obtient :

$$\frac{\partial S}{\partial B} = -2X^{'}Y + 2X^{'}X\ \widehat{B} = 0.$$

Par conséquent, on obtient :

$$\widehat{B} = (X^{'}X)^{-1}X^{'}Y. \tag{4}$$

Avec

$$(X^{'}X) = \begin{pmatrix} n & \sum X_{1t} & \sum X_{2t} & \ldots & \sum X_{kt} \\ \sum X_{1t} & \sum X_{1t}^2 & \sum X_{2t}X_{1t} & \ldots & \sum X_{kt}X_{1t} \\ \sum X_{2t} & \sum X_{2t}X_{1t} & \sum X_{1t}^2 & \ldots & \sum X_{kt}X_{2t} \\ . & . & . & . & \\ . & . & . & . & \\ . & . & . & . & \\ \sum X_{kt} & \sum X_{1t}X_{kt} & \sum X_{2t}X_{kt} & \ldots & \sum X_{kt}^2 \end{pmatrix}$$

Et

$$(X'Y) = \begin{pmatrix} \sum Y_t \\ \sum X_{1t}Y_t \\ . \\ . \\ . \\ \sum X_{kt}Y_t \end{pmatrix}$$

Cette solution existe car $(X'X)^{-1}$ existe par hypothèse.

Remarque

1. Si les variables sont centrées, alors $frac(X'X)n =$ *matrice de variance covariance*

2. Si les variables sont centrées et réduites $frac(X'X)n =$ matrice de corrélation

0.4 Propriétés des estimateurs

L'estimateur est Sans biais

Comme en regression simple, l'estimateur obtenu est sans biais. Car :

$$\widehat{B} = (X'X)^{-1}X'Y = (X'X)^{-1}X'(XB + \varepsilon) = B + (X'X)^{-1}X'\varepsilon$$

D'après l'hypothèse H_2, on déduit que $E[\widehat{B}] = B$

Matrice de variance covariance de $\widehat{B}$

On appelle la matrice de variance covariance du vecteur aleatoire la matrice de dispersion, qui est donnée par :

$\Omega_{\widehat{B}} = E\left[\left(B - t\widehat{B}\right)\left(B - t\widehat{B}\right)'\right]$

$$\Omega_{\widehat{B}} = \begin{pmatrix} V(\widehat{B_0}) & Cov(\widehat{B_0}, \widehat{B_1}) & ... \\ & V(\widehat{B_1}) & \\ & & V(\widehat{B_k}) \end{pmatrix}$$

la variance de l'estimateur de chaque coefficient, se trouve sur la diagonale de $\Omega_{\widehat{B}}$.

On sait que :

$\widehat{B} = B + (X'X)^{-1}X'\varepsilon$, ce qui implique que

$$\Omega_{\widehat{B}} = E\left[(X'X)^{-1}X'\varepsilon\left((X'X)^{-1}X'\varepsilon\right)'\right] = E\left[(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}\right]$$
$$= (X'X)^{-1}X'E\left[\varepsilon\varepsilon'\right]X(X'X)^{-1}$$

Et

$$E\left[\varepsilon\varepsilon'\right] = \begin{pmatrix} E[\varepsilon_1^2] & E[\varepsilon_1\varepsilon_2] & \dots & E[\varepsilon_1\varepsilon_n] \\ & E[\varepsilon_2^2] & & \vdots \\ & & & \vdots \\ & & & E[\varepsilon_n^2] \end{pmatrix}$$

Et d'après les hypothèses H_2 et H_3, on déduit que :

$$E\left[\varepsilon\varepsilon'\right] = \sigma_\varepsilon^2\mathbf{I}_n$$

D'où

$$\Omega_{\widehat{B}} = \sigma_\varepsilon^2(X'X)^{-1} \tag{5}$$

Cette matrice tend vers la matrice nulle (toutes les cellules à 0) lorsque $n \to +\infty$ (hypothèses H_7), ce qui signifie que les estimateurs des moindres carrés sont convergents et à variance minimale .

Remarque

Dans la plupart des cas la variance des erreurs (σ_ε^2) est inconnue.

