2015-16

Méthodes Econométries

Module M2:
-Econométrie I

Parcours :
Economie et Gestion

S6 -Section B-

Vol. Horaire : 50h = 34 C + 12.5 Td +3.5 éval

Un peu d'histoire

La naissance de l'économétrie moderne

- L'économétrie moderne est née à la fin des années 30 et pendant les années 40. Elle est la résultante de trois phénomènes : le développement de la théorie de l'inférence statistique à la fin du XIX ème siècle ; la théorie macro-économique et la comptabilité nationale qui offrent des agrégats objectivement mesurables (contrairement à la microéconomie fondée sur l'utilité subjective) ; enfin, et surtout , la forte demande de travaux économétriques, soit de la part d'organismes publics de prévision et de planification , soit de la part d'entreprises qui ont de plus en plus besoin de modéliser la demande et leur environnement économique général. A partir des années 60, l'introduction de l'informatique et des logiciels standardisés va rendre presque routinière l'utilisation de l'économétrie .

La naissance de l'économétrie moderne

- En simplifiant de façon sans doute abusive l'on peut distinguer deux grandes périodes de la recherche économétrique moderne . Jusqu'à la fin des années 70 l'économétrie va étudier la spécification et la solvabilité de modèles macroéconomiques à équations simultanées . Puis à la suite de ce que l'on a appelé la révolution des anticipations rationnelles et de la critique de Lucas, la recherche se tournera davantage vers la microéconomie et l'analyse des séries temporelles .

Introduction générale

- L'économétrie est la branche des sciences économiques qui traite des modèles et des méthodes mathématiques appliquées aux grandeurs et variations économiques.
- Le calcul infinitésimal, les probabilités, la statistique, et la théorie des jeux, ainsi que d'autres domaines des mathématiques, sont utilisés pour analyser, interpréter et prévoir divers phénomènes économiques tels que les variations de prix sur le marché, l'évolution des coûts de production, le taux de croissance, le niveau du chômage, les variations du taux de change, les grandes tendances de l'économie à court et moyen terme qui permettent d'orienter la conduite d'une politique économique.

Introduction générale

- Les modèles utilisés ne permettent pas de prévoir, au sens strict, l'évolution des phénomènes économiques, mais davantage de construire des hypothèses et d'extrapoler des tendances futures à partir d'éléments actuels.

PROGRAMME DE CE SEMESTRE

- **Chapitre 0: Rappel sur les Tests d'Hypothèses et moments conditionnels**
- **Chapitre 1: Introduction au modèle de régression linéaire simple**
- **Chapitre 2: Le modèle de régression linéaire multiple**

BIBLIOGRAPHIE

Titre	Auteurs	Code
ECONOMETRIE	William Greene	
ECONOMETRIE : Théorie et techniques de base- exercices	M.Cohen-J.Pradel	
ECONOMETRIE : manuel et exercices corrigés	R.Bourbonnais	
Probabilités-Analyse des données et statistique	G.Saporta	
ECONOMETRIE : Le modèle linéaire général	F.B. doucouré	Econ:46

CH 0 Rappels

I LES TESTS STATISTIQUES

Introduction

- Soit $X_1, X_2, \ldots, X_n$ un échantillon aléatoire relatif à la V.A. parente X de loi $\mathscr{L}(\theta)$, où $\theta \in \Theta$ est un paramètre réel inconnu.
- Le semestre précédent, on cherchait à estimer θ. Mais il arrive qu'on ait une idée préconçue sur sa valeur: $\theta = \theta_0$
- On désire alors tester la validité de cette hypothèse, en la confrontant à une hypothèse alternative.

- Cette dernière exprime une tendance différente au sujet du paramètre.
- **Exemple :**
  - Est-ce que le taux de chômage au Maroc est p_0 ?
  - Est-ce que l'espérance de vie au Maroc est μ_0 ?...ou a augmenté ?

Méthodologie du test d'hypothèse

- On suppose que Θ est partitionné en Θ_0 et Θ_1:

$$\Theta = \Theta_0 \bigcup \Theta_1 \quad et \quad \Theta_0 \bigcap \Theta_1 = \emptyset$$

- Exprimons le fait que $\theta \in \Theta_0$ par

l'hypothèse H_0 et le fait que $\theta \in \Theta_1$ par H_1

➢ H_0 : « $\theta \in \Theta_0$ »

➢ H_1 : « $\theta \in \Theta_1$ »

➢ H_0 : s'appelle **l'hypothèse nulle**.

➢ H_1 : s'appelle **l'hypothèse alternative**.

➢ Si Θ_0 se réduit au seul point θ_0: $\Theta_0 = \{\theta_0\}$

H_0 devient H_0: « $\theta = \theta_0$ » et sera appelée **l'hypothèse simple**.

➢ Soit l'hypothèse simple H_0: « $\theta = \theta_0$ »

➢ Si H_1 est telle que H_1: « $\theta > \theta_0$ » ; alors on dit que le test est unilatéral à droite:

$$T.U.D. \begin{cases} H_0 : "\theta = \theta_0" \\ \# \\ H_1 : "\theta > \theta_0" \end{cases}$$

➢ Si H_1 est telle que H_1: « $\theta < \theta_0$ » ; alors on dit que le test est unilatéral à gauche:

$$T.U.G. \begin{cases} H_0 : "\theta = \theta_0" \\ \# \\ H_1 : "\theta < \theta_0" \end{cases}$$

➢ Si H_1 est telle que H_1: « $\theta \neq \theta_0$ » ;

alors on dit que le test est bilatéral :

$$T.B. \begin{cases} H_0 : "\theta = \theta_0" \\ \# \\ H_1 : "\theta \neq \theta_0" \end{cases}$$

- **Définition:** Un test d'hypothèse, est une règle de décision permettant, au vu de la réalisation $(x_1, x_2, ..., x_n)$ de l'E.A., de répondre à la question « dans lequel des deux sous ensemble se trouve θ ? »

- Cette règle de décision peut conduire à deux types d'erreurs:

- On rejette H_0 alors que H_0 est vraie: $\boxed{RH_0/H_0 vraie}$
- On l'appelle « **erreur de première espèce** »
- On ne rejette pas H_0 alors que H_1 est vraie: $\boxed{NRH_0/H_1 vraie}$;
- c'est « **l'erreur de seconde espèce** ».
- On définit alors la probabilité de commettre l'une ou l'autre erreur:

- 1)-

$$\alpha = P\left(RH_0 / H_0 Vraie\right) = P_0\left(RH_0\right)$$

- C'est le **risque** de première espèce.
- 2)-

$$\beta = P\left(NRH_0 / H_1 Vraie\right) = P_1\left(NRH_0\right)$$

- C'est le **risque** de seconde espèce.

- Voici un tableau résumant toutes les situations

Décision / Réalité	RH_0	NRH_0
H_0 Vraie	Erreur de 1ère espèce	Bonne Décision
H_1 Vraie	Bonne Décision	Erreur de 2ème espèce

Procédure à suivre

1- définir le paramètre θ à tester
2- Formuler les deux hypothèses H_0 et H_1
3- Préciser le Genre du test
4- Choisir la Statistique du test (bon estimateur de θ)
5- Préciser la loi de la statistique sous H_0
6- Ecrire la règle de Décision
7- Faire l'application numérique et décider
8- Conclure

I- Tests Relatifs à Une Proportion

1Test Unilatéral à gauche pour la proportion

- *i)* Formulation des hypothèses

$$T.U.G.\begin{cases} H_0 : "p = p_0" \\ \# \\ H_1 : "p < p_0" \end{cases}$$

Procédure à suivre

1- paramètre $p=\theta$ à tester
2- Formulation des deux hypothèses
3- TUG pour la proportion
4- F est un bon estimateur de p
5- si le T.C.L. est vérifié sous H_0 alors

$$Z = \frac{F - p_0}{\sqrt{p_0 q_0 / n}} \leadsto \mathcal{N}(0,1)$$

I- Tests Relatifs à Une Proportion

1Test Unilatéral à gauche pour la proportion

R.D

$$\begin{cases} Si\ f < c_\alpha = p_0 - z_\alpha\sqrt{\dfrac{p_0q_0}{n}} & \text{on rejette } H_0 \\ Si\ f \ge c_\alpha = p_0 - z_\alpha\sqrt{\dfrac{p_0q_0}{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

En effet :

- Notation :

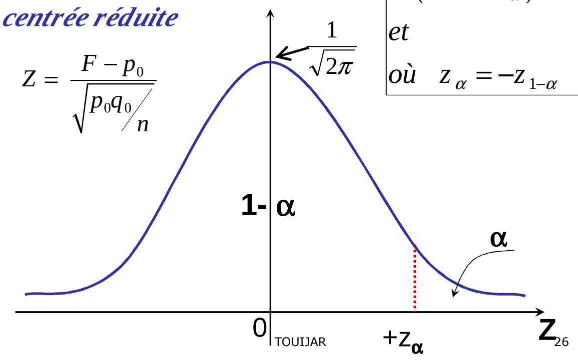
Courbe de densité de la loi normale centrée réduite

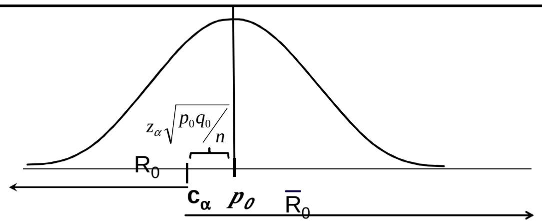
$$Z = \frac{F - p_0}{\sqrt{p_0q_0/n}}$$

où $P(Z > +z_\alpha) = \alpha$

et

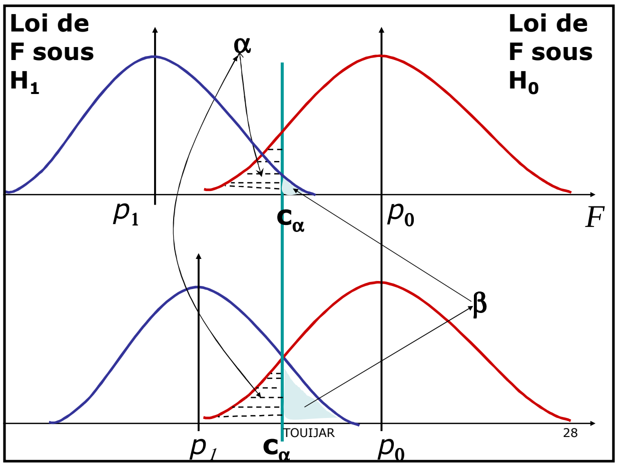
où $z_\alpha = -z_{1-\alpha}$





$$\alpha = P_0(RH_0) = P_0(F < c_\alpha) = P_0\left(\frac{F - p_0}{\sqrt{p_0q_0/n}} < \frac{c_\alpha - p_0}{\sqrt{p_0q_0/n}}\right)$$

$$\Rightarrow \frac{c_\alpha - p_0}{\sqrt{p_0q_0/n}} = -z_\alpha \Rightarrow \quad c_\alpha = p_0 - z_\alpha\sqrt{p_0q_0/n}$$



- **Remarque :** Les logiciels utilisent souvent **le niveau de signification** α_0 (*p-value*) d'une réalisation f de F : c'est le plus petit α à partir du quel on ne peut plus rejeter H_0 : $P_0(F < f) = \alpha_0$

Si dans notre exemple on suppose que $n=100$, $p_0=0.75$ $f=0{,}65$

Alors : $\alpha_0 = P_0(F < f)$

$= P_0(Z < -2{,}309) = 1\%$

Signe de la région de Rejet

Le test est donc significatif à 5%

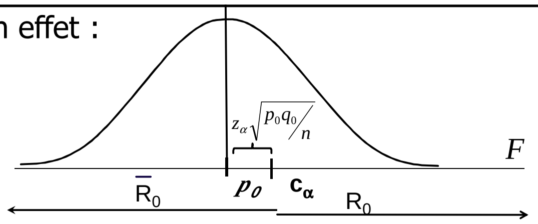
α_0 P-value	Interpretation
$\alpha_0 < 0{,}01$	very strong evidence against H_0
$0{,}01 \le \alpha_0 < 0{,}05$	moderate evidence against H_0
$0{,}05 \le \alpha_0 < 0{,}10$	suggestive evidence against H_0
$0{,}10 \le \alpha_0$	little or no real evidence against H_0

