# Modèles linéaires

Pr. Mokhtar KOUKI
mokhtar.kouki@essai.u-carthage.tn

Septembre-Octobre 2020

CHAPITRE 1

# Regression linéaire simple

## 1. Spécification

**1.1. Exemple.** Suppsons qu'on a $n$ observations sur la consommation $(C)$ et le revenu $(R)$ et qu'on cherche à expliquer la consommation. On sait que celle-ci dépend du revenu. On dit alors que la consommation est une variable dépendante ou expliquée (endogène) et le revenu est une variable indépendante ou explicative (exogène). Dans sa version la plus simple, la relation entre consommation et revenu est définie par :

$$C_i = \beta_0 + \beta_1 R_i, \; i = 1, \cdots, n$$

Les constantes $\beta_0$, $\beta_1$ correspondent respectivement à la consommation incompressible et la propension marginale à consommer. Elles sont positives et $0 < \beta_1 < 1$.

Cependant, la consommation d'un ménage est entachée d'aléas qui peuvent être expliqués par des consommations et/ou des revenus imprévisibles. Elle peut aussi dépendre d'autre facteurs tels que la taille du ménage, la catégorie socio-professionnelle du chef du ménage. La relation finale peut être établie sous la forme :



mh

$$C_i = \beta_0 + \beta_1 R_i + \varepsilon_i, \ i = 1, \cdots, n$$

$\varepsilon_i$ est un terme d'erreur qui représente les aléas ainsi que les facteurs non pris en compte dans la relation.

**1.2. Cas général et hypothèses.** Dans cas général, le modèle de regression simple peut être écrit sous la forme :

$$Y_i = \beta_0 + \beta_1 X_i, \ i = 1, \cdots, n$$

$Y$ : variable endogène, $X$ : variable exogène et $\varepsilon$ terme aléatoire qui vérifie les hypothèses suivantes :

: (1) $E(\varepsilon_i) = 0 \ \forall \ i$

(2) $V(\varepsilon_i) = \sigma^2 \ \forall \ i$ (les termes d'erreur sont homoskédatstiques)

(3) $Cov(\varepsilon_i, \varepsilon_j) = 0 \ \forall \ i \neq j$ (les termes d'erreur ne sont pas corrélés)

(4) $Cov(\varepsilon_i, X_j) = 0, \forall \ i, j$ (la variable $X$ est indépendante de $\varepsilon$)

## 2. Estimation des paramètres

**2.1. La méthode des moindres carrés ordinaires.** Les estimateurs par la méthodes des moindres carrés ordinaires (mco), sont définis par :

$$\left(\widehat{\beta}_0, \widehat{\beta}_1\right) = \arg \min_{\beta_0, \beta_1} \sum_{i=1}^{n} \varepsilon_i^2 \tag{2.1}$$

mk

Les conditions nécessaires de première ordre son définies par :

$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i^2}{\partial \beta_0} = 0$$
$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i^2}{\partial \beta_1} = 0$$

Aprés simplification, ces deux conditions peuvent être exprimées sous la forme

$$\sum_{i=1}^{n} \widehat{\varepsilon}_i = 0 \tag{2.2}$$
$$\sum_{i=1}^{n} X_i \widehat{\varepsilon}_i = 0 \tag{2.3}$$

avec $\widehat{\varepsilon}_i = Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i$.

Ces deux dernières équations définissent ce qu'on a convenu d'appeler les "équation normales".

: L'équation (2) :

: $\sum_{i=1}^{n} \widehat{\varepsilon}_i = 0 \Longrightarrow \sum_{i=1}^{n} Y_i = n\widehat{\beta}_0 + \widehat{\beta}_1 \sum_{i=1}^{n} X_i \Leftrightarrow \overline{Y} = \widehat{\beta}_0 + \widehat{\beta}_1 \overline{X}$.On dit que le point moyen de l'échantillon appartient à la droite de regression estimée.

L'équation (3) :
$\sum_{i=1}^{n} X_i \widehat{\varepsilon}_i = 0 \Longrightarrow \sum_{i=1}^{n} X_i \left( Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i \right) = 0$
$\Longrightarrow \sum_{i=1}^{n} X_i \left( Y_i - \overline{Y} - \widehat{\beta}_1 \left( X_i - \overline{X} \right) \right) = 0$
Ce qui permet d'experimer les estimateurs des parmètres :

$$\widehat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \text{ et } \widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1\overline{X} \qquad (2.4)$$

Exemple 1 :

On considère l'équation de la consommation avec :

$n = 20$, $\sum_{i=1}^{n} R_i = 1345.8$, $\sum_{i=1}^{n} C_i = 1221.1$, $\sum_{i=1}^{n} R_i^2 = 92520$, $\sum_{i=1}^{n} C_i^2 = 75969$, $\sum_{i=1}^{n} R_i C_i = 83826$.