Détermination d'un estimateur Sans biais de la variance de l'erreur σ_ε^2

L'estimateur Sans biais de σ_ε^2 est donnée par :

$$\widehat{\sigma}_\varepsilon^2 = \frac{\sum_{t=1}^{n} e_t^2}{(n-k-1)} \tag{6}$$

Preuve

Notons que $\sum_{t=1}^{n} e_t^2 = e'e$

Avec $e = Y - \widehat{Y}$, et $\widehat{Y} = X\widehat{B}$

Alors $e = \varepsilon - X\left(\widehat{B} - B\right) = \left[\mathbf{I}_n - X(X'X)^{-1}X'\right]\varepsilon = \Gamma\varepsilon$

On peut facilement montrer que la matrice Γ est symétrique ($\Gamma' = \Gamma$, car $(X'X)$ est symétrique) et elle est idempotente d'ordre 2 ($\Gamma^2 = \Gamma$), de taille (n, n).

On déduit, alors :

$\sum_{t=1}^{n} e_t^2 = \varepsilon'\Gamma\varepsilon$, et

$E\left[\sum_{t=1}^{n} e_t^2\right] = \sigma_\varepsilon^2 \mathbf{I}_n Tr(\Gamma)$ ($Tr(\Gamma)$ est la trace de la matrice Γ).

$E\left[\sum_{t=1}^{n} e_t^2\right] = \sigma_\varepsilon^2 \mathbf{I}_n (n-k-1)$. Par conséquent, on peut considérer l'estimateur sans biais de la variance de l'erreur suivant :

$$\widehat{\sigma}_\varepsilon^2 = \frac{\sum_{t=1}^{n} e_t^2}{(n-k-1)} \tag{7}$$

Remarque

D'après le résultat précédent, on déduit que l'estimateur de la matrice variance covariane est donné par :

$$\widehat{\Omega}\varepsilon = \widehat{\sigma}_\varepsilon^2 (X'X)^{-1} \tag{8}$$

Où, les estimateurs des variances des paramètres du modèle sont sur la diagonale de la matrice $\widehat{\Omega}\varepsilon$

0.5 Equation d'analyse de la variance et la qualité de l'ajustement

0.5.1 Equation d'analyse de la variance

L'equation d'analyse de la variance est donnée par :

$$\sum_{t=1}^{n} \underbrace{\left(Y_t - \overline{Y}\right)^2}_{SCT} = \sum_{t=1}^{n} \underbrace{\left(Y_t - \widehat{Y}_t\right)^2}_{SCR} + \sum_{t=1}^{n} \underbrace{\left(\widehat{Y}_t - \overline{Y}\right)^2}_{SCE} \tag{9}$$

SCT : variabilité totale

SCE : Variabilité expliquée par le modèle

SCR : Variabilité non-expliquée (Variabilité résiduelle)

Les estimateurs sont d'autant plus précis lorsque La variance de l'erreur est faible et quand la dispersion des X est forte.

Tableau d'analyse de la variance (ANOVA)

Source de variation	Somme des carrés	Degrés de liberté	Carré moyen
Régression $X_1,\ ,X_2,...,\ X_k$	$SCE = \sum_{t=1}^{n}\left(\widehat{Y}_t - \overline{Y}\right)^2$	k	$\frac{SCE}{k}$
Résiduelle	$SCR = \sum_{t=1}^{n}\left(Y_t - \widehat{Y}_t\right)^2$	$n-k-1$	$\frac{SCR}{n-k-1}$
Totale	$SCT = \sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)^2$	$n-1$	

0.5.2 La qualité de l'ajustement

Le coefficient de détermination R^2

Les valeurs données par l'equation d'analyse de la variance dépendent des unités de mesure, c'est pourquoi on préfère utiliser le nombre sans dimension R^2.

$$R^2 = \frac{\sum_{t=1}^{n}\left(\widehat{Y}_t - \overline{Y}\right)^2}{\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)^2} = 1 - \frac{\sum_{t=1}^{n}\left(Y_t - \widehat{Y}_t\right)^2}{\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)^2} \tag{10}$$

ou

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT} = 1 - \frac{\sum_{t=1}^{n} e_t^2}{SCT} \tag{11}$$

Le coefficient de détermination R^2, mesure la proportion de la variance de Y expliquée par la régression de Y sur X. La qualitée de l'ajustement est meilleur quand R^2 est très proche de 1.

