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

**Niveau** : $L3$ $EMB$

BERKAI née MECHERI Kheira

## Chapitre 1 : Régression Linéaire simple

L'objectif de la régression linéaire simple est d'établir un lien entre une variable dépendante $Y$ et une variable indépendante $X$, ce qui va nous permettre de faire ensuite des prévisions de Y lorsque X est mesurée.

## 0.1 *Présentation du modèle*

Un modèle de régression linéaire simple est défini par :

$$Y_t = B_0 + B_1 X_t + \varepsilon_t, \quad t = 1...n \tag{1}$$

*où*

- $Y_t$ est la variable dépendante ou endogène (une variable à expliquer) au temps $t$.
- $B_0$, $B_1$ sont les coefficients de la régression.
- $X_t$ est la variable indépendante ou exogène (variable explicative).
- $\varepsilon_t$ est une erreur aléatoire.
- $n$ est le nombre d'observations.

Ce modèle doit satisfaire certaines conditions :
$H_1$ : Le modéle est linéaire en $X_t$.
$H_2$ : Les valeurs de $X_t$ sont des grandeurs numériques mesurées sans erreur.
$H_3$ : Les $\varepsilon_t$ sont i.i.d. (indépendants et identiquement distribués).
$H_4$ : La moyenne des erreurs s'annulent, le modèle est bien spécifié, $E(\varepsilon_t) = 0$.
$H_5$ : La variance de l'erreur est constante et ne dépend pas de l'observation : homoscédasticité, $Var(\varepsilon_t) = \sigma_\varepsilon^2$.
$H_6$ : L'erreur est indépendante de la variable exogène, $COV(\varepsilon_t, X_t) = 0$.
$H_7$ : Indépendance des erreurs, les erreurs relatives à 2 observations sont indépendantes (on dit aussi que les erreurs sont non corrélées), $COV(\varepsilon_t, \varepsilon_s) = 0$.
$H_8$ : $\varepsilon_t$ suit une loi normale, $\varepsilon_t \rightsquigarrow N(0, \sigma_\varepsilon^2)$

## 0.2 *Estimation des paramètres par la méthode Moindres carrés ordinaires (MCO)*

Dans la réalité les valeurs de $B_0$ et $B_1$ sont inconnus, on doit les estimer. Leurs estimateurs sont donnés par $\widehat{B_0}$ et $\widehat{B_1}$ qui sont obtenus par la méthode Moindres carrés ordinaires (MCO). Cette méthode consiste à minimiser la somme des carrés des erreurs.

$$\min \sum_{t=1}^{n} \varepsilon_s^2 = \min \sum_{t=1}^{n} \left(Y_t - B_0 - B_1 X_t\right)^2 = \min S(B_0, B_1) \tag{2}$$

Pour minimiser $S(B_0$ et $B_1)$, il suffit de trouver $\widehat{B_0}$ et $\widehat{B_1}$.

$$\begin{cases} \frac{\partial S(B_0,B_1)}{\partial B_0} = 0 \\ \\ \frac{\partial S(B_0,B_1)}{\partial B_1} = 0 \end{cases}$$

Par conséquent, on obtient

$$\begin{cases} -2\sum_{t=1}^{n}\left(Y_t - \widehat{B_0} - \widehat{B_1}X_t\right) = 0 \\ \\ -2\sum_{t=1}^{n} X_t\left(Y_t - \widehat{B_0} - \widehat{B_1}X_t\right) = 0 \end{cases}$$

Et après calculs que,

$$\begin{cases} \widehat{B_1} = \frac{\sum_{t=1}^{n} X_t Y_t - n\ \overline{X}\ \overline{Y}}{\sum_{t=1}^{n} X_t^2 - n\overline{X}^2} = \frac{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)\left(Y_t-\overline{Y}\right)}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2} \\ \\ \widehat{B_0} = \overline{Y} - \widehat{B_1}\overline{X} \end{cases} \tag{3}$$