- **2- Test Unilatéral <u>à droite</u> pour la proportion**
- *i)-F.H.* $T.U.D.\begin{cases} H_0 : "p = p_0" \\ \# \\ H_1 : "p > p_0" \end{cases}$
- *ii)* si le T.C.L. est vérifié sous H_0 alors la règle de décision devient :

<u>R.D</u>

$$\begin{cases} Si\ f > c_\alpha = p_0 + z_\alpha \sqrt{\dfrac{p_0 q_0}{n}} & \text{on rejette } H_0 \\ Si\ f \le c_\alpha = p_0 + z_\alpha \sqrt{\dfrac{p_0 q_0}{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

En effet :



$$\alpha = P_0(RH_0) = P_0(F > c_\alpha) = P_0\left(\frac{F - p_0}{\sqrt{p_0q_0/n}} > \frac{c_\alpha - p_0}{\sqrt{p_0q_0/n}} \right)$$

$$\Rightarrow \frac{c_\alpha - p_0}{\sqrt{p_0q_0/n}} = +z_\alpha \Rightarrow c_\alpha = p_0 + z_\alpha\sqrt{p_0q_0/n}$$

- **Remarque :** La *p-value* ici vaut :

$$\alpha_0 = P_0(F > f)$$

Si on teste ***p = 0,75 # p > 0,75*** et si n=100 $f = 0{,}65$

Alors :

$$\alpha_0 = P_0(F > f) = P_0(Z > -2{,}309) = 99\%$$

Le test n'est donc pas significatif à 5% ni à 95%

3- Test Bilatéral pour la proportion

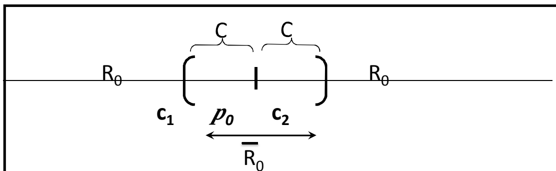
<u>F.H</u>

$$T.B. \begin{cases} H_0 : "p = p_0" \\ \# \\ H_1 : "p \neq p_0" \end{cases}$$

- Si T.C.L. On adopte la Règle de décision suivante :

<u>R.D</u>

$$\begin{cases} Si\ f \notin [c_1 ; c_2] & \text{on rejette } H_0 \\ Si\ f \in [c_1 ; c_2] & \text{on ne rejette pas } H_0 \end{cases}$$



or
$$f \in [c_1, c_2] \Leftrightarrow f \in [p_0 - c, p_0 + c] \Leftrightarrow |f - p_0| \le c$$
$$d'où\ f \notin [c_1, c_2] \Leftrightarrow |f - p_0| > c$$

$$\alpha = P_0(RH_0) = P_0(|F - p_0| > c) = P_0\left(\frac{|F - p_0|}{\sqrt{p_0q_0/n}} > \frac{c}{\sqrt{p_0q_0/n}} \right)$$

$$\Rightarrow \frac{c}{\sqrt{p_0q_0/n}} = z_{\alpha/2} \Rightarrow c = z_{\alpha/2}\sqrt{p_0q_0/n}$$

- D'où:
- <u>**R.D**</u>

$$\begin{cases} Si\ |f - p_0| > z_{\alpha/2}\sqrt{\dfrac{p_0q_0}{n}} & \text{on rejette } H_0 \\ Si\ |f - p_0| \le z_{\alpha/2}\sqrt{\dfrac{p_0q_0}{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

- Ou
$$\begin{cases} Si\ |z| > z_{\alpha/2} & \text{on rejette } H_0 \\ Si\ |z| \le z_{\alpha/2} & \text{on ne rejette pas } H_0 \end{cases}$$

- avec $z = \dfrac{f - p_0}{\sqrt{\dfrac{p_0q_0}{n}}}$

Remarque Si on teste $p = 0{,}75$ # $p \neq 0{,}75$

et si F vaut $f = 0{,}65$, alors la *p-value* vaut ici **:**

$$\alpha_0 = P_0\left(\left|\frac{F - p_0}{\sqrt{\frac{p_0 q_0}{n}}}\right| > \left|\frac{f - 0{,}75}{\sqrt{\frac{0{,}75 \times 0{,}25}{100}}}\right|\right)$$

$$= P_0\left(|Z| > +2{,}309\right) = 2 \times P_0\left(Z > +2{,}309\right)$$

$$= 2 \times 1\% = 2\%$$

Le test est donc significatif à **5%** et même à **2,5%**

II- Tests Relatifs à Une Moyenne

- **1- Test Unilatéral à droite pour la moyenne**

- **Question** : est-ce que l'espérance de vie des marocains a augmenté depuis le dernier recensement ?

- Pour répondre à cette question, on doit confronter 2 hypothèses:

•*i*)-

$$T.U.D.\begin{cases} H_0 : "\mu = \mu_0" \\ \# \\ H_1 : "\mu > \mu_0" \end{cases}$$

•*ii*)-On adopte ensuite une Règle de décision basée sur la statistique $\overline{X}$ et qui répondra à la question: à partir de quelle réalisation de $\overline{X}$ décidera-t-on du rejet de H_0 pour un risque α ?

$$\begin{cases} Si\ \overline{x} > c & \text{on rejette } H_0 \\ Si\ \overline{x} \leq c & \text{on ne rejette pas } H_0 \end{cases}$$

- **1)- Détermination de c en fonction du risque α :**

a)Si σ est connu et (X est normale ou $n \geq 30$)

$$\alpha = P_0(RH_0) = P_0\left(\overline{X} > c\right) = P_0\left(Z > \sqrt{n} \times \frac{c - \mu_0}{\sigma}\right)$$

$$\Rightarrow \frac{c - \mu_0}{\sigma} \times \sqrt{n} = z_\alpha \Rightarrow c = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$$

D'où la règle de décision :

$$\begin{cases} Si\ \overline{x} > c_\alpha = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ \overline{x} \leq c_\alpha = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

b)Si σ est inconnu et X est normale; on a la règle de décision :

$$\begin{cases} Si\ \overline{x} > c_\alpha = \mu_0 + t_{n-1;\alpha} \frac{s}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ \overline{x} \leq c_\alpha = \mu_0 + t_{n-1;\alpha} \frac{s}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

c)Si σ est inconnu et $n \geq 50$; on a la règle de décision :

$$\begin{cases} Si\ \overline{x} > c_\alpha = \mu_0 + z_\alpha \frac{s}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ \overline{x} \leq c_\alpha = \mu_0 + z_\alpha \frac{s}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

II- Tests Relatifs à Une Moyenne

- **2- Test Unilatéral à gauche pour la moyenne**

On doit confronter les deux hypothèses suivantes

i)- $T.U.G.\begin{cases} H_0 : "\mu = \mu_0" \\ \# \\ H_1 : "\mu < \mu_0" \end{cases}$

•*ii*)-On adopte ensuite une Règle de décision basée sur la statistique $\overline{X}$ et qui répondra à la question: à partir de quelle réalisation de $\overline{X}$ décidera-t-on du rejet de H_0 ?

$$\begin{cases} Si\ \overline{x} < c & \text{on rejette } H_0 \\ Si\ \overline{x} \geq c & \text{on ne rejette pas } H_0 \end{cases}$$

•De la même façon que précédemment, on a les règles de décision selon les cas :

a)Si σ est connu et (X est normale ou $n \geq 30$)

$$\begin{cases} Si\ \overline{x} < c_\alpha = \mu_0 - z_\alpha \dfrac{\sigma}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ \overline{x} \geq c_\alpha = \mu_0 - z_\alpha \dfrac{\sigma}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

b)Si σ est inconnu et X est normale; on a la règle de décision :

$$\begin{cases} Si\ \overline{x} < c_\alpha = \mu_0 - t_{n-1;\alpha} \dfrac{s}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ \overline{x} \geq c_\alpha = \mu_0 - t_{n-1;\alpha} \dfrac{s}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

c)Si σ est inconnu et $n \geq 50$; on a la règle de décision :

$$\begin{cases} Si\ \overline{x} < c_\alpha = \mu_0 - z_\alpha \dfrac{s}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ \overline{x} \geq c_\alpha = \mu_0 - z_\alpha \dfrac{s}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

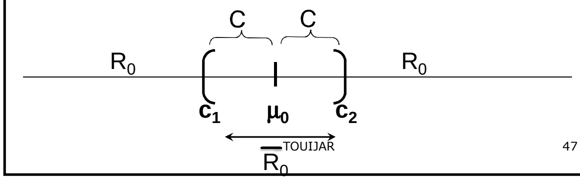
II- Tests Relatifs à Une Moyenne

3- Test Bilatéral pour la moyenne

i)- $T.B.\begin{cases} H_0 : "\mu = \mu_0" \\ \# \\ H_1 : "\mu \neq \mu_0" \end{cases}$

ii)-On adopte la Règle de décision suivante :

$$\begin{cases} Si\ \overline{x} \notin [c_1 ; c_2] & \text{on rejette } H_0 \\ Si\ \overline{x} \in [c_1 ; c_2] & \text{on ne rejette pas } H_0 \end{cases}$$



$$\overline{x} \in [c_1 ; c_2] \Leftrightarrow c_1 \leq \overline{x} \leq c_2 \Leftrightarrow \mu_0 - c \leq \overline{x} \leq \mu_0 + c$$

$$\Leftrightarrow -c \leq \overline{x} - \mu_0 \leq +c \Leftrightarrow |\overline{x} - \mu_0| \leq c$$

D'où la règle de décision suivante

$$\begin{cases} Si\ |\overline{x} - \mu_0| > c & \text{on rejette } H_0 \\ Si\ |\overline{x} - \mu_0| \leq c & \text{on ne rejette pas } H_0 \end{cases}$$

On a les règles de décision selon les cas :

a)Si σ est connu et (X est normale ou $n \geq 30$)

$$\begin{cases} Si\ |\bar{x}-\mu_0| > z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ |\bar{x}-\mu_0| \leq z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

b)Si σ est inconnu et X est normale

$$\begin{cases} Si\ |\bar{x}-\mu_0| > t_{n-1;\alpha/2}\dfrac{s}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ |\bar{x}-\mu_0| \leq t_{n-1;\alpha/2}\dfrac{s}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

c)Si σ est inconnu et $n \geq 50$; on a

$$\begin{cases} Si\ |\bar{x}-\mu_0| > z_{\alpha/2}\dfrac{s}{\sqrt{n}} & \text{on rejette } H_0 \\ Si\ |\bar{x}-\mu_0| \leq z_{\alpha/2}\dfrac{s}{\sqrt{n}} & \text{on ne rejette pas } H_0 \end{cases}$$

III- Tests Relatifs à Une Variance

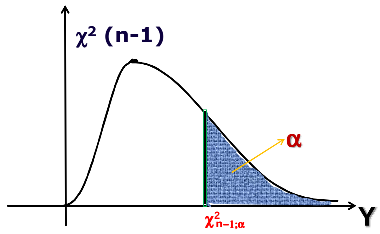
- Dans ce paragraphe on supposera la ***normalité*** de la population

- **1- Test Unilatéral à droite pour la variance**

$$T.U.D. \begin{cases} H_0 : "\sigma^2 = \sigma_0^2" \\ \# \\ H_1 : "\sigma^2 > \sigma_0^2" \end{cases}$$

RD.

$$\begin{cases} Si\ \ s^2 > c_\alpha & \text{on rejette } H_0 \\ Si\ \ s^2 \leq c_\alpha & \text{on ne rejette pas } H_0 \end{cases}$$



$Y = (n-1)\dfrac{S^2}{\sigma_0^2} \rightsquigarrow \chi^2$ (n-1) sous H_0

$$\alpha = P_0\left(Y > \chi^2_{n-1;\alpha}\right) = P_0\left(S^2 > \frac{\sigma_0^2}{n-1}\chi^2_{n-1;\alpha}\right) \Rightarrow c_\alpha = \frac{\sigma_0^2}{n-1}\chi^2_{n-1;\alpha}$$

De la même façon que précédemment, on a les règles de décision selon les cas :

a)Si μ est inconnue

La statistique utilisée est S²; d'où la R.D. est :

$$\begin{cases} Si\ \ s^2 > \chi^2_{n-1;\alpha}\dfrac{\sigma_0^2}{n-1} & \text{on rejette } H_0 \\ Si\ \ s^2 \leq \chi^2_{n-1;\alpha}\dfrac{\sigma_0^2}{n-1} & \text{on ne rejette pas } H_0 \end{cases}$$

2- Test Unilatéral à gauche pour la variance

$$T.U.G. \begin{cases} H_0 : "\sigma^2 = \sigma_0^2" \\ \# \\ H_1 : "\sigma^2 < \sigma_0^2" \end{cases}$$

R.D. (pour μ inconnue)

$$\begin{cases} Si\ \ s^2 < \chi^2_{n-1;1-\alpha}\dfrac{\sigma_0^2}{n-1} & \text{on rejette } H_0 \\ Si\ \ s^2 \geq \chi^2_{n-1;1-\alpha}\dfrac{\sigma_0^2}{n-1} & \text{on ne rejette pas } H_0 \end{cases}$$

3- Test Bilatéral pour la variance

$$T.B.\begin{cases} H_0 : "\sigma^2 = \sigma_0^2" \\ \# \\ H_1 : "\sigma^2 \neq \sigma_0^2" \end{cases}$$

Si μ est inconnue

$$\begin{cases} Si\ \ s^2 \notin \left[\chi^2_{n-1;1-\alpha/2} \dfrac{\sigma_0^2}{n-1} ; \chi^2_{n-1;\alpha/2} \dfrac{\sigma_0^2}{n-1} \right] \\ \text{on rejette } H_0 \\ Si\ \ s^2 \in \left[\chi^2_{n-1;1-\alpha/2} \dfrac{\sigma_0^2}{n-1} ; \chi^2_{n-1;\alpha/2} \dfrac{\sigma_0^2}{n-1} \right] \\ \text{on ne rejette pas } H_0 \end{cases}$$

- **1- Test Bilatéral de comparaison de 2 Variances :**
- On suppose la normalité des 2 populations
- F.H.