Remarque

Dans le cas de données centrées le coefficient de détermination R^2 est définé comme suit :

$$R^2 = \frac{\widehat{Y}'\widehat{Y}}{Y'Y} = 1 - \frac{\widehat{e}'\widehat{e}}{Y'Y} = 1 - \frac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n} Y_t^2} \tag{12}$$

Le coefficient de détermination corrigé $\overline{R}^2$

D'aprés la formule du coefficient de détermination R^2, nous constatons qu'il ne tient pas compte ni du nombre d'observations n, ni du nombre variables explicative k. Donc, il faut condidérer un autre coefficient afin de tenir compte n et k. Ce coefficient est appelé

le coefficient de détermination corrigé noté par $\overline{R}^2$ et qui est défini par :

$$\overline{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2) \tag{13}$$

Alors, la qualitée de l'ajustement et $\overline{R}^2$ est meilleur lorsque est très proche de 1.

Remarque

1. Si $k = 0 \Longrightarrow \overline{R}^2 = R^2$
2. Si $k \geq 1 \Longrightarrow \overline{R}^2 < R^2$
3. Quand $n \to +\infty \Longrightarrow \overline{R}^2 \simeq R^2$

0.6 Tests Statistiques

Afin de faire des tests Statistiques, on doit introduire une Hypothèse supplémentaire qui est celle de la normalité des erreurs : $H_9 : \varepsilon_t \rightsquigarrow N(0, \sigma_\varepsilon^2)$

0.6.1 Test de significativité globale du modèle (Fisher)

Le test de Fisher nous permet de tester s'il existe au moins une variable X_i qui explique la variable Y, ce test est défini par les deux hypothèses suivantes :
$H_0 : B_1 = B_2 = ... = B_k = 0$, aucune variable n'explique la variable Y
$H_1 : \exists i = \overline{1,k}/B_i \neq 0$, Le modèle est globalement significatif
Pour cela, nous devons calculer la statistique de Fisher empirique F_c afin de la comparer à celle lu sur la table $F_{\frac{\alpha}{2}}(k, n-k-1)$, où $(k, n-k-1)$ sont les degrés de liberté la statistique de Fisher au seuil $\frac{\alpha}{2}$.

$$F_c = \frac{SCE/k}{SCR/(n-k-1)} = \frac{R^2/k}{(1-R^2)/(n-k-1)} \tag{14}$$

Aolrs :

- si $F_c > F_{\frac{\alpha}{2}}(k, n-k-1)$, alors le modèle est globalement significatif.
- si $F_c < F_{\frac{\alpha}{2}}(k, n-k-1)$, alors le modèle n'est pas globalement significatif, et que la variable X n'est pas explicative (nous acceptons l'hypothèse H_0).

0.6.2 Test de Student

D'après l'hypothèse H_9, les erreurs suivent une loi normale, ce qui nous permet de vérifier que :

$$\begin{cases} \frac{\widehat{B_i}-B_i}{\sigma^2_{\widehat{B_i}}} \rightsquigarrow N(0,1), \\ \\ \frac{\sum\limits_{t=1}^{n} e_t^2}{\sigma_\varepsilon^2} = \frac{(n-k-1)\widehat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} \rightsquigarrow \chi^2_{(n-k-1)}, \end{cases}$$

où $\chi^2_{(n-k-1)}$ est le khi-deux à $(n-k-1)$ degrés de liberté.

Et $\frac{\widehat{B_i}-B_i}{\widehat{\sigma}^2_{\widehat{B_i}}} \rightsquigarrow t_{\frac{\alpha}{2}}(n-k-1)$.

où $t_{\frac{\alpha}{2}}(n-k-1)$ est la Student à $(n-k-1)$ degrés de liberté.

Par conséquent, nous pouvons mettre en place des tests statistiques (tests bilatéraux) pour vérifier la significativité des B_i.

Test de Student

$H_0 : B_i = 0$ contre $H_1 : B_i \neq 0$. Dans un premier, nous douvons calculer la statistique de Student empirique t_c afin de la comparer à celle lu sur la table $t_{\alpha/2}(n-k-1)$, où $(n-k-1))$ est le degré de liberté la statistique de Student au seuil $\alpha/2$, avec

$$t_c = \left| \frac{\widehat{B_i} - B_i}{\widehat{\sigma}^2_{\widehat{B_i}}} \right| \tag{15}$$

Aolrs :

- si $t_c > t_{\alpha/2}(n-k-1)$, alors B_i est significatif.
- si $F_c < t_{\alpha/2}(n-k-1)$, alors B_i n'est pas significatif, et que la variable X_i n'est pas explicative (nous acceptons l'hypothèse H_0).