***Remarque***

1. Si $B_0$ est constant, alors

$$\widehat{B_1} = \frac{\sum_{t=1}^{n} X_t Y_t}{\sum_{t=1}^{n} X_t^2} \tag{4}$$

2. Les estimateurs peuvent s'écrire :

$$\begin{cases} \widehat{B_1} = B_1 + \frac{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)\ \varepsilon_t}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2} \\ \\ \widehat{B_0} = B_0 - \left(\widehat{B_1} - B_1\right)\overline{X} + \overline{\varepsilon} \end{cases} \tag{5}$$

## 0.3 *Propriétés des estimateurs* $\widehat{B_0}$ *et* $\widehat{B_1}$

D'après la remarque précédente, nous allons montrer que ces estimateurs sont sans biais et convergents :

### *Les estimateurs sont sans biais*

Nous savons que les bruits $\varepsilon_t$ sont aléatoires, et puisqu'ils sont centrés, et D'après la formule précédente de $\widehat{B_1}$, on déduit bien que $E[\widehat{B_1}] = B_1$ ($\widehat{B_1}$ est sans biais) .
Pour $\widehat{B_0}$, on a : $E[\widehat{B_0}] = B_0 - E[\left(\widehat{B_1} - B_1\right)]\overline{X}$, et comme $\widehat{B_1}$ est sans biais, on déduit alors :$E[\widehat{B_0}] = B_0$, donc $\widehat{B_0}$ est sans biais.

### *Les estimateurs sont convergents*

$$Var(\widehat{B_1}) = E\left(\widehat{B_1} - B_1\right)^2 = E\left[\frac{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)\varepsilon_t}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2}\right]^2$$

$$Var(\widehat{B_1}) = \frac{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2}{\left(\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2\right)^2}E\left[\varepsilon_t\right]^2 + \sum_{t\leq s}\left[\frac{\left(X_t-\overline{X}\right)}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2}\right]\left[\frac{\left(X_s-\overline{X}\right)}{\sum_{s=1}^{n}\left(X_s-\overline{X}\right)^2}\right]E\left[\varepsilon_t\varepsilon_s\right]$$

$Var(\widehat{B_1}) = \frac{E[\varepsilon_t]^2}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2}$ car les erreurs sont non corrélées.

$$Var(\widehat{B_1}) = \frac{\sigma_\varepsilon^2}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2} \tag{6}$$

$Var(\widehat{B_0}) = Var(B_0) + Var(\widehat{B_1})\overline{X}^2 + Var(B_1)\overline{X}^2 + Var(\overline{\varepsilon}) = Var(\widehat{B_1})\overline{X}^2 + Var(\overline{\varepsilon})$,
car $Var(B_0) = Var(B_1) = 0$

$$Var(\widehat{B_0}) = \frac{1}{n}\sigma_\varepsilon^2 + \frac{\overline{X}^2\sigma_\varepsilon^2}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2} = \sigma_\varepsilon^2\left[\frac{1}{n} + \frac{\overline{X}^2}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2}\right] \tag{7}$$

Les deux estimateurs des moindres carrés sont convergents lorsque le nombre d'observations est important ou lorsque les valeurs de variable explicative sont très disperées autour de la moyenne.

***Remarque***

$$Cov(\widehat{B_1}, \widehat{B_0}) = -\overline{X}Var(\widehat{B_1}) = -\frac{\sigma_\varepsilon^2\ \overline{X}}{\sum_{t=1}^{n}\left(X_t-\overline{X}\right)^2}.$$

## 0.4 *Analyse de la variance*

Les estimateurs $\widehat{B_0}$ et $\widehat{B_1}$ des coefficients $B_0$ et $B_1$ prmettent de calculer pour chaque observation la réponse estimée. Celle-ci sera comparée à la valeur observée de $Y_t$ par l'intermédiaire des résidus $e_t$,
où $e_t = Y_t - \widehat{Y_t}$, avec