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ \# \\ H_1 : \sigma_1^2 \neq \sigma_2^2 \end{cases} \Leftrightarrow \begin{cases} H_0 : \sigma_1^2 / \sigma_2^2 = 1 \\ \# \\ H_1 : \sigma_1^2 / \sigma_2^2 \neq 1 \end{cases}$$

Cas où μ est inconnue

- Rapport critique (sous H_0):

$R_c = \dfrac{S_1^2}{S_2^2}$ Si $s_1^2 > s_2^2$

Sinon $R_c = \dfrac{S_2^2}{S_1^2}$ et F.H. devient alors $\begin{cases} H_0 : \sigma_2^2 / \sigma_1^2 = 1 \\ H_1 : \sigma_2^2 / \sigma_1^2 \neq 1 \end{cases}$

- **Loi de R_c** sous H_0 :

$$R_c \sim \mathcal{F}(\mathcal{V}_1 , \mathcal{V}_2)$$

- **R.D.**

$$\begin{cases} Si\ \ R_c \notin \left[F_{1-\alpha/2}(\nu_1;\nu_2) ; F_{\alpha/2}(\nu_1;\nu_2) \right] \\ \text{on rejette } H_0 \\ Si\ \ R_c \in \left[F_{1-\alpha/2}(\nu_1;\nu_2) ; F_{\alpha/2}(\nu_1;\nu_2) \right] \\ \text{on ne rejette pas } H_0 \end{cases}$$

$$où \quad \nu_i = n_i - 1$$

2- Test Bilatéral de comparaison de 2 moyennes

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ \# \\ H_1 : \mu_1 \neq \mu_2 \end{cases} \Leftrightarrow \begin{cases} H_0 : \ \mu_1 - \mu_2 = 0 \\ \# \\ H_1 : \ \mu_1 - \mu_2 \neq 0 \end{cases}$$

- On propose $\bar{X}_1 - \bar{X}_2$ comme estimateur de $\mu_1 - \mu_2$:

a) Si σ_1 et σ_2 connus et (X_1 normale ou $n_1 \geq 30$) et (X_2 normale ou $n_2 \geq 30$)

R.D.

$$\begin{cases} Si\,|\bar{x}_1-\bar{x}_2| > z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} & \text{on rejette } H_0 \\ Si\,|\bar{x}_1-\bar{x}_2| \le z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} & \text{on ne rejette pas } H_0 \end{cases}$$

b) Si σ_1 et σ_2 inconnus mais égales et X_1 et X_2 normales, alors la R.D. est la suivante :

$$\begin{cases} Si\,|\bar{x}_1-\bar{x}_2| > t_{\alpha/2;n_1+n_2-2}\hat{s}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \quad \text{on rejette } H_0 \\ Si\,|\bar{x}_1-\bar{x}_2| \le t_{\alpha/2;n_1+n_2-2}\hat{s}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \quad \text{on ne rejette pas } H_0 \end{cases}$$

• Où

$$\hat{S}^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{(n_1-1)+(n_2-1)}$$

• **c)** Si σ_1 et σ_2 inconnus et $n_1 \ge 50$, $n_2 \ge 50$ alors la R.D. est la suivante :

$$\begin{cases} Si\,|\bar{x}_1-\bar{x}_2| > z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} & \text{on rejette } H_0 \\ Si\,|\bar{x}_1-\bar{x}_2| \le z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} & \text{on ne rejette pas } H_0 \end{cases}$$

CH 0 Rappels II Moments Conditionnels

- RAPPELS
- Moments Conditionnels :
- On suppose que Y est une V.A.R. mais que X peut être qualitative;
- **L'Espérance conditionnelle :**
- On appelle espérance de Y sachant X=x; notée $E(Y/X=x)$; la quantité définie par

$$E(Y/X=x)=\sum_y y\,P(Y=y/X=x)$$

- ***Rem*** : $E(Y/X=x)=\varphi(x)$ est une fonction de x appelée fonction de régression de Y en X

- ***Définition*** :
- On appelle espérance conditionnelle de Y sachant X; notée $E(Y/X)$; la variable aléatoire définie par :

$$E(Y/X)=\varphi(X)$$

- ***Propriétés*** :
- 1) Linéarité: $E(Y_1+Y_2/X)=E(Y_1/X)+E(Y_2/X)$
- 2) Théorème de l'espérance Totale :

$$E\left[E(Y/X)\right]=E(Y)$$

- **La Variance conditionnelle :**
- ***Définition*** :
- On appelle Variance conditionnelle de Y sachant X=x; notée $V(Y/X=x)$; la quantité

$$V(Y/X=x)=E\left[\left(Y-E(Y/X=x)\right)^2/X=x\right]$$

- De là on peut définir la V.A. variance conditionnelle :

$$V(Y/X)=E\left[\left(Y-E(Y/X)\right)^2/X\right]$$

- Théorème de la variance Totale :

$$V(Y)=E\left[V(Y/X)\right]+V\left[E(Y/X)\right]$$

- **Approche géométrique :**
- Soit le produit scalaire de 2 V.A. X et Y, noté <X , Y>, définie comme suit

<X , Y>= E(XY)

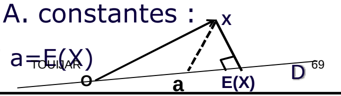
Et soit ||.|| la norme associée à ce produit scalaire : $\|X\|=\sqrt{\langle X,X\rangle}=\sqrt{E\left(X^2\right)}$

- Si Y et X sont deux variables centrées, alors

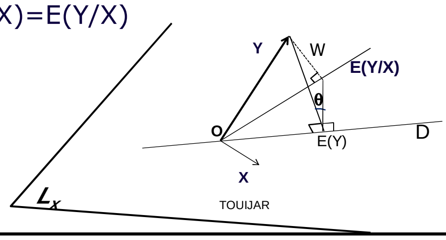
<X,Y>=E(XY)=cov(X,Y) et $\|X\|=\sqrt{E\left(X^2\right)}=\sigma_X$

- E(X) est la projection orthogonale de X sur la droite D des V.A. constantes :

E[(X-a)²] minimale si a=E(X)



- **Approche géométrique :**
- Soit maintenant l'espace L_x des V.A. fonctions de X (i.e. du type φ(X)); alors E(Y/X) est la projection orthogonale de Y sur L_x (les propriétés d'un projecteur orthogonale sont vérifiées grâce aux propriétés de l'espérance conditionnelle).D'où E[(Y-φ(X))²] minimale si φ(X)=E(Y/X)



Chapitre 1

Introduction à La Régression Linéaire Simple :

MRLS

Introduction :

- La RLS (régression linéaire simple) permet d'étudier la liaison (supposée *linéaire*) entre *deux* variables *quantitatives* X et Y; où Y (variable endogène) sera expliquée par X (variable exogène).
- Autrement, on cherche à prévoir le comportement moyen de Y en fonction de la V.A. X : $\hat{Y}=\varphi(X)$ tel que cet estimateur soit sans biais : $E(\hat{Y})=E(Y)$, et à mesurer l'erreur de prévision : $V(Y-\hat{Y})$

- Nous allons nous intéressé à l'aspect *inférentiel* de cette régression.
- On supposera, dorénavant, que la fonction φ de la V.A. X est linéaire (Régression Linéaire) et on traitera alors cette régression sous deux volets; le volet théorique ensuite le volet où on ne travaillera qu'avec les réalisations de X et Y sur un échantillon.

- **Modèle théorique de la régression simple**
- On mesure la qualité de cette approximation par le rapport

$$\eta^2_{Y/X} = \frac{V\left(E\left(Y/X\right)\right)}{V(Y)} = \frac{Variance\ expliquée}{Variance\ Totale} = Cos^2(\theta)$$

- La fonction qui à $x \mapsto E(Y/X=x)$ est appelé fonction de régression, dont la représentation graphique s'appelle courbe de régression de Y en X : on écrit $Y = E(Y/X) + W$. Or puisque

$$E(Y) = \underbrace{E\left[E(Y/X)\right]}_{=E(Y)} + E(W) \Rightarrow E(W) = 0$$

et d'après le graphique ci-dessus W est non corrélée ni avec X ni avec E(Y/X) (car$\perp$) d'où

$$V(Y) = V\left[E(Y/X)\right] + V(W)$$

$$\Rightarrow \eta^2_{Y/X} = \frac{V(Y) - V(W)}{V(Y)} = 1 - \frac{V(W)}{V(Y)}$$

$$\Rightarrow V(W) = \sigma^2(W) = \left(1 - \eta^2_{Y/X}\right)V(Y)$$

- **Cas où la régression est Linéaire**
- C'est le cas où $E(Y/X)=\alpha_0+\alpha_1 X$; on a alors

$$Y = \alpha_0 + \alpha_1 X + W$$

$$E\left(Y/X\right) = \alpha_0 + \alpha_1 X$$

$$E(Y) = \alpha_0 + \alpha_1 E(X)$$

$$\Rightarrow Y - E(Y) = \alpha_1\left(X - E(X)\right) + W$$

$$\Rightarrow E\left[\left(Y - E(Y)\right)\left(X - E(X)\right)\right] = \alpha_1 E\left[\left(X - E(X)\right)^2\right] + \underbrace{E\left[W\left(X - E(X)\right)\right]}_{=0}$$

$$\Rightarrow Cov(X,Y) = \alpha_1 V(X) \Rightarrow \alpha_1 = \frac{Cov(X,Y)}{V(X)} = \rho\frac{\sigma_Y}{\sigma_X}$$

- **Cas où la régression est Linéaire et simple**
- L'équation de la droite s'écrit alors

$$E\left(Y/X\right) - E(Y) = \frac{Cov(X,Y)}{V(X)}\left(X - E(X)\right)$$

$$\Rightarrow Y = E(Y) + \rho\frac{\sigma_Y}{\sigma_X}\left(X - E(X)\right) + W$$

$$\Rightarrow V(Y) = \rho^2\frac{\sigma_Y^2}{\sigma_X^2}V(X) + V(W)$$

$$= \rho^2 V(Y) + V(W)$$

$$\Rightarrow \rho^2 = \eta^2_{Y/X}$$

- **Modèle de la régression linéaire avec données Expérimentales**
- **Il y a trois types différents de données :**

Données en coupes instantanées (ou transversales : Cross-sections)

- Exemple : Si on prend le modèle keynésien classique:

$$C_i = \alpha_0 + \alpha_1 R_i + W_i \quad \forall i = 1,\ldots,n$$

- *C_i et R_i observés pour $i = 1, \ldots, n$*
- *i =individu, ménage, entreprise,... et n = nombre total* d'observations