Intervalle de confiance

Nous pouvons confirmer les résultats obtenus par le test bilatéral de Student à partir de l'intervalle de confiance de chaque paramètre. Dans la plus part des σ_ε est inconnu, alors

L'intervalle de confiance pour B_i (I_{CB_i})

$$B_i = \widehat{B_i} \pm t_{\alpha/2}(n-k-1)\widehat{\sigma}_{\widehat{B_i}} \tag{16}$$

où $t_{\alpha/2}(n-k-1)$ est la valeur de la statistique de Student au seuil $\alpha/2$ à $(n-k-1)$ est le degré de liberté lu sur la table.

Remarque

$H_0 : B_i = 0$ contre $H_1 : B_i \neq 0$.

- si $0 \notin I_{CB_i}$, alors B_i est significatif.
- si $0 \in I_{CB_i}$, alors B_i n'est pas significatif, et que la variable X_i n'est pas explicative (nous acceptons l'hypothèse H_0).

0.7 Prévision

Une fois le modèle est validé (toutes les variables sont significatives), nous pouvons faire des prévisions pour Y_{n+h} lorsque nous connaisons les valeurs $X_{i\ n+h}$.

$$Y_{n+h} = \widehat{B}_0 + \widehat{B}_1 X_{1\ n+h} + \widehat{B}_2 X_{2\ n+h} + ... \widehat{B}_k X_{k\ n+h} + \varepsilon_{n+h} \tag{17}$$

Sachant que : $E[\varepsilon_{n+h}] = 0$, et $Cov(\varepsilon_{n+h}, \varepsilon_t) = 0$ et $Var(\varepsilon^2_{n+h}) = \sigma^2_\varepsilon$
Dans un premier temps, nous douvons calculer $\widehat{Y}_{n+h}$ (la valeur ponctuelle ajustée de la prévision) qui est donné par :

$$\widehat{Y}_{n+h} = \widehat{B}_0 + \widehat{B}_1 X_{1\ n+h} + \widehat{B}_2 X_{2\ n+h} + ... \widehat{B}_k X_{k\ n+h} \tag{18}$$

Et

$$Y_{n+h} = \widehat{Y}_{n+h} + e_{t+h} \tag{19}$$

L'erreur de prévision est donnée par :

$$e_{t+h} = Y_{n+h} - \widehat{Y}_{n+h}$$

Remarque

$E[e_{n+h}] = 0$ et la valeur ajustée de la prévision $\widehat{Y}_{n+h}$ est sans biais. On peut facilement montrer que $E\left[\widehat{Y}_{n+h}\right] = Y_{n+h}$.

La variance de l'erreur de prévision

On sait que :
$e_{n+h} = \widehat{Y}_{n+h} - Y_{n+h} = X'_{n+h}\left(B - \widehat{B}\right) + \varepsilon_{n+h}$
Où :

$$X'_{n+h} = (1, X_{1\ n+h}, , X_{1\ n+h}, .., X_{k\ n+h})$$

Alors
$Var(e_{n+h}) = Var\left[X'_{n+h}\left(B - \widehat{B}\right) + \varepsilon_{n+h}\right] = Var\left[X'_{n+h}\left(B - \widehat{B}\right)\right] + Var\left[\varepsilon_{n+h}\right]$

$Var(e_{n+h}) = X'_{n+h} Var\left[\left(B - \widehat{B}\right)\right] X_{n+h} + Var\left[\varepsilon_{n+h}\right] = X'_{n+h} Var\left[\widehat{B}\right] X_{n+h} + Var\left[\varepsilon_{n+h}\right]$

D'après les résultats précédents, on déduit que :

$$Var(e_{n+h}) = \sigma^2_\varepsilon \left[X'_{n+h}(X'X)^{-1} X_{n+h} + 1\right] \tag{20}$$

Et comme la variance de l'erreur σ^2_ε est inconnue, alors la variance de l'erreur de prévision est donnée par :

$$Var(e_{n+h}) = \widehat{\sigma}^2_\varepsilon \left[X'_{n+h}(X'X)^{-1} X_{n+h} + 1\right] \tag{21}$$

Remarque

Nous constatons, comme pour le modèle de régression simple que la valeur de la variance de prévision est faible lorsque les valeurs prévues des variables exogènes se rapprochent de leurs moyennes.