- $E[e_t] = 0$
  car $\sum_{t=1}^{n} e_t = 0$
- $Var(e_t) = \sum_{t=1}^{n} e_t^2$
  $Var(e_t) = \sum_{t=1}^{n} (Y_t - \widehat{Y_t})^2 = (n-2)\sigma_\varepsilon^2$.
  Ce qui nous permet de déduire que :

$$\widehat{\sigma_\varepsilon^2} = \frac{\sum_{t=1}^{n} e_t^2}{(n-2)} = \frac{SCR}{(n-2)} \tag{8}$$

Alors les estimateurs des variances des deux paramètres $\widehat{B_1}$ et $\widehat{B_0}$ sont donnés par

$$\begin{cases} Var(\widehat{B_1}) = \dfrac{\widehat{\sigma_\varepsilon^2}}{\sum_{t=1}^{n} (X_t - \overline{X})^2}, \\ \\ Var(\widehat{B_0}) = \widehat{\sigma_\varepsilon^2} \left[ \dfrac{1}{n} + \dfrac{\overline{X}^2}{\sum_{t=1}^{n} (X_t - \overline{X})^2} \right], \end{cases} \tag{9}$$

### 0.4.1 *Décomposition de la variance*

L'objectif de la régression est de minimiser $\sum_{t=1}^{n} e_t^2$ et ce qui nous permet de juger la qualité de l'ajustement de ce modèle.

$$\sum_{t=1}^{n} (Y_t - \overline{Y})^2 = \sum_{t=1}^{n} \left(Y_t - \widehat{Y_t} + \widehat{Y_t} - \overline{Y}\right)^2 = \sum_{t=1}^{n} \left(Y_t - \widehat{Y_t}\right)^2 + \sum_{t=1}^{n} \left(\widehat{Y_t} - \overline{Y}\right)^2 + 2 \underbrace{\sum_{t=1}^{n} \left(Y_t - \widehat{Y_t}\right)\left(\widehat{Y_t} - \overline{Y}\right)}_{=0}$$

$$\underbrace{\sum_{t=1}^{n} (Y_t - \overline{Y})^2}_{SCT} = \underbrace{\sum_{t=1}^{n} \left(Y_t - \widehat{Y_t}\right)^2}_{SCR} + \underbrace{\sum_{t=1}^{n} \left(\widehat{Y_t} - \overline{Y}\right)^2}_{SCE} \tag{10}$$

$SCT$ : somme des carrés totaux (variabilité totale)

$SCE$ : somme des carrés expliqués par le modèle (variabilité expliquée)

$SCR$ : somme des carrés résiduels, non expliqués par le modèle (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 (expliquée) | $SCE = \sum_{t=1}^{n} \left(\widehat{Y}_t - \overline{Y}\right)^2$ | 1 | $SCE$ |
| Résiduelle | $SCR = \sum_{t=1}^{n} \left(Y_t - \widehat{Y}_t\right)^2$ | $n-2$ | $\frac{SCR}{n-2}$ |
| Totale | $SCT = \sum_{t=1}^{n} \left(Y_t - \overline{Y}\right)^2$ | $n-1$ | |

***Le coefficient de détermination*** $R^2$

Le coefficient de détermination $R^2$, nous donne une idée sur la qualitée de l'ajustement et il est meilleur lorsque la variabilitée expliquée est très proche de la variabilitée totale ($c-\grave{a}-d$ $R^2$ est très proche de 1) qui est donné par la formule suivante :

$$R^2 = \frac{\sum_{t=1}^{n} \left(\widehat{Y}_t - \overline{Y}\right)^2}{\sum_{t=1}^{n} \left(Y_t - \overline{Y}\right)^2} = 1 - \frac{\sum_{t=1}^{n} \left(Y_t - \widehat{Y}_t\right)^2}{\sum_{t=1}^{n} \left(Y_t - \overline{Y}\right)^2} \tag{11}$$

ou

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT} = 1 - \frac{\sum_{t=1}^{n} e_t^2}{SCT} \tag{12}$$