- Modèle de la régression linéaire avec données Expérimentales

- **Séries chronologiques** **(ou séries temporelles** :Time series**)**
- Exemple :

$$C_t = \alpha_0 + \alpha_1 R_t + W_t \quad \forall t = 1, \ldots, T$$

- *C_t et R_t observés pour $t = 1, \ldots, T$*
- *T = année, trimestre, mois,... et T = nombre total d'observations*

TOUIJAR 79

- Modèle de la régression linéaire avec données Expérimentales

- ***Données de panel*** *:* Panel data **(ou données individuelles-temporelles)**
- Exemple :

$$C_{it} = \alpha_0 + \alpha_1 R_{it} + W_{it} \quad \forall t = 1, \ldots, T \quad \forall i = 1, \ldots, n$$

- *C_{it} et R_{it} observés pour $i = 1, \ldots, n$; $t = 1, \ldots, T$*

TOUIJAR 80

- Modèle de la régression linéaire avec données Expérimentales : 1er type
- Soient les données (x_i, y_i) un échantillon de taille n, relatif à (X, Y).
- Régression linéaire $\Rightarrow E\left(Y/X\right) = \alpha_0 + \alpha_1 X$
- Le modèle s'écrit alors :

$$y_i = \alpha_0 + \alpha_1 x_i + w_i \quad \forall i = 1, \ldots, n$$

- *Où w_i sont des réalisations indépendantes de W d'espérance nulle et de variance* σ^2 constante $\forall x_i$
- On l'appelle *modèle linéaire*

TOUIJAR 81

Terminology for Simple Regression

y	*x*
Dependent Variable	Independent Variable
Explained Variable	Explanatory Variable
Response Variable	Control Variable
Predicted Variable	Predictor Variable
Regressand	Regressor

TOUIJAR 82

- Les coefficients(paramètres) α_0 et α_1 du modèle, ainsi que $\sigma^2(W)$ seront estimés puis testés à partir des données (x_i, y_i) de l'échantillon.

TOUIJAR 83

- ***Exemples de Modèles linéaires et non Linéaires:***

1)- Soit la fonction keynésienne :

$$C_t = \alpha_0 + \alpha_1 R_t + W_t \quad \forall t = 1, \ldots, n$$

Où C : consommation
R : le revenu
α_1 : propension marginale à consommer
α_0 : consommation incompressible.
w : l'erreur (bruit)

C'est un modèle linéaire :MRLS

TOUIJAR 84

- ***Exemples de Modèles linéaires et non Linéaires:***

- *2)-* $y = \alpha x^{\beta}$ un modèle non-linéaire mais qu'on peut rendre linéaire en composant par le logarithme :

$$Ln(y) = Ln(\alpha) + \beta Ln(x)$$

- Modèle très utilisé en économétrie (élasticité constante de y/x ; où β est le coefficient d'élasticité

- ***Exemples de Modèles linéaires et non Linéaires:***

- *3)-* $y = \alpha \exp(\beta x)$ un modèle non-linéaire mais qu'on peut linéariser en posant :

$$y' = Ln(y) \Rightarrow y' = \alpha' + \beta x$$

- Modèle à croissance exponentielle

- ***Exemples de Modèles linéaires et non Linéaires:***

- *4)-* $y = \dfrac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)}$ un modèle non-linéaire

mais qu'on peut rendre linéaire en posant :

$$y' = Ln\left(\frac{y}{1-y}\right) \Rightarrow y' = \alpha + \beta x$$

- Modèle logistique qui étudie les variations d'un taux de réponse y en fonction d'une excitation x

- ***Exemples de Modèles linéaires et non Linéaires:***

- *5)-* $y = \alpha + \exp(\beta x)$

  est un modèle non-linéarisable

- *6)-* $y = \alpha + \beta x + \gamma x^2$ est un modèle linéaire mais pas simple : à 2 variables explicatives.

- ***Estimations des paramètres α , β et σ²:***

- ***Exemple Introductif***

Le tableau ci-dessous représente le produit national brute (PNB) et la demande des biens de première nécessité obtenus pour la période 1969-1980 dans un certain pays.

Nous cherchons à estimer la demande des biens de première nécessité en fonction du PNB suivant le modèle linéaire simple:

$$y_i = \alpha_0 + \alpha_1 x_i + w_i \;;\quad i = 1,\ldots,n$$

Année	PNB: X	La Demande : Y
1969	50	6
1970	52	8
1971	55	9
1972	60	12
1973	57	8
1974	58	10
1975	62	12
1976	65	9
1977	68	11
1978	69	10
1979	70	11
1980	78	14

- ***Estimations des paramètres α , β et σ^2:***

- ***Exemple Introductif***

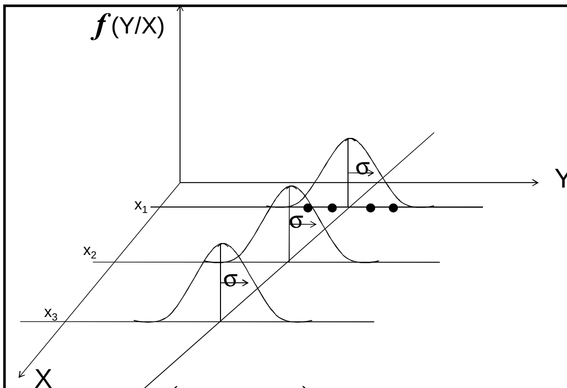
Pour effectuer une telle étude, on doit préciser certaines conditions d'application (hypothèses fondamentales)

Hypothèses

- **H1:** Le modèle est linéaire en X
- **H2**: Exogénéité de la Variable indépendante:
  $E(W_i\ /X)=0$
- **H3**: Homoscédasticité et absence d'autocorrélation:
  - $\sigma^2 = V(W_i)$ est constante quelque soit i
  - $Cov(W_i\ ,\ W_j)=0$ pour $i\neq j$
- **H4**: Les W_i sont de même loi normale :
  $W_i \rightsquigarrow \mathcal{N}(0,\sigma^2)$

Remarques

- **1)-** La normalité des erreurs W_i entraîne celle des Y_i ; en effet
- $W_i \rightsquigarrow \mathcal{N}(0,\sigma^2) \Rightarrow Y_i/_{X_i=x_i} \rightsquigarrow \mathcal{N}\left(\alpha_0+\alpha_1 x_i\ ;\ \sigma^2\right)$
- Graphiquement **:**



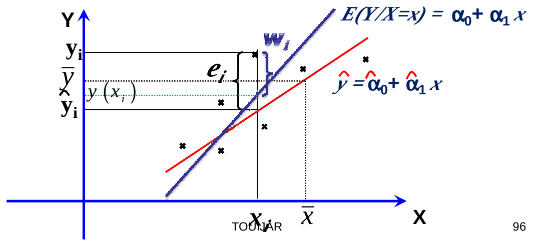
$E\left(Y\ /X\ =x\right)=\alpha_0+\alpha_1 x$

Remarques

- **2)-** $E(W_i\ /X)=0 \ \Rightarrow\ E(W_i)=0$ et $cov(W_i\ ,X)=0$
- **3)-** La normalité des erreurs W_i n'est pas nécessaire pour l'estimation des paramètres α_0 et α_1 par la méthode des moindres carrés.
- **4)-** Il suffit d'avoir la non corrélation des erreurs et leur normalité pour avoir leur indépendance
- **5)-** On pourra ajouter une autre hypothèse : Les X_i sont certaines (non aléatoires); ceci nous évitera d'utiliser les espérances conditionnelles

I- Estimation des paramètres du MRLS

- 1- La méthode des MCO
  - On détermine la droite passant, le plus proche, de tous les points du nuage (x_i,y_i)



- Notons par $\hat{y} = \hat{\alpha}_0 + \hat{\alpha}_1 x$ l'équation de la droite de régression recherchée
- Et $e_i = y_i - \hat{y}_i$ la valeur résiduelle.
- On désire calculer les valeurs $\hat{\alpha}_0$ et $\hat{\alpha}_1$ qui minimisent la somme des carrés des résidus :

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{\alpha}_0 - \hat{\alpha}_1 x_i)^2 = F(\hat{\alpha}_0, \hat{\alpha}_1)$$

- On notera désormais α_0 et α_1 à la place de $\hat{\alpha}_0$ et $\hat{\alpha}_1$
- En annulant les dérivées par rapport à α_0 et α_1 :

$$\frac{\partial F(\alpha_0, \alpha_1)}{\partial \alpha_0} = 0 \Leftrightarrow -2\sum_{i=1}^{n}(y_i - \alpha_0 - \alpha_1 x_i) = 0 \quad (1)$$

$$\Leftrightarrow \bar{y} = \alpha_0 + \alpha_1 \bar{x}$$

- Donc la droite passe par le point moyen G $(\bar{x}; \bar{y})$

$$\frac{\partial F(\alpha_0, \alpha_1)}{\partial \alpha_1} = 0 \Leftrightarrow -2\sum_{i=1}^{n} x_i (y_i - \alpha_0 - \alpha_1 x_i) = 0 \quad (2)$$

$$\Leftrightarrow \sum_{i=1}^{n} x_i y_i - \alpha_0 \sum_{i=1}^{n} x_i - \alpha_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\Leftrightarrow \overline{xy} - \alpha_0 \bar{x} - \alpha_1 \overline{x^2} = 0$$

$$\Leftrightarrow \alpha_1 = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{Cov(x, y)}{V(x)}$$

- (1) et (2) nous donnent les équations :

$$\sum_{i=1}^{n} e_i = 0 \quad et \quad \sum_{i=1}^{n} x_i e_i = 0$$

- Et d'autre part :

$$\hat{\alpha}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad et \quad \hat{\alpha}_0 = \bar{y} - \hat{\alpha}_1 \bar{x}$$

- $\hat{\alpha}_1$ s'écrit aussi

$$\hat{\alpha}_1 = \frac{n\sum (x_i y_i) - \sum x_i \sum y_i}{n\sum (x_i)^2 - \left(\sum x_i\right)^2}$$

Les Données centrées

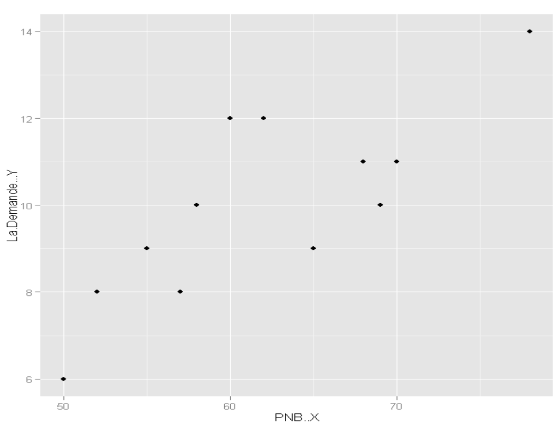
- En utilisant les données centrées:

$$x_{0i} = x_i - \bar{x} \quad et \quad y_{0i} = y_i - \bar{y}$$

d'où A_1 s'écrit

$$A_1 = \frac{\sum_{i=1}^{n} (x_{0i} Y_{0i})}{\sum_{i=1}^{n} x_{0i}^2}$$

X	Y	X0	X02	Y0	X0 Y0	Y^	ei	ei xi	ei2
50	6	-12	144	-4	48	7,511	-1,511	-75,532	2,282
52	8	-10	100	-2	20	7,926	0,074	3,872	0,006
55	9	-7	49	-1	7	8,548	0,452	24,867	0,204
60	12	-2	4	2	-4	9,585	2,415	144,894	5,832
57	8	-5	25	-2	10	8,963	-0,963	-54,878	0,927
58	10	-4	16	0	0	9,170	0,830	48,128	0,689
62	12	0	0	2	0	10,000	2,000	124,000	4,000
65	9	3	9	-1	-3	10,622	-1,622	-105,452	2,632
68	11	6	36	1	6	11,245	-0,245	-16,638	0,060
69	10	7	49	0	0	11,452	-1,452	-100,197	2,109
70	11	8	64	1	8	11,660	-0,660	-46,170	0,435
78	14	16	256	4	64	13,319	0,681	53,106	0,464
744	**120**	**0**	**752**	**0**	**156**	**120,000**	**0,000**	**0,000**	**19,638**
62	10	0		0		10	α^1 = 0,2074	α^0 = -2,862	



- Exemple: la droite de régression empirique dans notre cas, a pour équation : $\hat{y} = 0.2074 \times x - 2.862$
- Elle permet d'obtenir une estimation de la demande *moyenne* des biens de 1ère nécessité pour un PNB x.
- Rem : 1)- si x (PNB)=0 alors $\hat{y} = -2.862$ valeur qui n'a aucune signification concrète.
- 2) si x'=x+1, alors $y' = \hat{y} + 0.2074$; une aug$^{\theta}$ de 1 du PNB entraine une aug$^{\theta}$ de la demande moyenne de 0,21.