Notons que $\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right) = \sum_{i=1}^{n} X_i Y_i - n\overline{XY}$,

$\widehat{\beta}_1 = 0.84$ et $\widehat{\beta}_0 = 4.531$. L'équation estimée est définie par $\widehat{C} = 4.531 + 0.84R$

**2.2. La méthode du maximum de vraisemblance.** L'utilisation de cette méthode requiert la connaissance de la loi de probabilité qui régit les termes d'erreurs. D'où l'hypothèse H5 : $\varepsilon_i \curvearrowright N(0, \sigma^2)$. Le logarithme de la vraisemblance de l'échantillon est donné par :

$$L(\varepsilon_1, \cdots \varepsilon_n, \beta_0, \beta_1, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2}\sum_{i=1}^{n}\frac{\varepsilon_i^2}{\sigma^2} \qquad (2.5)$$

La méthode du maximum de vraisemblance permet d'estimer simultanément les paramètres de la regression et de celui de la variance du terme d'erreur :

$$\left(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\sigma}^2\right) = \arg\min_{\beta_0, \beta_1, \sigma^2} L(\varepsilon_1, \cdots \varepsilon_n, \beta_0, \beta_1, \sigma^2) \qquad (2.6)$$

Les conditions nécessaire de première orde donnent :

$$\sum_{i=1}^{n} \widehat{\varepsilon}_i = 0 \tag{2.7}$$

$$\sum_{i=1}^{n} X_i \widehat{\varepsilon}_i = 0 \tag{2.8}$$

$$-\frac{n}{\widehat{\sigma}^2} + \sum_{i=1}^{n} \frac{\widehat{\varepsilon}_i^2}{\widehat{\sigma}^4} = 0 \tag{2.9}$$

Les deux premières équations sont identiques à celles de la méthode MCO et la dernière équation fournit un estimateur de la variance des erreurs :

$$\widehat{\sigma}^2 = \frac{\sum_{i=1}^{n} \widehat{\varepsilon}_i^2}{n} \tag{2.10}$$

## 3. Propriétés statistiques des estimateurs

**3.1. Estimateurs sans biais.** Par définition $\widehat{\theta}$ est un estimateur sans bias de $\theta$, si seulement si $E(\widehat{\theta}) = \theta$. Ceci nous ramène au calcul des espérance des estimateurs. $\widehat{\beta}_1$ peut être exprimé sous la forme :

$\widehat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{S_{xx}} = \beta_1 + \frac{\sum_{i=1}^{n}(X_i-\overline{X})\varepsilon_i}{S_{xx}}$

$\Longrightarrow E(\widehat{\beta}_1) = \beta_1 + \frac{\sum_{i=1}^{n}(X_i-\overline{X})E(\varepsilon_i)}{S_{xx}} = \beta_1$ car $E(\varepsilon_i) = 0 \forall i$

De la même façon $E(\widehat{\beta}_0) = E(\overline{Y}) - E(\widehat{\beta}_1)\overline{X} = E(\beta_0 + \beta_1\overline{X} + \overline{\varepsilon}) - \beta_1\overline{X} = \beta_0$

Conclusion : $\widehat{\beta}_0$ et $\widehat{\beta}_1$ sont des estimateurs sans biais de $\beta_0$ et $\beta_1$

### 3.2. Calcul de la matrice de variance covariance :

$$
\begin{aligned}
V(\widehat{\beta}_1) &= V\left(\beta_1 + \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\varepsilon_i}{S_{xx}}\right) \\
&= \frac{1}{S_{xx}^2}\left(\sum_i (X_i - \overline{X})^2 V(\varepsilon_i) + \sum_{i<j}(X_i - \overline{X})(X_j - \overline{X})cov(\varepsilon_i, \varepsilon_i)\right) \\
&= \frac{1}{S_{xx}^2}\sum_i (X_i - \overline{X})^2\sigma^2 = \frac{\sigma^2}{S_{xx}}
\end{aligned}
$$

$$
V(\widehat{\beta}_0) = V(\overline{\varepsilon}) + \overline{X}^2 V(\widehat{\beta}_1) - 2\overline{X}cov(\overline{\varepsilon}, \widehat{\beta}_1)
$$

$$
\begin{aligned}
cov(\overline{\varepsilon}, \widehat{\beta}_1) &= cov(\overline{\varepsilon}, \beta_1 + \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\varepsilon_i}{S_{xx}}) \\
&= \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)}{S_{xx}} cov(\overline{\varepsilon}, \varepsilon_i) \\
&= \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)}{S_{xx}} \frac{1}{n} cov(\varepsilon_i, \varepsilon_i) \\
&= \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)}{S_{xx}} \frac{1}{n}\sigma^2 = 0
\end{aligned}
$$