***Remarque***

Le coefficient decrrélation linéaire simple est défini par

$$r = \rho = R = \frac{\sum_{t=1}^{n} \left(Y_t - \overline{Y}\right)\left(X_t - \overline{X}\right)}{\sqrt{\sum_{t=1}^{n} \left(X_t - \overline{X}\right)^2}\sqrt{\sum_{t=1}^{n} \left(Y_t - \overline{Y}\right)^2}} \tag{13}$$

## 0.5 *Tests de significativités du modèle*

### 0.5.1 *Test de significativité globale du modèle (Fisher)*

Le test de Fisher nous permet de tester la significativité de la régression, ce test est défini par les deux hypothèses suivantes :

$H_0 : B_1 = 0$, Le modèle n'explique pas la variable $Y$
$H_1 : B_1 \neq 0$, Le modèle est pertinent (globalement significatif)
Pour cela, nous devons calculer la statistique de Fisher empirique $F_c$ afin de la comparer à celle lu sur la table $F_{\frac{\alpha}{2}}(1, n-2)$, où $(1, n-2)$ sont les degrés de liberté la statistique de Fisher au seuil $\frac{\alpha}{2}$.

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

Aolrs :

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

### 0.5.2 *Test de Student*

D'après l'hypothèse $H_8$, les erreurs suivent une loi normale, ce qui nous permet de vérifier que :

$$\begin{cases} \frac{\widehat{B_0}-B_0}{\sigma^2_{\widehat{B_0}}} \rightsquigarrow N(0,1), \\ \\ \frac{\widehat{B_1}-B_1}{\sigma^2_{\widehat{B_1}}} \rightsquigarrow N(0,1), \\ \\ \frac{\sum_{t=1}^{n} e_t^2}{\sigma_\varepsilon^2} = \frac{(n-2)\widehat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} \rightsquigarrow \chi^2_{(n-2)}, \end{cases}$$

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

$$\text{ET} \begin{cases} \frac{\widehat{B_0}-B_0}{\widehat{\sigma}^2_{\widehat{B_0}}} \rightsquigarrow t_{\frac{\alpha}{2}}(n-2), \\ \\ \frac{\widehat{B_1}-B_1}{\widehat{\sigma}^2_{\widehat{B_1}}} \rightsquigarrow t_{\frac{\alpha}{2}}(n-2), \end{cases}$$

où $t_{\frac{\alpha}{2}}(n-2)$ est la Student à $(n-2)$ degrés de liberté.
Par conséquent, nous pouvons mettre en place des tests statistiques pour voir si les paramètres sont significatifs ?

#### *Test bilatéral*

$H_0 : B_1 = 0$ contre $H_1 : B_1 \neq 0$ . Dans un premier, nous douvons calculer la statistique de Student empirique $t_c$ afin de la comparer à celle lu sur la table $t_{\alpha/2}(n-2)$, où $(n-2)$ est le degré de liberté la statistique de Student au seuil $\alpha/2$, avec

$$t_c = \left| \frac{\widehat{B_1} - B_1}{\widehat{\sigma}^2_{\widehat{B_1}}} \right| \tag{15}$$

Aolrs :

- si $t_c > t_{\alpha/2}(n-2)$, alors $B_1$ est significatif.

- si $F_c < t_{\alpha/2}(n-2)$, alors $B_1$ n'est pas significatif, et que la variable $X$ n'est pas explicative (nous acceptons l'hypothèse $H_0$ ).

De la même maniere, nous pouvons vérifier si $B_0$ est significatif ou non ?