L'intervalle de prévision

D'après l'hypothèse H_9, nous déduisons que :

$$\varepsilon_{n+h} \rightsquigarrow N\left(0, \sigma^2_{e_{n+h}}\right).$$

L'intervalle de prédiction est défini par :

$$Y_{n+h} = \widehat{Y}_{n+h} \pm t_{\frac{\alpha}{2}}(n-k-1)\ \widehat{\sigma}_\varepsilon \sqrt{\left[X'_{n+h}(X'X)^{-1}X_{n+h}+1\right]}. \tag{22}$$

Où $t_{\frac{\alpha}{2}}(n-k-1)$ est la valeur de la loi de Student à $(n-k-1)$ degrés de liberté au seuil de signification $\frac{\alpha}{2}$.

0.8 *Exemple*

exemple 1. Pendant dix ans, de 1995 à 2004, une ferme a expérimenté le rendement du maïs Y associé à l'emploi de quantités croissantes d'un fertilisant X_1 et d'un insecticide X_2. Les données sont :

X_{1t}	6	10	12	14	16	18	22	24	26	32
X_{2t}	4	4	5	7	9	12	14	20	21	24
Y_t	40	44	46	48	52	58	60	68	74	80

Soit le modèle : $Y_t = B_0 + B_1X_{1t} + B_2X_{2t} + \varepsilon_t,\ t = \overline{1,n}$

1. Mettre le modèle sous forme matricielle en spécifiant les dimensions de chacune des matrices.

2. Estimer par la méthode des moindres carrés ordinaires les paramètres du modèle.

3. Calculer la variance résiduelle ainsi que les écarts-types de chacun des paramètres.

4. Calculer le coefficient de détermination et le coefficient de détermination corrigé. Conclusion ?

5. Le modèle est-il globalement significatif au seuil 5% ?

6. Les variables explicatives sont-elles significatives au seuil 5% ?

7. Donner la valeur de Y à la date 11 sachant que :$X_{1\ 11} = 36$, $X_{2\ 11} = 27$

8. Donner le tableau de l'analyse de la variance.

Corrigé 1.

1. La forme matricielle est par l'equation suivante :

$$\underset{(n=10,1)}{Y} = \underset{(n=10,k+1=3)}{X}\ \underset{(k+1=3,1)}{B} + \underset{(n=10,1)}{\varepsilon} \tag{23}$$

où

- Y est un vecteur aléatoire de dimension $n = 10$,
- X est une matrice de taille $n \times (k+1) = 10(3) = 30$, connue, appelée matrice du plan d'expérience,
- B est le vecteur des paramètres inconnus du modèle,
- ε est le vecteur des erreurs.

Avec

$$\underset{(n=10,1)}{Y} = \begin{pmatrix} Y_1 = 40 \\ Y_2 = 44 \\ Y_3 = 46 \\ Y_4 = 48 \\ Y_5 = 52 \\ Y_6 = 58 \\ Y_7 = 60 \\ Y_8 = 68 \\ Y_9 = 74 \\ Y_{10} = 80 \end{pmatrix}; \quad \underset{(n=10,k+1=3)}{X} = \begin{pmatrix} 1 & X_{1\ 1} = 6 & X_{2\ 1} = 4 \\ 1 & X_{1\ 2} = 10 & X_{2\ 2} = 4 \\ 1 & X_{1\ 3} = 12 & X_{2\ 3} = 5 \\ 1 & X_{1\ 4} = 14 & X_{2\ 4} = 7 \\ 1 & X_{1\ 5} = 16 & X_{2\ 5} = 9 \\ 1 & X_{1\ 6} = 18 & X_{2\ 6} = 12 \\ 1 & X_{1\ 7} = 22 & X_{2\ 7} = 14 \\ 1 & X_{1\ 8} = 24 & X_{2\ 8} = 20 \\ 1 & X_{1\ 9} = 26 & X_{2\ 9} = 21 \\ 1 & X_{1\ 10} = 32 & X_{2\ 10} = 24 \end{pmatrix};$$

$$\underset{(k+1=3,1)}{B} = \begin{pmatrix} B_0 \\ B_1 \\ B_2 \end{pmatrix}; \quad \underset{(n=10,1)}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_{n=10} \end{pmatrix}.$$