Ce qui permet d'exprimer la variance de $\widehat{\beta}_0$ sous la forme :

$$
V(\widehat{\beta}_0) = \left\{\frac{1}{n} + \frac{\overline{X}^2}{S_{xx}}\right\}\sigma^2
$$

$$
\begin{aligned}
cov(\widehat{\beta}_0, \widehat{\beta}_1) &= cov(\overline{Y}, \widehat{\beta}_1) - \overline{X}var(\widehat{\beta}_1) \\
&= cov(\overline{\varepsilon}, \widehat{\beta}_1) - \overline{X}var(\widehat{\beta}_1) \\
&= -\frac{\overline{X}}{S_{xx}}\sigma^2
\end{aligned}
$$

mli

Ainsi la matrice de variance covariance des paramètres est définie par :

$$
\begin{aligned}
V(\widehat{\beta}_0, \widehat{\beta}_1) &= \begin{pmatrix} v(\widehat{\beta}_0) & cov(\widehat{\beta}_0, \widehat{\beta}_1) \\ cov(\widehat{\beta}_0, \widehat{\beta}_1) & v(\widehat{\beta}_1) \end{pmatrix} \\
&= \sigma^2 \begin{pmatrix} \frac{1}{n} + \frac{\overline{X}^2}{S_{xx}} & -\frac{\overline{X}}{S_{xx}} \\ -\frac{\overline{X}}{S_{xx}} & \frac{1}{S_{xx}} \end{pmatrix}
\end{aligned}
$$

**3.3. Estimateur sans biais de la variance des erreurs.** L'estiamteur sans biais de la variance des erreurs est définie par :

$$
\widehat{\sigma}^2 = \frac{\sum_{i=1}^{n} \widehat{\varepsilon}_i^2}{n-2} \tag{3.1}
$$

Sachant que :

$$
\begin{aligned}
\widehat{\varepsilon}_i &= Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i = (\varepsilon_i - \overline{\varepsilon}) - (\widehat{\beta}_1 - \beta_1)(X_i - \overline{X}) \\
\sum_{i=1}^{n} \widehat{\varepsilon}_i^2 &= \sum_{i=1}^{n} (\varepsilon_i - \overline{\varepsilon})^2 - 2(\widehat{\beta}_1 - \beta_1) \sum_{j=1}^{n} (\varepsilon_i - \overline{\varepsilon})(X_i - \overline{X}) + (\widehat{\beta}_1 - \beta_1)^2 \sum_{i=1}^{n} (X_i - \overline{X})^2 \\
&= \sum_{i=1}^{n} (\varepsilon_i - \overline{\varepsilon})^2 - 2(\widehat{\beta}_1 - \beta_1) \sum_{i=1}^{n} \varepsilon_i (X_i - \overline{X}) + (\widehat{\beta}_1 - \beta_1)^2 \sum_{i=1}^{n} (X_i - \overline{X})^2 \\
&= \sum_{i=1}^{n} (\varepsilon_i - \overline{\varepsilon})^2 - (\widehat{\beta}_1 - \beta_1)^2 \sum_{i=1}^{n} (X_i - \overline{X})^2
\end{aligned}
$$

$$
\begin{aligned}
E(\sum_{i=1}^{n} \widehat{\varepsilon}_i^2) &= E(\sum_{i=1}^{n} (\varepsilon_i - \overline{\varepsilon})^2) - \sum_{i=1}^{n} (X_i - \overline{X})^2 V(\widehat{\beta}_1) \\
&= (n-1)\sigma^2 - \sigma^2 = (n-2)\sigma^2
\end{aligned}
$$

Ce qui permer de montrer que $E(\widehat{\sigma}^2) = E(\frac{\sum_{i=1}^n \widehat{\varepsilon}_i^2}{n-2}) = \sigma^2$. Faisons remarquer que la méthode du maximum de vraisemblance fournit un estimateur biaisé de $\sigma^2$.