- Théorème 1:

A_0, A_1 *et* $\hat{Y}$ *sont des estimateurs sans Biais de* α_0, α_1 *et de* $E^x(Y)$

$$\boxed{A_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}}$$

Où $E^x(Y)$ *désignera* $E(Y / X = x)$

- Remarque :

$$\boxed{\hat{\alpha}_0, \hat{\alpha}_1 \text{ et } \hat{y} \text{ sont des réalisations de } A_0, A_1 \text{ et } \hat{Y}}$$

- Dem : on doit montrer que $E(A_1) = \alpha_1$

Pour ce on doit montrer $E^{x_i}(A_1) = \alpha_1$

En effet $E(E^{x_i}(A_1)) = E(A_1) = \alpha_1$

$$\boxed{E^{x_i}(A_1) = \frac{\sum (x_i - \bar{x}) E^{x_i}((Y_i - \bar{Y}))}{\sum (x_i - \bar{x})^2}}$$

Or $\boxed{E^{x_i}(Y_i) = \alpha_0 + \alpha_1 x_i}$

d'où

$$E^{x_i}(\bar{Y}) = \alpha_0 + \alpha_1 \bar{x} \Rightarrow E^{x_i}(Y_i - \bar{Y}) = \alpha_1 (x_i - \bar{x})$$

$$\Rightarrow E^{x_i}(A_1) = \alpha_1$$

- De même puisque :

$\hat{\alpha}_0 = \bar{y} - \hat{\alpha}_1 \bar{x}$, $\hat{\alpha}_0$ étant une réalisation de A_0 ; on a :

$$E^{x_i}(A_0) = E^{x_i}(\bar{Y}) - \bar{x}_1 \underbrace{E^{x_i}(A_1)}_{\alpha_1}$$

$$= \alpha_0 + \alpha_1 \bar{x} - \bar{x}_1 \alpha_1 = \alpha_0$$

$$\Rightarrow E(E^{x_i}(A_0)) = E(A_0) = \alpha_0$$

- Pour la troisième estimation :

$$\boxed{\begin{array}{l} E^x(Y) = \alpha_0 + \alpha_1 x \Rightarrow \hat{\alpha}_0 + \hat{\alpha}_1 x = \hat{y} \\ \text{est une estimation sans Biais de } \alpha_0 + \alpha_1 x = E^x(Y) \\ \Rightarrow \hat{Y} = A_0 + A_1 X \quad \text{est un estimateur sans} \\ \text{Biais de } E^x(Y) \end{array}}$$

- On montre de plus que A_1 et $\bar{Y}$ sont non corrélés; pour ce, on commence par montrer qu'elles le sont conditionnellement à x_i :

- En utilisant les données centrées:

$\boxed{x_{0i} = x_i - \bar{x} \text{ et } Y_{0i} = Y_i - \bar{Y}}$ d'où A_1 s'écrit

$$A_1 = \frac{\sum x_{0i} Y_{0i}}{\sum (x_{0i})^2} = \frac{\sum x_{0i} Y_i}{\sum (x_{0i})^2} = \sum \omega_i Y_i \quad \text{où} \quad \omega_i = \frac{x_{0i}}{\sum (x_{0i})^2}$$

D'où $Cov^{x_i}(A_1, \bar{Y}) = Cov^{x_i}\left(\sum \omega_i Y_i, \bar{Y}\right)$

$$= \sum \omega_i \, Cov^{x_i}(Y_i, \bar{Y}) = \sum \omega_i \, Cov^{x_i}\left(Y_i, \frac{1}{n} \sum Y_j\right)$$

$$= \frac{\sigma^2}{n} \underbrace{\sum \omega_i}_{=0} = 0$$

Par conséquent elles sont non corrélées marginalement

Propriétés et Distributions des Estimateurs A_1 et A_0

- **Théorème 2 (GAUSS-MARKOV) :**

A_0 *et* A_1 sont des estimateurs BLUE de α_0 *et* α_1

Dém : on a déjà montré qu' ils sont sans biais.
De plus

$$V^{x_i}(A_1) = \sum\left[\omega_i^2 V^{x_i}(Y_i)\right] = \sigma^2 \sum \omega_i^2 = \frac{\sigma^2}{\sum (x_{0i})^2}$$

car

$$V^{x_i}(Y_i) = V(\alpha_0 + \alpha_1 X_i + W_i / X_i = x_i) = V(W_i) = \sigma^2$$

Pour la linéarité, on a déjà vu que :

$$A_1 = \sum \omega_i Y_i \quad où \quad \omega_i = \frac{x_{0i}}{\sum (x_{0i})^2}$$

Reste à montrer que la variance est minimale; pour ce supposons qu'il existe un autre estimateur A' linéaire et sans Biais (où on notera E à la place de E^{x_i} et V au lieu de V^{x_i}) :

$$A' = \sum k_i Y_i \quad et \quad E(A') = \alpha_1 \Rightarrow \sum k_i E(Y_i) = \alpha_1$$

$$\Rightarrow \sum k_i (\alpha_0 + \alpha_1 x_i) = \alpha_1$$

$$\Rightarrow \alpha_0 \sum k_i + \alpha_1 \sum k_i x_i = \alpha_1$$

$$\Rightarrow \sum k_i = 0 \quad et \quad \sum k_i x_i = 1$$

Or

$$V(A') = \sum k_i^2 V(Y_i) = \sigma^2 \sum k_i^2$$

$$= \sigma^2 \sum (k_i - \omega_i + \omega_i)^2 = \sigma^2 \sum (k_i - \omega_i)^2 + \sigma^2 \sum (\omega_i)^2 + 2\sigma^2 \sum \omega_i (k_i - \omega_i)$$

$$= \sigma^2 \sum (k_i - \omega_i)^2 + \sigma^2 \sum (\omega_i)^2 - 2\sigma^2 \sum (\omega_i)^2 + 2\sigma^2 \sum \omega_i k_i$$

$$= \sigma^2 \sum (k_i - \omega_i)^2 - \underbrace{\sigma^2 \sum (\omega_i)^2}_{V(A_1)} + \underbrace{\frac{2\sigma^2}{\sum x_{0i}^2}}_{2V(A_1)} \underbrace{\sum x_i k_i}_{=1} - \frac{2\sigma^2}{\sum x_{0i}^2} \bar{x} \underbrace{\sum k_i}_{=0}$$

$$= \sigma^2 \sum (k_i - \omega_i)^2 + 2V(A_1) - V(A_1)$$

$$= \underbrace{\sigma^2 \sum (k_i - \omega_i)^2}_{\geq 0} + V(A_1) \geq V(A_1)$$

De même pour A_0; on a vu que :

$$A_0 = \bar{Y} - \bar{X} A_1 = \frac{1}{N} \sum Y_i - \bar{X} \frac{\sum X_{0i}(Y_i - \bar{Y})}{\sum X_{0i}^2}$$

$$= \frac{1}{N} \sum Y_i - \bar{X} \frac{\sum X_{0i} Y_i - \bar{Y} \underbrace{\sum X_{0i}}_{=0}}{\sum X_{0i}^2} = \sum \left(\frac{1}{N} - \bar{X} \Omega_i \right) Y_i$$

une réalisation de A_0 *est* :

$$\hat{\alpha}_0 = \sum \left(\frac{1}{N} - \bar{x} \omega_i \right) y_i$$

On vient de voir que A_0; est une combinaison linéaire des Y_i et qu'il est sans Biais de α_0 reste à montrer qu'il est de variance minimale :

On calcule maintenant la variance de A_0:
(désormais, on travaille avec les x_i fixées et V au lieu de V^{x_i})

$$V(A_0) = \sum \left(\frac{1}{N} - \bar{x} \omega_i \right)^2 V(Y_i) = \sigma^2 \sum \left(\frac{1}{N} - \bar{x} \omega_i \right)^2$$

$$= \sigma^2 \sum \left(\frac{1}{N^2} + \bar{x}^2 \frac{x_{0i}^2}{\left(\sum_{i=1}^{N} x_{0i}^2 \right)^2} - 2 \frac{\bar{x}}{N} \omega_i \right)$$

$$= \sigma^2 \left(\frac{1}{N} + \bar{x}^2 \frac{1}{\left(\sum_{i=1}^{N} x_{0i}^2 \right)} - 2 \frac{\bar{x}}{N} \frac{\overbrace{\sum x_{0i}}^{=0}}{\sum_{i=1}^{N} x_{0i}^2} \right) = \sigma^2 \left(\frac{1}{N} + \frac{\bar{x}^2}{\left(\sum_{i=1}^{N} x_{0i}^2 \right)} \right)$$

- Considérons un autre estimateur linéaire sans biais A' de α_0, d'où

$$A' = \sum k_i Y_i \quad et \quad E(A') = \alpha_0 \Rightarrow \sum k_i E(Y_i) = \alpha_0$$

$$\Rightarrow \sum k_i (\alpha_0 + \alpha_1 x_i) = \alpha_0$$

$$\Rightarrow \alpha_0 \sum k_i + \alpha_1 \sum k_i x_i = \alpha_0$$

$$\Rightarrow \sum k_i = 1 \quad et \quad \sum k_i x_i = 0$$

- Or posons $f_i = \frac{1}{N} - \bar{x}\,\omega_i$

$$V(A') = \sum k_i^2 V(Y_i) = \sigma^2 \sum k_i^2$$
$$= \sigma^2 \sum (k_i - f_i + f_i)^2 = \sigma^2 \sum (k_i - f_i)^2 + \sigma^2 \sum (f_i)^2$$
$$+ 2\sigma^2 \sum f_i (k_i - f_i)$$
$$= \sigma^2 \sum (k_i - f_i)^2 - V(A_0) + 2\sigma^2 \sum f_i k_i$$
$$= \sigma^2 \sum (k_i - f_i)^2 - V(A_0) + \frac{2\sigma^2}{N} \underbrace{\sum k_i}_{=1} - \frac{2\sigma^2}{\sum x_{0i}^2}\bar{x} \underbrace{\sum x_i k_i}_{=0}$$
$$+ \frac{2\sigma^2 \bar{x}^2}{\sum x_{0i}^2} = \sigma^2 \sum (k_i - f_i)^2 - V(A_0) + 2V(A_0)$$
$$= \underbrace{\sigma^2 \sum (k_i - f_i)^2}_{\geq 0} + V(A_0) \geq V(A_0)$$

- **Théorème 3 :**

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-2}$$ est une estimation sans Biais de σ^2

Dém : en injectant $\bar{Y}$ dans la différence, on a :

$$\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} (Y_i - \bar{Y} + \bar{Y} - \hat{Y}_i)^2;$$

et puisque $\bar{Y} = \bar{\hat{Y}}$ d'où :

$$= \underbrace{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}_{(1)} + \underbrace{\sum_{i=1}^{n} (\hat{Y}_i - \bar{\hat{Y}})^2}_{(2)} - 2\underbrace{\sum_{i=1}^{n} (Y_i - \bar{Y})(\hat{Y}_i - \bar{\hat{Y}})}_{(3)}$$

Dém : (suite)

$$\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} (Y_i - \bar{Y} + \bar{Y} - \hat{Y}_i)^2$$
$$= \underbrace{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}_{(1)} + \underbrace{\sum_{i=1}^{n} (\hat{Y}_i - \bar{\hat{Y}})^2}_{(2)} - 2\underbrace{\sum_{i=1}^{n} (Y_i - \bar{Y})(\hat{Y}_i - \bar{\hat{Y}})}_{(3)}$$

D'autre part : pour calculer (1), on a :

$$\begin{matrix} y_i = \alpha_0 + \alpha_1 x_i + w_i \\ \bar{y} = \alpha_0 + \alpha_1 \bar{x} + \bar{w} \end{matrix} \Rightarrow y_i - \bar{y} = (w_i - \bar{w}) + \alpha_1 (x_i - \bar{x})$$