2. Nous allons estimer le vecteur des paramètres B par la méthode Moindres carrés ordinaires (MCO), en supposant que : $Y = X\ B\ +\ \varepsilon$.
   L'estimer du vecteur des paramètres B est donné par :

$$\widehat{B} = (X'X)^{-1}X'Y. \tag{24}$$

Afin de caculer les estimateurs des coefficients de régression, nous devons calculer dans un premier temps $(X'X)^{-1}$.

$$(X'X) = \begin{pmatrix} n & \sum X_{1t} & \sum X_{2t} \\ 10 & 180 & 120 \\ \sum X_{1t} & \sum X_{1t}^2 & \sum X_{1t}X_{2t} \\ 180 & 3816 & 2684 \\ \sum X_{2t} & \sum X_{2t}X_{1t} & \sum X_{2\ t}^2 \\ 120 & 2684 & 1944 \end{pmatrix}$$

Où, les résultats utilisés sont donnés dans le tableau suivant :

n	Y_t	$X_{1\ t}$	$X_{2\ t}$	$X_{1\ t}X_{2\ t}$	$X_{1\ t}^2$	$X_{2\ t}^2$	$X_{1\ t}Y_t$	$X_{2\ t}Y_t$	Y_t^2
1	40	6	4	24	36	16	240	160	1600
2	44	10	4	40	100	16	440	176	1936
3	46	12	5	60	144	25	552	230	2116
4	48	14	7	98	196	49	672	336	2304
5	52	16	9	144	256	81	832	468	2704
6	58	18	12	216	324	144	1044	696	3364
7	60	22	14	308	484	196	1320	840	3600
8	68	24	20	480	576	400	1632	1630	4624
9	74	26	21	546	676	441	1924	1554	5476
10	80	32	24	786	1024	676	2560	1920	6400
$\sum$	570	180	120	2684	3816	1944	11216	7740	34124

$$\det(X'X) = \begin{vmatrix} 10 & 180 & 120 \\ 180 & 3816 & 2684 \\ 120 & 2684 & 1944 \end{vmatrix}$$

$$\det(X^{'}X) = 10 \begin{vmatrix} 3816 & 2684 \\ 2684 & 1944 \end{vmatrix} - 180 \begin{vmatrix} 180 & 120 \\ 2684 & 1944 \end{vmatrix} + 120 \begin{vmatrix} 180 & 120 \\ 3816 & 2684 \end{vmatrix} = 157280$$

La comatrice de $(X^{'}X)$

$$C_{(X'X)} = \begin{pmatrix} + \begin{vmatrix} 3816 & 2684 \\ 2684 & 1944 \end{vmatrix} & - \begin{vmatrix} 180 & 120 \\ 2684 & 1944 \end{vmatrix} & + \begin{vmatrix} 180 & 120 \\ 3816 & 2684 \end{vmatrix} \\ - \begin{vmatrix} 180 & 120 \\ 2684 & 1944 \end{vmatrix} & + \begin{vmatrix} 10 & 120 \\ 120 & 1944 \end{vmatrix} & - \begin{vmatrix} 10 & 180 \\ 120 & 2624 \end{vmatrix} \\ + \begin{vmatrix} 180 & 120 \\ 3816 & 2684 \end{vmatrix} & - \begin{vmatrix} 10 & 120 \\ 180 & 2684 \end{vmatrix} & + \begin{vmatrix} 10 & 180 \\ 180 & 3816 \end{vmatrix} \end{pmatrix}$$

$$C_{(X'X)} = \begin{pmatrix} 214448 & -27840 & 25200 \\ -27840 & 5040 & -5240 \\ 25200 & -5240 & 5760 \end{pmatrix}$$

$$(X^{'}X)^{-1} = \frac{C^{'}_{(X'X)}}{\det(X'X)} = \frac{1}{157280} \begin{pmatrix} 214448 & -27840 & 25200 \\ -27840 & 5040 & -5240 \\ 25200 & -5240 & 5760 \end{pmatrix}$$

Et par conséquent

$$\widehat{B} = \frac{C'_{(X'X)}}{\det(X'X)}(X'Y) = \frac{1}{157280}\begin{pmatrix} 214448 & -27840 & 25200 \\ -27840 & 5040 & -5240 \\ 25200 & -5240 & 5760 \end{pmatrix}\begin{pmatrix} 570 \\ 11216 \\ 7740 \end{pmatrix}$$