**3.4. Théorème de Gauss-Markov.** Théorème : $\widehat{\beta}_0$ et $\widehat{\beta}_1$sont les meilleurs estimateurs linéaires et sans biais de $\beta_0$ et $\beta_1$.

## 4. Analyse de la variance.

**4.1. Equation de la variance.** On a :
$Y_i - \overline{Y} = \widehat{\beta}_0 + \widehat{\beta}_1 X_i + \widehat{\varepsilon}_i - \left(\widehat{\beta}_0 + \widehat{\beta}_1 \overline{X}\right) = \widehat{\beta}_1 (X_i - \overline{X}) + \widehat{\varepsilon}_i$
Ce qui implique que :

$$\begin{aligned}
\sum_{i=1}^{n} (Y_i - \overline{Y})^2 &= \sum_{i=1}^{n} \left(\widehat{\beta}_1^2 (X_i - \overline{X})^2 + 2\widehat{\beta}_1 (X_i - \overline{X})\widehat{\varepsilon}_i + \widehat{\varepsilon}_i^2\right) \\
&= \widehat{\beta}_1^2 \sum_{i=1}^{n} (X_i - \overline{X})^2 + 2\widehat{\beta}_1 \sum_{i=1}^{n} (X_i - \overline{X})\widehat{\varepsilon}_i + \sum_{i=1}^{n} \widehat{\varepsilon}_i^2 \\
&= \widehat{\beta}_1^2 \sum_{i=1}^{n} (X_i - \overline{X})^2 + \sum_{i=1}^{n} \widehat{\varepsilon}_i^2
\end{aligned}$$

On pose $SCT = \sum_{i=1}^n (Y_i - \overline{Y})^2$, $SCE = \widehat{\beta}_1^2 \sum_{i=1}^n (X_i - \overline{X})^2$, $SCR = \sum_{i=1}^n \widehat{\varepsilon}_i^2$. L'équation de la variance est définie par :

$$SCT = SCE + SCR$$

La somme des carrés totales est égale à la somme des carrés expliquées et la somme des carrés résiduelles. La $SCE$ a une importance dans l'appréciation de la qualité d'ajustement. Plus la SCE est élevée

Mh

mieux est la qualité d'ajustement. Faisons remarquer que la somme des carrés expliquées peut être définie autrement :

$$SCE = \widehat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \overline{X})^2 = \widehat{\beta}_1 \frac{S_{xy}}{S_{xx}} \sum_{i=1}^{n}(X_i - \overline{X})^2 = \widehat{\beta}_1 S_{xy}$$

**4.2. Le coefficient de détermination.** La qualité d'ajustement est mieux appréciée par la part de la variation expliquée dans la variation totale. cette part est appelée coefficient de détermination mesuré par :

$$R^2 = \frac{SCE}{SCT} = \frac{SCT - SCR}{SCT} = 1 - \frac{SCR}{SCT}$$

Dans le cadre de ce modèle simple, ce coefficient de détermination est égale au carré du coefficient de corrélation linéaire entre la variable $Y$ et la variable $X$. En effet :

$$R^2 = \frac{SCE}{SCT} = \frac{\widehat{\beta}_1 S_{xy}}{S_{yy}} = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{cov(X,Y)^2}{v(X)v(Y)} = \rho^2(X,Y)$$

Exemple :

$n = 20$, $\sum_{i=1}^{n} R_i = 1345.8$, $\sum_{i=1}^{n} C_i = 1221.1$, $\sum_{i=1}^{n} R_i^2 = 92520$, $\sum_{i=1}^{n} C_i^2 = 75969$, $\sum_{i=1}^{n} R_i C_i = 83826$. $SCT = \sum_{i=1}^{n} C_i^2 - n\overline{C}^2 = 1414.739$, $SCE = \widehat{\beta}_1 S_{CR} = 0.84 \times 1658.181 = 1392.872 \Longrightarrow R^2 = 0.984$

## 5. Inférence statistique

### 5.1. Lois de probabilité :

**:** Théorème : Sous les hypothèses H1-H5, on a :

mh

: i) $\widehat{\beta}_0 \curvearrowright N\left(\beta_0, \left\{\frac{1}{n} + \frac{\overline{X}^2}{S_{xx}}\right\}\sigma^2\right)$

: ii) $\widehat{\beta}_1 \curvearrowright N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right)$

: iii) $(n-2)\frac{\widehat{\sigma}^2}{\sigma^2} = \frac{SCR}{\sigma^2} \curvearrowright \chi^2(n-2)$

: iv) $\widehat{\beta}_0$ et $\widehat{\beta}_1$ sont indépendants de $\widehat{\sigma}^2$.