***Intervalle de confiance***

Nous pouvons aussi donner un intervalle de confiance pour chaque paramètre. Dans la plus part des cas $\sigma_\varepsilon$ est inconnu, alors

***L'intervalle de confiance pour*** $B_0$ $(I_{cB_0})$

$$B_0 = \widehat{B_0} \pm t_{\alpha/2}(n-2)\widehat{\sigma}_{\widehat{B_0}} \tag{16}$$

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

***Remarque***

$H_0 : B_0 = 0$ contre $H_1 : B_0 \neq 0$ .

- si $0 \notin I_{cB_0}$, alors $B_0$ est significatif.

- si $0 \in I_{cB_0}$, alors $B_0$ n'est pas significatif, et que ce modèle ne contient de constante.

La même chose pour l'intervalle de confiance pour $B_1$, si $0 \in I_{cB_1}$, alors $B_1$ n'est pas significatif, et que la variable $X$ n'est pas explicative (nous acceptons l'hypothèse $H_0$ ). Sinon, nous acceptons l'hypothèse $H_1$ et que la variable $X$ est une variable explicative pour la variable endogène $Y$.

## 0.6 *Prévision*

L'un des buts de la régression est de proposer des prévisions pour la variable endogène $Y$. Une fois le modèle est validé, nous pouvons faire des prévisions pour une valeur non observée $X_{n+h}$. Dans un premier temps, nous douvons calculer $Y_{n+h}$ qui est donné par

$$Y_{n+h} = B_0 + B_1 X_{n+h} + \varepsilon_{n+h} \tag{17}$$

avec $E[\varepsilon_{n+h}] = 0$ et $Cov(\varepsilon_{n+h}, \varepsilon_t) = 0$ et $Var(\varepsilon^2_{n+h}) = \sigma^2_\varepsilon$.

• La variance de l'erreur de prévision est donnée par

$$Var(\widehat{\varepsilon}_{n+h}) = Var(B_0 - \widehat{B_0}) + Var(B_1 - \widehat{B_1})X_{n+h}^2 + 2X_{n+h}Cov(\widehat{B_0}, \widehat{B_1}) + Var(\varepsilon_{n+h})$$

$$Var(\widehat{\varepsilon}_{n+h}) = Var(\widehat{B_0}) + Var(\widehat{B_1})X_{n+h}^2 + Var(\varepsilon_{n+h})$$

$$Var(\widehat{\varepsilon}_{n+h}) = \frac{\sigma_\varepsilon^2}{n} + (X_{n+h} - \overline{X})^2 \frac{\sigma_\varepsilon^2}{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2} + \sigma_\varepsilon^2$$

$$Var(\widehat{\varepsilon}_{n+h}) = \sigma_\varepsilon^2 \left[\frac{1}{n} + \frac{(X_{n+h} - \overline{X})^2}{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2} + 1\right] \tag{18}$$

• La variance de l'erreur de prévision est fonction de l'ecart quadratique entre la variable exogène prévue et la moyenne de la même variable. Donc plus la valeur prévue s'éloigne de cette moyenne, plus le risque d'erreur est important.

***Remarque***

D'après l'hypothèse $H_8$, nous déduisons que :

$$\widehat{\varepsilon}_{n+h} \rightsquigarrow N\left(0, \sigma_\varepsilon^2 \left[\frac{1}{n} + \frac{(X_{n+h} - \overline{X})^2}{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2} + 1\right]\right).$$

• La valeur de prévision correspondante au modèle estimé est défini par

$$\widehat{Y}_{n+h} = \widehat{B_0} + \widehat{B_1}X_{n+h} \tag{19}$$

Et sa variance est donnée par

$$Var(\widehat{Y}_{n+h}) = \widehat{\sigma}_\varepsilon^2 \left[\frac{1}{n} + \frac{(X_{n+h} - \overline{X})^2}{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2}\right] \tag{20}$$

• L'intervalle de prédiction est défini par

$$Y_{n+h} = \widehat{Y}_{n+h} \pm t_{\frac{\alpha}{2}}(n-2)\ \widehat{\sigma}_\varepsilon \sqrt{\left[\frac{1}{n} + \frac{(X_{n+h} - \overline{X})^2}{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2} + 1\right]}. \tag{21}$$