Dém : (suite)

D'où :

$$\sum_{i=1}^{n} (Y_i - \bar{Y})^2 = \sum_{i=1}^{n} (W_i - \bar{W})^2 + \alpha_1^2 \sum_{i=1}^{n} (x_i - \bar{x})^2$$
$$+ 2\sum_{i=1}^{n} (W_i - \bar{W})(x_i - \bar{x})$$
$$\Rightarrow \sum_{i=1}^{n} E\left[(Y_i - \bar{Y})^2\right] = \sum_{i=1}^{n} E\left[(W_i - \bar{W})^2\right] + \alpha_1^2 \sum_{i=1}^{n} (x_{0i})^2$$
$$+ 2\sum_{i=1}^{n} (x_i - \bar{x}) \underbrace{E(W_i - \bar{W})}_{=0}$$
$$= \sum_{i=1}^{n} E\left[W_i^2\right] + n\,E\left[\bar{W}^2\right] - 2\sum_{i=1}^{n} E\left[W_i \bar{W}\right] + \alpha_1^2 \sum_{i=1}^{n} (x_{0i})^2$$

Dém : (suite)

$$= n\sigma^2 + n\frac{\sigma^2}{n} - 2\sum_{i=1}^{n} E\left[W_i \frac{1}{n}\left(W_i + \sum_{j \neq i} W_j\right)\right] + \alpha_1^2 \sum_{i=1}^{n} (x_{0i})^2$$
$$= n\sigma^2 + \sigma^2 - \frac{2}{n}\sum_{i=1}^{n} \left(\underbrace{E\left[W_i^2\right]}_{=\sigma^2} + \underbrace{E\left(\sum_{j \neq i} W_i W_j\right)}_{=0} \right)$$
$$+ \alpha_1^2 \sum_{i=1}^{n} (x_{0i})^2$$
$$= n\sigma^2 + \sigma^2 - 2\sigma^2 + \alpha_1^2 \sum_{i=1}^{n} (x_{0i})^2 = \boxed{(n-1)\sigma^2 + \alpha_1^2 \sum_{i=1}^{n} (x_{0i})^2}$$

Dém : (suite) Pour ce qui est de (2) :

$$\sum_{i=1}^{n} (\hat{Y}_i - \bar{\hat{Y}})^2 = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 = A_1^2 \sum_{i=1}^{n} (x_i - \bar{x})^2$$
$$\Rightarrow E\left(\sum_{i=1}^{n} (\hat{Y}_i - \bar{\hat{Y}})^2\right) = \left(\sum_{i=1}^{n} (x_i - \bar{x})^2\right) E(A_1^2)$$
$$= \left(\sum_{i=1}^{n} (x_{0i})^2\right) \times \left(V(A_1) + \alpha_1^2\right)$$
$$= \left(\sum_{i=1}^{n} (x_{0i})^2\right) \times \left(\frac{\sigma^2}{\sum_{i=1}^{n} (x_{0i})^2} + \alpha_1^2\right) = \boxed{\sigma^2 + \alpha_1^2 \left(\sum_{i=1}^{n} (x_{0i})^2\right)}$$

Dém : (suite) Pour ce qui est de (3) :

$$\sum_{i=1}^{n}(Y_i-\bar{Y})(\hat{Y}_i-\bar{\hat{Y}})=\sum_{i=1}^{n}Y_i(\hat{Y}_i-\bar{\hat{Y}})-\bar{Y}\sum_{i=1}^{n}(\hat{Y}_i-\bar{\hat{Y}})$$

$$=A_1\underbrace{\sum_{i=1}^{n}Y_i(x_i-\bar{x})}_{=\sum_{i=1}^{n}(x_{0i})^2\sum\omega_i Y_i}-A_1\bar{Y}\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})}_{=0}$$

$$\Rightarrow E\left(\sum_{i=1}^{n}(Y_i-\bar{Y})(\hat{Y}_i-\bar{\hat{Y}})\right)=E\left(A_1\left(\sum_{i=1}^{n}(x_{0i})^2\right)A_1\right)$$

$$=\left(\sum_{i=1}^{n}(x_{0i})^2\right)\times\left(V(A_1)+\alpha_1^2\right)=\boxed{\sigma^2+\alpha_1^2\left(\sum_{i=1}^{n}(x_{0i})^2\right)}$$

Dém : (suite)

Pour ce qui est de (1)+(2)-2*(3) :

$$E\left(\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2\right)=(n-2)\sigma^2$$

$$\Rightarrow E\left(\frac{1}{n-2}\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2\right)=\sigma^2$$

CQFD

On pose donc :

$$S_\sigma^2=\frac{1}{n-2}\sum e_i^2 \quad ESB\ de\ \ \sigma^2$$

Distributions des estimateurs $A_0, A_1\ et\ \hat{Y}$

Puisque W_i est normale, on a

$Y_i/X_i=x_i \sim \mathcal{N}(\alpha_1 x_i+\alpha_0\ ,\ \sigma^2)$

Par conséquent $A_0, A_1\ et\ \hat{Y}$ suivent aussi, pour des x_i fixées, des lois Normales :

$$A_1 \sim \mathcal{N}\left(\alpha_1\ ,\ \frac{\sigma^2}{\sum(x_{0i})^2}\right)$$

$$A_0 \sim \mathcal{N}\left(\alpha_0\ ,\ \sigma^2\left(\frac{1}{N}+\frac{\bar{x}^2}{\sum(x_{0i})^2}\right)\right)$$

$$\hat{Y}_i \sim \mathcal{N}\left(\alpha_1 x_i+\alpha_0\ ,\ \sigma^2\left(\frac{1}{N}+\frac{x_{0i}^2}{\sum(x_{0i})^2}\right)\right)$$

- Remarques :
- 1- Les lois des estimateurs vont nous servir pour la construction des intervalles de confiance et pour les Tests.
- 2- On démontre que :

$$\frac{(n-2)S_\sigma^2}{\sigma^2}=\frac{\sum e_i^2}{\sigma^2}\sim\chi^2(n-2)$$

$$est\ indépendante\ de\ A_0, A_1\ et\ \bar{Y}$$

II- Tests sur les coefficients du MRLS

- 1- Test sur la régression
- On veut tester si la régression de Y sur X est significative ?
- Sous l'hypothèse $\alpha_1=0$ le modèle s'écrit :

$$Y_i=\alpha_0+W_i \quad \forall i=1,\ldots,n$$

et Y ne dépend plus alors de X.

i)- Formulation des Hypothèses :

$$\begin{cases} H_0 : \alpha_1 = 0 \\ \# \\ H_1 : \alpha_1 \neq 0 \end{cases}$$

ii)- Statistique utilisée et sa loi :

On Propose alors A_1 estimateur sans biais de $\boldsymbol{\alpha_1}$

$$\frac{A_1}{\sigma / \sqrt{\sum x_{0i}^2}} \rightsquigarrow \mathcal{N}(0\,,1) \text{ sous } H_0$$

- Or σ est inconnu, on doit l'estimer;
- Si α_0 et α_1 étaient connus, la meilleure estimation de σ^2 est :

$$\frac{1}{n}\sum w_i^2 = \frac{1}{n}\sum (y_i - \alpha_0 - \alpha_1 x_i)^2$$

- Sinon, En remplaçant α_0 et α_1 par A_0 et A_1, on obtient une bonne estimation de σ^2 :

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum (y_i - \hat{\alpha}_0 - \hat{\alpha}_1 x_i)^2 = \frac{1}{n-2}\sum e_i^2$$

- Le $n-2$ vient du nombre de paramètres estimés
- D'où, la statistique du test sera, sous H_0

$$T_{A_1} = \frac{A_1}{S_{A_1}} \quad ;où \quad s_{A_1} = \frac{\hat{\sigma}}{\sqrt{\sum (x_{0i})^2}} = \frac{\sqrt{\sum e_i^2}}{\sqrt{n-2}\sqrt{\sum (x_{0i})^2}}$$

qui suit une loi de student à $n-2$ d.d.l.

- Sous H_0, T devient :

$$T_{A_1} = \frac{A_1}{S_{A_1}} \rightsquigarrow t_{n-2}$$

iii)- R.D.

$$\begin{cases} Si\ |t_{A_1}| > t_{n-2;\alpha/2} & \text{on rejette } H_0 \\ Si\ |t_{A_1}| \le t_{n-2;\alpha/2} & \text{on ne rejette pas } H_0 \end{cases}$$

Exercice :

Pour l'exemple du PNB,

➢ Tester si la régression est significative à 5% ?

- Solution :

✓ F.H.
$$\begin{cases} H_0 : \alpha_1 = 0 \\ \# \\ H_1 : \alpha_1 \neq 0 \end{cases}$$

✓ Estimateur et sa loi $T_{A_1} = \dfrac{A_1}{S_{A_1}} \rightsquigarrow$ t(10)

✓ R.D.

$$\begin{cases} Si\ |t_{\hat{\alpha}_1}| > t_{10;0,025} & \text{on rejette } H_0 \\ Si\ |t_{\hat{\alpha}_1}| \le t_{10;0,025} & \text{on ne rejette pas } H_0 \end{cases}$$

- **A.N.**

$$t_{\hat{\alpha}_1} = \frac{\hat{\alpha}_1}{s_{\hat{\alpha}_1}} = \frac{\hat{\alpha}_1\sqrt{n-2}\sqrt{\sum (x_{0i})^2}}{\sqrt{\sum e_i^2}}$$

$$= \frac{0{,}2074\sqrt{10}\sqrt{752}}{\sqrt{19{,}638}} \approx 4{,}06 \quad et \quad t_{10;0,025} = 2{,}23$$

- On a t *calculé* supérieur au t *tabulé* donc on RH_0 la régression est donc significative

- 2- Intervalle de confiance de α_1

- D'après ce qui précède, on a :

$$I_C(\alpha_1) = [\hat{\alpha}_1 - t_{n-2;\alpha/2} s_{\hat{\alpha}_1} ; \hat{\alpha}_1 + t_{n-2;\alpha/2}\, s_{\hat{\alpha}_1}]$$

- A.N.

$$t_{10;0,025} = 2{,}23 \quad \hat{\alpha}_1 = 0{,}2074 \quad s_{\hat{\alpha}_1} = \frac{\sqrt{\frac{19{,}638}{10}}}{\sqrt{752}} = 0{,}05$$

$$I_C(\alpha_1) = [0{,}2074 - 2{,}23 \times 0{,}05 \ ; \ 0{,}2074 + 2{,}23 \times 0{,}05]$$
$$= [0{,}096 \ ; \ 0{,}319]$$

- Remarque :

$$\alpha_1 = 0 \notin I_C(\alpha_1) = [0,096;\ 0,319]$$
$$\Rightarrow RH_0$$

- Interprétation: à 95% de confiance, l'augmentation de la demande moyenne se situe entre 0,1 et 0,32 pour une augmentation de 1 du PNB
- Pour α_0 on a :

$$I_C(\alpha_0) = \left[\hat{\alpha}_0 - t_{n-2;\alpha/2} s_{\hat{\alpha}_0} ; \hat{\alpha}_0 + t_{n-2;\alpha/2} s_{\hat{\alpha}_0}\right]$$
$$= [-9,979 ;\ 4,255] \text{ où } s^2_{\hat{\alpha}_0} = S^2_{\sigma}\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum x_{0i}^2}\right] = 3,194 \wedge 2$$