Avec

$$(X'Y) = \begin{pmatrix} \sum Y_t = 570 \\ \sum X_{1\ t} Y_t = 11216 \\ \sum X_{2\ t} Y_t = 7740 \end{pmatrix}$$

$$\widehat{B} = \frac{1}{157280}\begin{pmatrix} 5029920 \\ 102240 \\ 174560 \end{pmatrix} = \begin{pmatrix} \frac{5029920}{157280} \\ \frac{102240}{157280} \\ \frac{174560}{157280} \end{pmatrix}$$

$$\widehat{B} = \begin{pmatrix} 31.98 \\ 0.65 \\ 1.11 \end{pmatrix} \tag{25}$$

Donc, le modèle s'écrit comme suit :

$$\widehat{Y_t} = 31.98 + 0.65X_{1\ t} + 1.11X_{2\ t}, \quad t = \overline{1,10} \tag{26}$$

3. Pour le calcul de la variance des paramètres du modèle, on doit calculer la matrice variance covariance $\widehat{\Omega}_{\widehat{B}}$

$$\widehat{\Omega}_{\widehat{B}} = \widehat{\sigma}_{\varepsilon}^{2}(X^{'}X)^{-1} \tag{27}$$

Où

$$\widehat{\sigma}_{\varepsilon}^{2} = \frac{\sum_{t=1}^{n} e_i^2}{n-(k+1)}$$

Avec

$$\sum_{t=1}^{n} e_i^2 = \sum_{t=1}^{n} Y_t^2 - \widehat{B}^{'}\,(X^{'}Y) = 34124 - (31.98, 0.65, 1.11)\begin{pmatrix} 570 \\ 11216 \\ 7740 \end{pmatrix} = 34124 - 34110.4 = 13.6$$

Et on déduit que :

$$\widehat{\sigma}_{\varepsilon}^{2} = \frac{13.6}{10-3} = 1.94 \Rightarrow \widehat{\sigma}_{\varepsilon} = 1.39$$

La variance des paramètres du modèle se trouve sur la diagonale de la matrice variance covariance $\widehat{\Omega}_{\widehat{B}}$, c'est à dire :

$$\widehat{\sigma}_{\widehat{B}_0}^{2} = 1.94 \times 1.363 = 2.64 \Rightarrow \widehat{\sigma}_{\widehat{B}_0} = 1.63$$

$$\widehat{\sigma}_{\widehat{B}_1}^{2} = 1.94 \times 0.032 = 0.062 \Rightarrow \widehat{\sigma}_{\widehat{B}_1} = 0.25$$

$$\widehat{\sigma}_{\widehat{B}_2}^{2} = 1.94 \times 0.036 = 0.07 \Rightarrow \widehat{\sigma}_{\widehat{B}_2} = 0.26$$

4. ♣ Le calcul du coefficient de détermination R^2

$$R^2 = 1 - \frac{SCR}{SCT}. \tag{28}$$

♣ Le calcul du SCT

$$SCT = \sum_{t=1}^{n} Y_t^2 - n\overline{Y}^2 = 34124 - 10(57)^2 = 1634$$

Alors

$$R^2 = 1 - \frac{SCR}{SCT} = 1 - \frac{13.6}{1634} = 0.99$$

♣ Le calcul du coefficient de détermination corrigé $\overline{R}^2$

$$\overline{R}^2 = 1 - \frac{n-1}{n-(k+1)}(1-R^2) = 1 - \frac{10-1}{10-3}(1-0.99) = 0.99 \tag{29}$$

D'après le résultat de R^2 et $\overline{R}^2$, on conclu que la qualité de l'ajustement est très importante.

5. Le test globale
   $H_0 : B_1 = B_2 = ... = B_k = 0$, aucune variable n'explique la variable Y
   $H_1 : \exists i = \overline{1,k} / B_i \neq 0$

$$F_c = \frac{SCE/k}{SCR/(n-(k+1))} = \frac{R^2/k}{(1-R^2)/(n-(k+1))} \tag{30}$$

$$F_c = \frac{0.99/2}{(1-0.99)/(10-3)} = 346.5$$

On accepte H_1, car $F_c = 346.5 > F_{\frac{\alpha}{2}}(k, n-(k+1)) = F_{0.05}(2,7) = 4.74$, ce qui signifie que le modèle est globalement significatif au seuil de 10%