## 0.7 *Exemple*

Soit le modèle : $Y_i = B_0 + B_1 X_i + \varepsilon_i, \quad i = \overline{1,8}$
Avec les données suivantes :

| $X_i$ | 11.2 | 13.2 | 16 | 16 | 16.8 | 17.2 | 18.6 | 19 |
|---|---|---|---|---|---|---|---|---|
| $Y_i$ | 28 | 33 | 30.4 | 34 | 32.2 | 38 | 40 | 36.4 |

1. Représenter le nuage de points.
2. Déterminer les estimateurs $B_0$, $B_1$ et étudier la significativité de chaque paramètre.
3. Calculer le $SCE$, $SCT$, $SCR$ et le coefficient de détermination.Conclusion ?
4. Donner la prévision à la date 9 sachant que $X_9 = 19.5$
5. Donner le tableau de l'analyse de la variance.

**Corrigé 1.**

1. Représentation du nuage de points

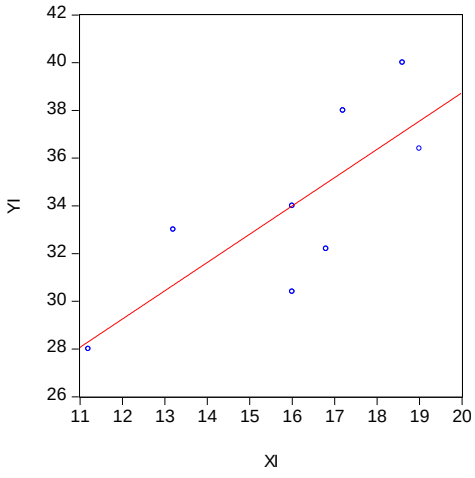

Figure 1 - Nuage de points Xt et Yt

D'après la figure, nous constatons que le nuage de points ne donne pas une idée claire sur le type de relation qu'il existe entre les deux variables, et que la droite de régression ne passe par l'origine.
$\widehat{B_0}$ et $\widehat{B_1}$ sont les quantités qui minimisent les distances verticales entre les observations $Y_t$ et la droite de régression théorique $Y_t = B_0 + B_1 X_t + \varepsilon_t$, et la droite de régression estimée est donnée par :
$\widehat{Y_t} = \widehat{B_0} + \widehat{B_1} X_t$.

Afin de caculer les estimateurs des coefficients de régression, nous réalisons le tableau suivant :

| $n$ | $X_i$ | $Y_i$ | $X_iY_i$ | $X_i^2$ | $Y_i^2$ | $\widehat{Y_i}$ | $e_i = Y_i - \widehat{Y_i}$ | $e_i^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 11.2 | 28 | 313.6000 | 125.44 | 784.0000 | 28.309359 | -0.309359 | 0.095703 |
| 2 | 13.2 | 33 | 435.6000 | 174.24 | 1089.0000 | 30.680460 | 2.319540 | 5.380265 |
| 3 | 16 | 30.4 | 486.4000 | 256.0000 | 924.1600 | 33.99999 | -3.599999 | 12.960002 |
| 4 | 16 | 34 | 544.0000 | 256.0000 | 1156.0000 | 34.00000 | 0.000000 | 0.000000 |
| 5 | 16.8 | 32.2 | 540.9600 | 282.2400 | 1036.8400 | 34.00000 | -2.748440 | 7.553924 |
| 6 | 17.2 | 38 | 653.6000 | 295.8400 | 1444.0000 | 34.948440 | 2.577339 | 6.6426795 |
| 7 | 18.6 | 40 | 744.0000 | 345.9600 | 1600.0000 | 35.422660 | 2.917569 | 8.512211 |
| 8 | 19 | 36.4 | 691.6000 | 361.0000 | 1324.9600 | 37.556650 | -1.156650 | 1.337840 |
| $\sum$ | 128 | 272 | 4409.7600 | 2096.7200 | 9358.9600 | | 0 | 42.48263 |