III- Tests et Tableau d'Analyse de la Variance (*ANOVA*) d'un MRLS

- **1- Équation fondamentale de l' ANOVA:**
- La notion de liaison entre Y et X, signifie q'une variation de x implique celle de Y. La formule de décomposition de la variance permet de connaître la part de variation de Y expliquée par celle de X :

$$\underbrace{\sum\left(Y_i - \bar{Y}\right)^2}_{SCT} = \underbrace{\sum\left(\hat{Y}_i - \bar{Y}\right)^2}_{SCE} + \underbrace{\sum\left(Y_i - \hat{Y}_i\right)^2}_{SCR=\sum e_i^2}$$

en effet :

$$\sum(Y_i - \bar{Y})^2 = \sum(Y_i - \hat{Y}_i + \hat{Y}_i - \bar{Y})^2$$
$$= \sum(Y_i - \hat{Y}_i)^2 + \sum(\hat{Y}_i - \bar{Y})^2 + 2\sum(Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y})$$
$$= \sum(Y_i - \hat{Y}_i)^2 + \sum(\hat{Y}_i - \bar{Y})^2 + 2\sum e_i(\hat{Y}_i - \bar{Y})$$
$$= \sum(Y_i - \hat{Y}_i)^2 + \sum(\hat{Y}_i - \bar{Y})^2 + 2\sum e_i(\hat{Y}_i - \bar{\hat{Y}})$$
$$= \sum(Y_i - \hat{Y}_i)^2 + \sum(\hat{Y}_i - \bar{Y})^2 + 2A_1\sum e_i(x_i - \bar{x})$$
$$= \sum(Y_i - \hat{Y}_i)^2 + \sum(\hat{Y}_i - \bar{Y})^2 + 2A_1\underbrace{\left(\underbrace{\sum e_i x_i}_{=0} - \bar{x}\underbrace{\sum e_i}_{=0}\right)}_{=0}$$
$$= \sum(Y_i - \hat{Y}_i)^2 + \sum(\hat{Y}_i - \bar{Y})^2$$

- La variabilité totale(SCT)est égale à la variabilité expliquée(SCE) augmentée de la variabilité résiduelle(SCR).
- Cette décomposition va nous permettre de décider de la qualité de l'ajustement du modèle.
- Remarque : si la variance expliquée tend vers la variance totale (SCR faible), la qualité de l'ajustement tend à être meilleure.
- Ceci nous donne une idée de tester, d'une autre façon, la signification de la régression: un test équivalent au T-test sur α_1

- La variabilité expliquée SCE n'est autre qu'un estimateur de $\mathbb{E}$(SCE/1) et SCR estimateur de $\mathbb{E}$(SCR/n-2); on a :

$$\mathbb{E}\left(\frac{SCE}{1}\right) = \mathbb{E}\left(A_1^2 \sum x_{0i}^2\right) = \sigma^2 + \alpha_1^2 \sum x_{0i}^2$$
$$\mathbb{E}\left(\frac{SCR}{n-2}\right) = \mathbb{E}\left(\frac{\sum e_i^2}{n-2}\right) = \sigma^2$$

- De là, on peut affirmer que si la régression est significative $(\alpha_1 \neq 0)$ alors: $\boxed{\mathbb{E}(SCE/1) > \mathbb{E}(SCR/n-2)}$
- Par contre, si la régression n'est pas significative $(\alpha_1 = 0)$ alors: $\boxed{\mathbb{E}(SCE/1) = \mathbb{E}(SCR/n-2)}$

- D'où

$$\begin{cases} H_0 : \alpha_1 = 0 \\ \# \\ H_1 : \alpha_1 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0 : E\left(SCE/1\right) = E\left(SCR/n-2\right) \ (rég\ non\ sign.) \\ \# \\ H_1 : E\left(SCE/1\right) > E\left(SCR/n-2\right) \ (rég\ sign.) \end{cases}$$

- La statistique utilisée est SCE/SCR
- Sa loi sous H_0?

On montre que

$$SCE/\sigma^2 \rightsquigarrow \chi^2(1)$$
$$SCR/\sigma^2 \rightsquigarrow \chi^2(n-2)$$

- D'où sous H_0
- $F = \dfrac{SCE/1}{SCR/n-2} \rightsquigarrow \mathcal{F}(1;n-2)$
- **R.D.** $\begin{cases} Si\ F > \mathcal{F}_\alpha(1,n-2) & \text{on rejette } H_0 \\ Si\ F \le \mathcal{F}_\alpha(1,n-2) & \text{on ne rejette pas } H_0 \end{cases}$
- **Remarque** :**on a vu** $\mathcal{F}(1;n-2) = [t(n-2)]^2$
- En effet : $(T_{A_1})^2 = \left(\dfrac{A_1}{S_{A_1}}\right)^2 = SCE / SCR/n-2 = F$

Tableau de l' *ANOVA*

Source de var$^\theta$	Somme des carrés	d.d.l	Carrés moyens	Fisher
x	SCE=$\sum(\hat{Y}_i - \bar{Y})^2$	1	SCE/1	F= SCE/ (SCR/(n-2))
Résidu	SCR= $\sum e_i^2$	n-2	SCR/(n-2)	
Totale	SCT=$\sum(Y_i - \bar{Y})^2$	n-1		

•**Exercice**: on reprend l'exo. Précédent. A l'aide de l'ANOVA, tester si la régression est significative

- **Solution :** sous H_0
- $F = \dfrac{SCE}{SCR/n-2} \rightsquigarrow \mathcal{F}(1;n-2)$
- **R.D.** $\begin{cases} Si\ F > F_{0,05}(1;10) & \text{on rejette } H_0 \\ Si\ F \le F_{0,05}(1;10) & \text{on ne rejette pas } H_0 \end{cases}$

Or $F = \dfrac{SCE}{SCR}\times 10 = \dfrac{\hat{\alpha}_1^2 \sum x_{0i}^2}{\sum e_i^2}\times 10 = \dfrac{0,2074^2 \times 752}{1,9638} = 16,472$

$\mathcal{F}_{0,05}(1;10) = 4,96$ d'où $F > \mathcal{F}_{0,05}(1;10)$ donc on RH_0

la régression est donc significative

Remarque :F $= [T]^2 = (4,06)^2 = 16,484$

- Une troisième façon de tester la régression est de tester le coefficient de corrélation théorique ρ :

✓ F.H. $\begin{cases} H_0 : \rho = 0 \\ \# \\ H_1 : \rho \ne 0 \end{cases}$

- Soit R le coefficient de corrélation simple entre X et Y estimateur de ρ; R^2 est appelé le coefficient de détermination

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

- En effet, $R = \dfrac{\operatorname{cov}(x,Y)}{\sigma_X \sigma_Y} = \dfrac{\sum x_{0i}Y_{0i}}{\sqrt{\sum x_{0i}^2 \sum Y_{0i}^2}}$

$$\Rightarrow R^2 = \frac{\left(\sum x_{0i}Y_{0i}\right)^2}{\sum x_{0i}^2 \sum Y_{0i}^2} = \hat{\alpha}_1^2 \frac{\sum x_{0i}^2}{\sum Y_{0i}^2} = \frac{\hat{\alpha}_1^2 \sum x_{0i}^2}{SCT} = \frac{SCE}{SCT}$$

- Car

$$SCE = \sum\left(\hat{Y}_i - \bar{Y}\right)^2 = \sum\left(\hat{Y}_i - \bar{\hat{Y}}\right)^2 = \sum\left(\hat{\alpha}_1(x_i - \bar{x})\right)^2 = \hat{\alpha}_1^2 \sum x_{0i}^2$$

✓ Or on a vu que :

$$F = \frac{SCE/1}{SCR/n-2} = \frac{R^2 SCT}{SCR/n-2} = \frac{R^2}{\dfrac{SCR/SCT}{n-2}} = \frac{R^2}{(1-R^2)/n-2}$$

✓ Et sous H_0

$$F = T_{\hat{\alpha}_1}^2 \Rightarrow T_{\hat{\alpha}_1} = \frac{R\sqrt{n-2}}{\sqrt{(1-R^2)}} \rightsquigarrow t_{\alpha/2}(n-2)$$

✓ **R.D.** $\begin{cases} Si\ |r| > r_{\alpha/2;n-2} & on\ rejette\ H_0 \\ Si\ |r| \le r_{\alpha/2;n-2} & on\ ne\ rejette\ pas\ H_0 \end{cases}$

$r_{\alpha/2;n-2}$ se trouve sur la table du coefficient de corrélation

- ***A.N. :*** $r^2 = \dfrac{SCE}{SCT} = \dfrac{32,347}{52} = 0,622 \Rightarrow r = 0,789$
- ***Or*** $\mathrm{r} = 0,789 > r_{0,025;10} \approx 0,576$ on rejette H_0

- ***Remarques:***
- - On peut tester H_0 en utilisant le rapport critique $T = \dfrac{R\sqrt{n-2}}{\sqrt{(1-R^2)}}$, on lit alors t dans la table de student à n-2 ddl
- le r-test sur le coefficient de corrélation ρ est équivalent au T-test sur la pente(α_1) équivalent au F-test sur les variances.

IV- PREVISION Sur Un MRLS

- On sait que $\hat{Y}_i$ est un estimateur de la moyenne des (Y_i) donc de $\mathbb{E}$(Y):

$$Y_t = \hat{Y}_t + e_t = \hat{\alpha}_0 + \hat{\alpha}_1 x_t + e_t \quad \forall t = 1,\cdots,n$$

- À l'instant k=n+1, on a :

$$\hat{Y}_k = \hat{\alpha}_0 + \hat{\alpha}_1 x_k = \overline{Y} + \hat{\alpha}_1 (x_k - \overline{x})$$

Intervalle de confiance autour de $\mathbb{E}$(Y)

- **Propriétés de $\hat{Y}_k$:** la loi de $\hat{Y}_k$ est normale car combinaison des Y_i

$$\mathbb{E}(\hat{Y}_k) = \mathbb{E}(\overline{Y}) + (x_k - \overline{x})\mathbb{E}(\hat{\alpha}_1)$$
$$= (\alpha_0 + \alpha_1 \overline{x}) + (x_k - \overline{x})\alpha_1 = \mathbb{E}(Y_k)$$
$$\mathrm{V}(\hat{Y}_k) = \mathrm{V}(\overline{Y}) + (x_k - \overline{x})^2\, \mathrm{V}(\hat{\alpha}_1) =$$
$$\frac{\sigma^2}{n} + \frac{\sigma^2}{\sum x_{0i}^2}(x_k - \overline{x})^2 = \sigma^2\left(\frac{1}{n} + \frac{(x_k - \overline{x})^2}{\sum x_{0i}^2}\right)$$

- **Or σ^2** est inconnue, on l'estime par $\hat{\sigma}^2 = \dfrac{\sum e_i^2}{n-2}$

✓ on a alors :

$$T_{\hat{Y}_k} = \frac{\hat{Y}_k - \mathbb{E}(Y_k)}{\hat{\sigma}\sqrt{\left(\dfrac{1}{n} + \dfrac{(x_k - \overline{x})^2}{\sum (x_i - \overline{x})^2}\right)}} \rightsquigarrow t_{\alpha/2}(n-2)$$

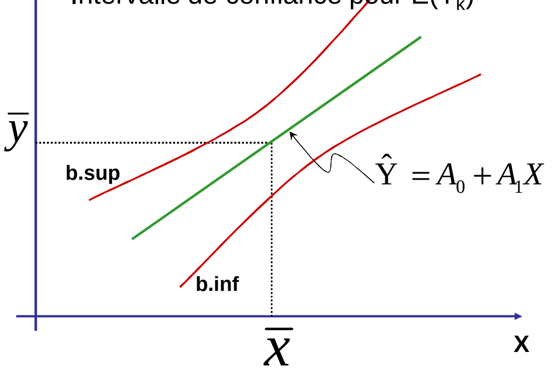
D'où

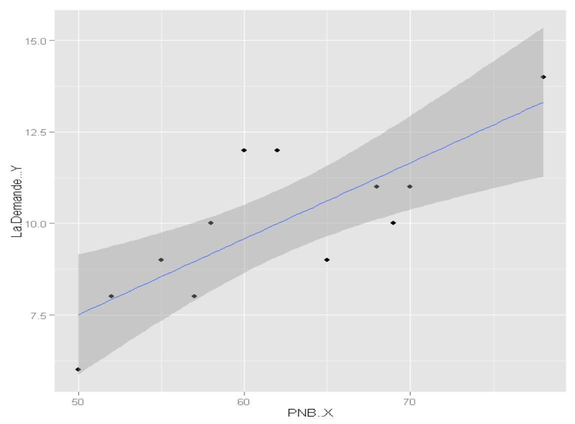
$$I_C(E(Y_k)) = \left[\hat{Y}_k - \hat{\sigma}\, t_{\alpha/2}(n-2)\sqrt{\left(\frac{1}{n} + \frac{(x_k - \overline{x})^2}{\sum (x_i - \overline{x})^2}\right)};\right.$$
$$\left.\hat{Y}_k + \hat{\sigma}\, t_{\alpha/2}(n-2)\sqrt{\left(\frac{1}{n} + \frac{(x_k - \overline{x})^2}{\sum (x_i - \overline{x})^2}\right)}\right]$$

✓ Si $n \longrightarrow \infty$:

$$I_C(E(Y_k)) = \left[\hat{Y}_k - \hat{\sigma}\, z_{\alpha/2}\sqrt{\left(\frac{1}{n} + \frac{(x_k - \overline{x})^2}{\sum (x_i - \overline{x})^2}\right)};\right.$$
$$\left.\hat{Y}_k + \hat{\sigma}\, z_{\alpha/2}\sqrt{\left(\frac{1}{n} + \frac{(x_k - \overline{x})^2}{\sum (x_i - \overline{x})^2}\right)}\right]$$

Intervalle de confiance pour E(Y_k)





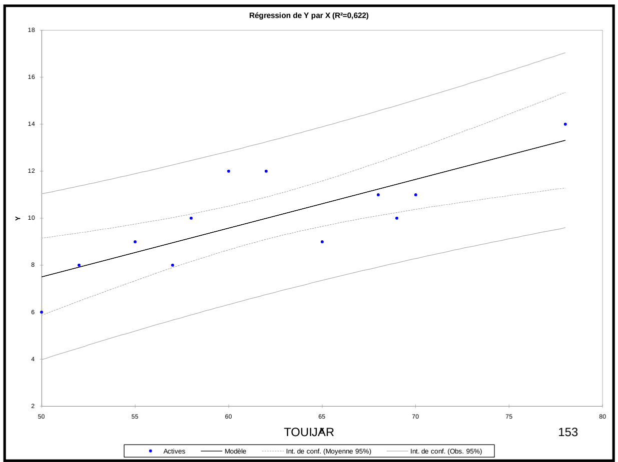
Exercice: on reprend toujours le même exo.