6. On utilise le test de Student pour la significativité de B_i, qui est définie par :

   $H_0 : \ B_i = 0$ contre $H_1 : \ B_i \neq 0$
   Pour cela, on doit calculer $t_{c\ B_i}$

$$t_{c_{B_i}} = \left| \frac{\widehat{B_i} - B_i}{\widehat{\sigma}_{\widehat{B_i}}} \right| \tag{31}$$

On obtient alors les résultats suivants :

$$t_{c_{B_0}} = \frac{30.98 - 0}{1.63} = 19.62$$

$$t_{c_{B_1}} = \frac{0.65 - 0}{0.25} = 2.62$$

$$t_{c_{B_2}} = \frac{1.11 - 0}{0.26} = 4.27$$

On constate d'après la valeur de la Statistique de Student calculer pour chaque B_i qu'elle est superieure à la valeur de la statistique de Student ($t_T(7, 0.025) = 2.365$) lu sur la table avec 7 est son degré de liberté et 0.025 est le seuil de confiance. On conclut donc que tous les paramètres sont significatifs et que les deux variables sont significatives. Ce qui nous permis de dire que ce modèle est validé.

7. Le calcul de la prévision à la date 11 sachant que $X_{1\ 11} = 36$ et $X_{2\ 11} = 27$

   Du moment le modèle est validé, on peut donc calculer la prévision à la date 11 .

   ♣ La valeur ponctuelle ajustée de la prévision est donné par :

$$\widehat{Y}_{n+h} = \widehat{B}_0 + \widehat{B}_1 X_{1\ n+h} + \widehat{B}_2 X_{2\ n+h} + ...\widehat{B}_k X_{k\ n+h} \tag{32}$$

   Donc

$$\widehat{Y}_{11} = \widehat{B}_0 + \widehat{B}_1 X_{1\ 11} + \widehat{B}_2 X_{2\ 11} \tag{33}$$

   Donc

$$\widehat{Y}_{11} = 30.98 + 0.65 X_{1\ 11} + 1.11 X_{2\ 11} = 30.98 + 0.65(36) + 1.11(27) = 85.35$$

   ♣ L'intervalle de prédiction est défini par :

$$Y_{11} = \widehat{Y}_{11} \pm t_{0.025}(7)\ \sigma_{e_{n+h}} \tag{34}$$

Où $\sigma^2_{e_{n+h}}$ est la variance de l'erreur de prévision qui est donnée par :

$$\sigma^2_{e_{n+h}} = Var(e_{n+h}) = \widehat{\sigma}^2_{\varepsilon}\left[X'_{n+h}(X'X)^{-1}X_{n+h} + 1\right] \tag{35}$$

$$\sigma^2_{e_{n+h}} = \widehat{\sigma}^2_{\varepsilon}\left[X'_{11}(X'X)^{-1}X_{11} + 1\right]$$

Avec $X'_{11} = (1, 36, 27)$. Alors

$$\sigma^2_{e_{n+h}} = 1.94\left[\frac{1}{157280}(1, 36, 27)\begin{pmatrix} 214448 & -27840 & 25200 \\ -27840 & 5040 & -5240 \\ 25200 & -5240 & 5760 \end{pmatrix}\begin{pmatrix} 1 \\ 36 \\ 27 \end{pmatrix} + 1\right]$$

$$\sigma^2_{e_{n+h}} = 1.94\left[\frac{1}{157280}(1, 36, 27)\begin{pmatrix} -107392 \\ 12120 \\ -7920 \end{pmatrix} + 1\right] = 1.94\,[0.823 + 1] = 3.537$$

♣ L'intervalle de prédiction est défini par :

$$Y_{11} = 85.35 \pm 2.62(1.881) = [81.02\ ;\ 89.68]. \tag{36}$$

8. Le tableau de l'analyse de la variance

Source de variation	Somme des carrés	Degrés de liberté	Carré moyen
Régression X_1, X_2	$SCE = \underset{=1634-13.6=1620.4}{SCT - SCR}$	2	$\frac{SCE}{2} = \frac{1620.4}{2} = 810.2$
Résiduelle	$SCR = 13.6$	$n - 3 = 10 - 3 = 7$	$\frac{SCR}{7} = \frac{13.6}{7} = 1.94$
Totale	$SCT = 1634$	$n - 1 = 10 - 1 = 9$	