$\overline{X} = \frac{128}{8} = 16$ et $\overline{Y} = \frac{272}{8} = 34$

$\widehat{B_1} = \frac{\sum_{t=1}^{n} X_t Y_t - n\ \overline{X}\ \overline{Y}}{\sum_{t=1}^{n} X_t^2 - n\overline{X}^2} = \frac{4409.76 - 8(16)(34)}{2096.72 - 8(16)^2} = 1.18555008$

$\widehat{B_0} = \overline{Y} - \widehat{B_1}\overline{X} = 34 - 1.18555008(16) = 15.031199$
On écrit

$$\widehat{Y_i} = 15.031199 + 1.18555008 X_i \tag{22}$$

2. ♣ On a $\sum_{t=1}^{n} e_i^2 = 42.48263$. Et $\widehat{\sigma_\varepsilon^2} = \frac{\sum_{t=1}^{n} e_t^2}{(n-2)} = \frac{42.48263}{8-2} = 7.080438$

♣ $Var(\widehat{B_1}) = \frac{\widehat{\sigma_\varepsilon^2}}{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2} = \frac{7.080438}{2096.7200 - 8(16)^2} = 0.145329$

d'ou

$\widehat{\sigma}_{\widehat{B_1}} = 0.381220$

♣ $Var(\widehat{B_0}) = \widehat{\sigma_\varepsilon^2}\left[\frac{1}{n} + \frac{\overline{X}^2}{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2}\right]$

$$Var(\widehat{B_0}) = 7.08044\left[\frac{1}{8} + \frac{(16)^2}{2096.7200 - 8(16)^2}\right] = 38.089324$$

$$\widehat{\sigma}_{\widehat{B}_0} = 6.171655$$

♣ On utilise le test de Student pour la significativité de $B_0$, qui est définie par :

$H_0 : \ B_0 = 0$ contre $H_1 : \ B_0 \neq 0$
Pour cela, on doit calculer $t_{c\ B_0}$

$$t_{c_{B_0}} = \left| \frac{\widehat{B_0} - B_0}{\widehat{\sigma}_{\widehat{B_0}}} \right| = \frac{15.03120 - 0}{6.171655} = 2.435522 \tag{23}$$

D'après la valeur de la Statistique de Student calculer pour $B_0$ ($t_{c_{B_0}} = 2.435522$) est infieure à $t_T(6, 0.025) = 2.447$ ($t_T(6, 0.025)$ est la valeur de la statistique de Student lu sur la table avec 6 est degré de liberté et 0.025 est le seuil de confiance). on conclut donc que le paramètre n'est pas significatif.

♣ De même pour étudier la significativité de $B_1$. Le test de Student est défini par :

$H_0 : \ B_1 = 0$ contre $H_1 : \ B_1 \neq 0$
Pour cela, on doit calculer $t_{c\ B_1}$

$$t_{c_{B_1}} = \left| \frac{\widehat{B_1} - B_1}{\widehat{\sigma}_{\widehat{B_1}}} \right| = \frac{1.185550 - 0}{0.381221} = 3.109879 \tag{24}$$

D'après la valeur de la Statistique de Student calculer pour $B_0$ ($t_{c_{B_1}} = 3.109879$) est superieure à $t_T(6, 0.025) = 2.447$ ($t_T(6, 0.025)$ est la valeur de la statistique de Student lu sur la table avec 6 est degré de liberté et 0.025 est le seuil de confiance). on conclut donc que le paramètre est significatif.