Demontration :

i) et ii) : $\widehat{\beta}_0$ et $\widehat{\beta}_1$s'écrivent comme combinaisons linéaires de lois normales $(\varepsilon_i)$

iv) L'indépendance entre les estimateurs de l'équation de celui de la variance, viend du fait de l'indépendance entre ces estimateurs et $\widehat{\varepsilon}_i$ pour tout $i$. Sachant que ces trois variables suivent la loi normales, l'indépendance est équivalente à une covariance nulle.

En effet,

$$\widehat{\varepsilon}_i = Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i = \varepsilon_i - \overline{\varepsilon} - \widehat{\beta}_1(X_i - \overline{X}) + \beta_1(X_i - \overline{X})$$
$$cov(\widehat{\varepsilon}_i, \widehat{\beta}_1) = cov(\varepsilon_i, \widehat{\beta}_1) - cov(\overline{\varepsilon}, \widehat{\beta}_1) - (X_i - \overline{X})v(\widehat{\beta}_1)$$
$$cov(\varepsilon_i, \widehat{\beta}_1) = cov(\varepsilon_i, \beta_1 + \frac{\sum_{j=1}^n (X_j - \overline{X})\varepsilon_j}{S_{xx}}) = \frac{X_i - \overline{X}}{S_{xx}}\sigma^2$$
$$cov(\overline{\varepsilon}, \widehat{\beta}_1) = \frac{1}{n}\sum_{i=1}^n \frac{X_i - \overline{X}}{S_{xx}}\sigma^2 = 0$$
$$\Longrightarrow cov(\widehat{\varepsilon}_i, \widehat{\beta}_1) = \frac{X_i - \overline{X}}{S_{xx}}\sigma^2 - 0 - (X_i - \overline{X})\frac{\sigma^2}{S_{xx}} = 0 \quad \forall i$$

### 5.2. Intervalles de confiance.

5.2.1. *Intervalles de confiance pour les paramètres de l'équation.* On sait que $\widehat{\beta}_1 \curvearrowright N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right) \Longrightarrow \sqrt{S_{xx}}\frac{\widehat{\beta}_1 - \beta_1}{\sigma} \curvearrowright N(0,1)$. En remplaçant $\sigma^2$ par son estimateur on a :

$$\begin{aligned} T &= \sqrt{S_{xx}}\frac{\widehat{\beta}_1-\beta_1}{\widehat{\sigma}} = \sqrt{S_{xx}}\frac{\widehat{\beta}_1-\beta_1}{\sigma\sqrt{\frac{\widehat{\sigma}^2}{\sigma^2}}} \\ &= \sqrt{S_{xx}}\frac{\widehat{\beta}_1-\beta_1}{\sigma\sqrt{(n-2)\frac{\widehat{\sigma}^2}{\sigma^2}\big/(n-2)}} \\ &= \frac{N(0,1)}{\sqrt{\frac{\chi^2(n-2)}{n-2}}} \curvearrowright ST(n-2) \end{aligned}$$

L'intervalle de confiance pour $\beta_1$ au nivreau $1-\alpha$ est tel que :

$$P\left(\left|\sqrt{S_{xx}}\frac{\widehat{\beta}_1-\beta_1}{\widehat{\sigma}}\right|<t\right)=1-\alpha \Leftrightarrow \widehat{\beta}_1 - t\frac{\widehat{\sigma}}{S_{xx}} < \beta_1 < \widehat{\beta}_1 - t\frac{\widehat{\sigma}}{S_{xx}}$$

Conclusion : les intervalles de confiance de $\beta_0$ et $\beta_1$ au niveau $1-\alpha$ sont définis par :

$$\begin{aligned} IC_{\beta_0}^{1-\alpha} &= \widehat{\beta}_0 \pm t \times \widehat{\sigma}(\widehat{\beta}_0) \\ IC_{\beta_1}^{1-\alpha} &= \widehat{\beta}_1 \pm t \times \widehat{\sigma}(\widehat{\beta}_1) \end{aligned}$$

où $\widehat{\sigma}(\widehat{\beta}_0)$ et $\widehat{\sigma}(\widehat{\beta}_1)$ définissent respectivment les écart-types estimés de $\widehat{\beta}_0$ et $\widehat{\beta}_1$ et la $t$ donné par la table de la loi de student à $n-2$ degré de liberté.