3. Le calcul $SCT$, $SCE$ et $R^2$

♣ On sait que : $SCR = \sum_{t=1}^{n} e_t^2 = 42.48263$

♣

$$SCT = \sum_{t=1}^{n} \left[ Y_t - \overline{Y} \right]^2 = \sum_{t=1}^{n} Y_t^2 - n\overline{Y}^2 = 9358.96000 - 8(34)^2 = 110.96000 \tag{25}$$

♣ On a $SCT = SCE + SCR$, donc

$$SCE = SCT - SCR = 110.960000 - 42.48263 = 68.47737 \tag{26}$$

***Remarque***

On peut également calculer $SCE$ en utilisant la formule suivante :

$$SCE = \sum_{t=1}^{n} \left[\widehat{Y}_t - \overline{Y}\right]^2 = \widehat{\beta}_1^2 \left[\sum_{t=1}^{n} X_t^2 - n\overline{X}^2\right] = 1.185550^2(48.72) = 68.477363 \quad (27)$$

♣

$$R^2 = \frac{SCE}{SCT} = \frac{68.47737}{110.960000} = 0.617136. \quad (28)$$

D'aprsè la valeur de $R^2$, le modèle est moyennement significatif

4. Le calcul de la prévision à la date 9 sachant que $X_9 = 19.5$

   D'après la Statistique de Student calculer pour chaque paramètre calculer précidament est superieure à $t_T(6, 0.05) = 1.943$ ($t_T(6, 0.05)$ est la valeur de la statistique de Student lu sur la table avec 6 est degré de liberté et 0.05 est le seuil de confiance). on conclut donc que les deux paramètres sont significatifs. Ce qui nous permis de dire que ce modèle est validé et de calculer la prévision à la date 9 .

   ♣ La valeur de prévision correspondante au modèle estimé est défini par

   $$\widehat{Y}_{n+1} = \widehat{B_0} + \widehat{B_1} X_{n+1} \quad (29)$$

   Donc

   $$\widehat{Y}_9 = 15.03120 + 1.185550X_9 = 15.03120 + 1.185550(19.5) = 38.149425$$

   ♣ La variance de l'erreur de prévision $Var(\widehat{\varepsilon}_{n+h})$ est donnée par :

   $$Var(\widehat{\varepsilon}_{n+h}) = \widehat{\sigma}_\varepsilon^2 \left[\frac{1}{n} + \frac{(X_{n+1} - \overline{X})^2}{\sum_{t=1}^{n} \left(X_t - \overline{X}\right)^2} + 1\right] \quad (30)$$

   $$Var(\widehat{\varepsilon}_9) = \widehat{\sigma}_\varepsilon^2 \left[\frac{1}{8} + \frac{(X_9 - \overline{X})^2}{\sum_{t=1}^{n} \left(X_t - \overline{X}\right)^2} + 1\right] = 2.660909 \left[\frac{1}{8} + \frac{(19.5 - 16)^2}{48.72} + 1\right] = 3.662569$$

   Alors $\sigma_{\widehat{\varepsilon}_9} = 1.913784$

♣ L'intervalle de prédiction est défini par :

$$Y_{n+1} = \widehat{Y}_{n+1} \pm t_{\frac{\alpha}{2}}(n-2)\ \sqrt{Var(\widehat{\varepsilon}_{n+1})}. \tag{31}$$

$$Y_9 = \widehat{Y}_9 \pm t_{0.05}(6)\ \sigma_{\widehat{\varepsilon}_9} = 38.149425 \pm 1.943(1.913784) = [34.43094\ ;\ 41.86791]$$

5. Le tableau de l'analyse de la variance

| Source de variation | Somme des carrés | Degrés de liberté | Carré moyen |
|---|---|---|---|
| Régression $X_1$ | $SCE = 68.47737$ | $1$ | $\frac{SCE}{1} = 68.47737$ |
| Résiduelle | $SCR = 42.48263$ | $n-2 = 8-2 = 6$ | $\frac{SCR}{n-2} = \frac{42.48263}{6} = 7.08044$ |
| Totale | $SCT = 110.96$ | $n-1 = 8-1 = 7$ | |