Exemple : $\widehat{\beta}_1 = 0.84$, $\widehat{\sigma}^2 = 1.214$, $\widehat{v}(\widehat{\beta}_1) = \frac{1.214}{1962.118}$, $n = 20$, $\alpha = 5\%$, $t = 2.101$ et $IC_{\beta_1}^{0.95} = \widehat{\beta}_1 \pm t \times \widehat{\sigma}(\widehat{\beta}_1) = 0.84 \pm 2.101 \times 0.024 = 0.84 \pm 0.05 = [0.789; 0.891]$. Avec une probabilité égale à 0.95, la propension marginale à consommer est comprise entre 0.789 et 0.891.

5.2.2. *Intervalle de confiance pour l'estimateur de la variance des erreurs.* On sait que $(n-2)\frac{\widehat{\sigma}^2}{\sigma^2} = \frac{SCR}{\sigma^2} \curvearrowright \chi^2(n-2)$. L'intervalle de confiance pour $\widehat{\sigma}^2$ au niveau $1-\alpha$ est tel que $P\left(\chi_1 < \frac{SCR}{\sigma^2} < \chi_2\right) = 1-\alpha$. $\chi_1$ et $\chi_2$ sont définis par :

$$P\left(\chi^2(n-2) < \chi_1\right) = \frac{\alpha}{2} et P\left(\chi^2(n-2) < \chi_2\right) = 1 - \frac{\alpha}{2}$$

Ce qui permet d'écrire :

$$IC_{\widehat{\sigma}^2}^{1-\alpha} = \left[\frac{SCR}{\chi_2}, \frac{SCR}{\chi_1}\right]$$

Exemple :$SCR = 0.548$, $n = 25$, $\alpha = 5\%$, $\chi_1 = 11.69$ et $\chi_2 = 38.08$ : $IC_{\widehat{\sigma}^2}^{0.95} = [0.014, 0.046]$

**5.3. Tests d'hypothèses.** $H_0 : \beta_1 = c$

Sous l'hypothèse $H_0$ au seuil $\alpha$, $\widehat{t} = \frac{\widehat{\beta}_1 - c}{\widehat{\sigma}(\widehat{\beta}_1)} \curvearrowright ST(n-2)$. On adopte la règle suivante : si $|\widehat{t}| > t$ on rejette $H_0$ et on dit que le paramètre est signifiactivement différent de $c$ au seuil $\alpha$ ; et inversement. Dans le cas où $c = 0$, ce test est appelé test de significativité. si $|\widehat{t}| > t$, on dit que le paramètre est significativement différent de zéro (la variable $X$ est pertinente dans l'explication de $Y$).

## 6. La prévision

Pour une valeur $X_p$, on peut construire une prévison de la variable $Y$ correspondant, noté $\widehat{Y}_p = \widehat{\beta}_0 + \widehat{\beta}_1 X_p$. L'erreur de prévision est définie par :

mh

$$\begin{aligned}
\widehat{\varepsilon}_p &= Y_p - \widehat{Y}_p = \beta_0 + \beta_1 X_p + \varepsilon_p - \widehat{\beta}_0 - \widehat{\beta}_1 X_p \\
E(\widehat{\varepsilon}_p) &= \beta_0 + \beta_1 X_p + E(\varepsilon_p) - E(\widehat{\beta}_0) - E(\widehat{\beta}_1) X_p = 0
\end{aligned}$$

$$\begin{aligned}
v(\widehat{\varepsilon}_p) &= v(\varepsilon_p - \widehat{\beta}_0 - \widehat{\beta}_1 X_p) \\
&= v(\varepsilon_p) + v(\widehat{\beta}_0) + X_p^2 v(\widehat{\beta}_1) + 2X_p cov(\widehat{\beta}_0, \widehat{\beta}_1) \\
&= \left[1 + \frac{1}{n} + \frac{(X_p - \overline{X})^2}{S_{xx}}\right] \sigma^2
\end{aligned}$$

Ceci nous permet de construire un intervalle de confiance pour $Y_p$ analogue à ceux des paramètres de l'équation de la regression :

$$IC_{Y_p}^{1-\alpha} = \widehat{Y}_p \pm t \left[1 + \frac{1}{n} + \frac{(X_p - \overline{X})^2}{S_{xx}}\right] \widehat{\sigma}^2$$

mh