- Estimer à l'aide d'un intervalle de confiance la moyenne des Y pour un $x = 78$.

Solution : on a d'abord $\hat{Y}_{78} = -2,862 + 0,2074 \times 78 = 13,319$

$$I_C\left(E\left(Y_{78}\right)\right) = \left[13,319 - 2,23\sqrt{1,964}\sqrt{\left(\frac{1}{10}+\frac{(78-62)^2}{752}\right)} ; 13,319 + 2,23\sqrt{1,964}\sqrt{\left(\frac{1}{10}+\frac{(78-62)^2}{752}\right)}\right]$$

$$\boxed{= [11,287 ; 15,352]}$$



V- PREVISION Sur Un MRLS

Intervalle de prévision autour de Y_k

Soit x_k une valeur non observée de X (exple $k=n+1$) alors la prévision naturelle de Y_k est

$$\hat{Y}_k = A_0 + A_1 x_k$$

Or $Y_k = (Y/X=x_k) \leadsto \mathcal{N}(\alpha_1 x_k + \alpha_0 , \sigma^2)$

Vu que les 2 V.A. Y_k(ne dépend que de ε_k) et $\hat{Y}_k$ (ne dépend que des valeurs précédentes($\varepsilon_1,\dots,\varepsilon_n$) on a :

IV- PREVISION Sur Un MRLS

Intervalle de prévision autour de $Y_{k(p)}$

$$Y_{k(p)} = \hat{Y}_k + e_k \Rightarrow e_k = Y_{k(p)} - \hat{Y}$$

$$\Rightarrow V(e_k) = V\left(Y_{k(p)}\right) + V\left(\hat{Y}\right) = \sigma^2\left(1+\frac{1}{n}+\frac{(x_k-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right)$$

$$\Rightarrow \hat{\sigma}(e_k) = S_\sigma\sqrt{\left(1+\frac{1}{n}+\frac{(x_k-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right)}$$

$$\Rightarrow \frac{Y_{k(p)} - \hat{Y}}{S_\sigma\sqrt{\left(1+\frac{1}{n}+\frac{(x_k-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right)}} \mapsto t_{\alpha/2}(n-2))$$

IV- PREVISION Sur Un MRLS

Intervalle de prévision autour de $Y_{k(p)}$

$$I_p\left(Y_{k(p)}\right) = \left[\hat{Y}_k - \hat{\sigma}\, t_{\alpha/2}(n-2)\sqrt{\left(1+\frac{1}{n}+\frac{(x_k-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right)} ; \hat{Y}_k + \hat{\sigma}\, t_{\alpha/2}(n-2)\sqrt{\left(1+\frac{1}{n}+\frac{(x_k-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right)}\right]$$

$$I_p\left(Y_{78}\right) = [9,568 ; 17,07]$$

IV- PREVISION Sur Un MRLS

Remarque :

Cet intervalle de prévision permet de détecter les points aberrants(ou extrêmes) afin de les écarter pour refaire une nouvelle régression sans leurs influence.

Chapitre 2

Régression Multiple

Introduction :

- Le but premier de ce deuxième chapitre est la modélisation (l'explication) dans un but prédictif, d'une variable quantitative Y par plusieurs variables quantitatives X_1, X_2, …, X_p . Ces dernières sont liées linéairement avec Y. Il s'agit là de ce qu'on appelle : la régression linéaire multiple.

Le modèle :

- Le modèle de régression linéaire multiple est une généralisation de la régression simple.
- C'est un outil statistique mis en œuvre pour l'étude de données multidimensionnelles.

Le modèle : Objectifs

- Estimer les paramètres du modèle $\alpha_0; \alpha_1; \alpha_2; \cdots; \alpha_p$

Avec des estimateurs de meilleur qualité.

- Mesurer le pouvoir explicatif global du modèle.
- Faire de la prévision en construisant des intervalles de prévision.
- Ce dernier point nous permettra de repérer les points aberrants et de les supprimer.

Le modèle : (aspect empirique)

- Une variable quantitative Y (V. à expliquer ou endogène) est mise en relation avec p variables quantitatives X_1, X_2, …, X_p (V. explicatives , exogènes ou régresseurs).
- On mesure sur n individus ces p+1 variables représentées par des vecteurs de R^n: y, x_1, x_2, …, x_p (où $n > p+1$).
- L'écriture du modèle linéaire est alors comme suit :

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \cdots + \alpha_p x_{ip} + w_i \quad 1 \le i \le n$$

Le modèle : Ecriture matricielle

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix}}_{Y\ (n,1)} = \underbrace{\begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{i1} & x_{i2} & \cdots & x_{ip} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}}_{X\ (n,p+1)} \underbrace{\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_p \end{pmatrix}}_{\alpha(p+1,1)} + \underbrace{\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{pmatrix}}_{W\ (n,1)}$$

$$Y = X\alpha + W$$

Le modèle : Les Hypothèses

On supposera vrai les hypothèses suivantes :

H1- Linéarité : $y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \cdots + \alpha_p x_{ip} + w_i \quad 1 \le i \le n$
La relation entre y et $x_1, \ldots, x_p$ est linéaire.

H2: Plein rang : La matrice X'X est inversible; autrement det(X'X)≠0. On peut l'exprimer par le fait que les X_i sont indépendantes linéairement (pas statistiquement). Cette hyp. est nécessaire pour l'estimation des paramètres.

H3: Exogénéité des variables indépendantes : Les W_i sont des termes d'erreur d'espérance conditionnelle aux réalisations des x_i est nulle : $E(W_i \mid x_1, \ldots, x_p) = 0$.
Les x_i n'interviennent pas dans la prédiction de W_i

Le modèle : Les Hypothèses (suite)

H4: Homoscédasticité et absence d'autocorrélation: $V(W_i) = \boldsymbol{\sigma^2}$; où $\boldsymbol{\sigma^2}$ est cste et w_i n'est pas corrélé avec w_j pour $i \neq j$: $\mathrm{cov}(w_i, w_j) = 0$

H5: Génération des données: Les X_i qu'elle soient aléatoires ou déterministes (facteurs contrôlés) ne changent en rien les résultats.

H6: Distribution Normale : Les W sont distribués selon la loi Normale.

Le modèle : Méthode des Moindres Carrés

- Comme on a vu dans le chapitre 1, il s'agit , afin d'estimer $\boldsymbol{\alpha}$, de minimiser la somme des carrées des résidus (e_i) (voir graphique du chapitre 1) :

$$\min \sum_{i=1}^{n} e_i^2$$

$$\text{où} \quad e = Y - \hat{Y} = Y - X\hat{\alpha}$$

Le modèle : Méthode des Moindres Carrés

- Autrement :

$$F(\alpha) = \sum_{i=1}^{n} e_i^2 = e'e$$

$$= \sum \left(y_i - \alpha_0 - \alpha_1 x_{i1} - \alpha_2 x_{i2} + \cdots - \alpha_p x_{ip} \right)^2$$

$$= (Y - X\alpha)'(Y - X\alpha) = Y'Y - 2\alpha' X'Y + \alpha' X'X\alpha$$

$$\frac{\partial F}{\partial \alpha} = 0 \Leftrightarrow -2X'Y + 2X'X\alpha = 0$$

$$\Leftrightarrow \boxed{\alpha = (X'X)^{-1} X'Y}$$

Le modèle : Méthode des Moindres Carrés

- Où:

$$X'X = \underbrace{\begin{pmatrix} n & \sum x_{i1} & \sum x_{i2} & \cdots & \sum x_{ip} \\ \sum x_{i1} & \sum x_{i1}^2 & \sum x_{i1}x_{i2} & \cdots & \sum x_{i1}x_{ip} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum x_{ip} & \sum x_{ip}x_{i1} & & \cdots & \sum x_{ip}^2 \end{pmatrix}}_{(p+1,p+1)}$$

$$X'Y = \begin{pmatrix} \sum y_i \\ \sum x_{i1} y_i \\ \vdots \\ \sum x_{ip} y_i \end{pmatrix}$$

Le modèle : Méthode des Moindres Carrés

- Exemple : (Modèle avec constante)

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad Y = \begin{pmatrix} 1.5 \\ 2 \end{pmatrix} \Rightarrow 2 \textit{observations} : (1, 1.5) \textit{ et } (2,2)$$

$$X'Y = \begin{pmatrix} 3.5 \\ 5.5 \end{pmatrix} \quad X'X = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \Rightarrow (X'X)^{-1} = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}$$

$$\hat{\alpha} = (X'X)^{-1} X'Y = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix} \qquad \underbrace{\hat{Y} = X\hat{\alpha} = \begin{pmatrix} 1.5 \\ 2 \end{pmatrix}}_{\hat{y}_i = 1 + 0.5x_i}$$

$$e = \begin{pmatrix} 1.5 \\ 2 \end{pmatrix} - \begin{pmatrix} 1.5 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- $e = 0_{R^2}$: signifie que la droite estimée passe par les deux points (1 , 1.5) et (2 , 2).

Le modèle : Méthode des Moindres Carrés

- Exemple : (Modèle sans constante)

$$X = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad Y = \begin{pmatrix} 1.5 \\ 2 \end{pmatrix} \Rightarrow 2 \textit{observations} : (1, 1.5) \textit{ et } (2,2)$$

$$X'Y = 5.5 \quad X'X = 5 \Rightarrow (X'X)^{-1} = \frac{1}{5}$$

$$\hat{\alpha} = (X'X)^{-1} X'Y = \frac{5.5}{5} = 1.1 \quad \underbrace{\hat{Y} = X\hat{\alpha}}_{\hat{y}_i = 1.1x_i} = \begin{pmatrix} 1.1 \\ 2.2 \end{pmatrix}$$

$$e = \begin{pmatrix} 1.5 \\ 2 \end{pmatrix} - \begin{pmatrix} 1.1 \\ 2.2 \end{pmatrix} = \begin{pmatrix} 0.4 \\ -0.2 \end{pmatrix}$$

- On remarque que $\sum_{i=1}^{n} e_i \neq 0$ car le modèle est sans cste.

PROPRIETES ET DISTRIBUTION DE L'ESTIMATEUR

- Théorème (GAUSS-MARKOV) :

  $\hat{\alpha}$ est BLUE de α

- Remarque :

➢ $\hat{\alpha}_0 ; \hat{\alpha}_1 ; \hat{\alpha}_2 ; \cdots ; \hat{\alpha}_p$ sont des fonctions linéaires des Y_i

➢ La matrice Var-Cov (variance-covariance) de $\hat{\alpha}$, *notée* $\Omega_{\hat{\alpha}}$, s'écrit : $\boxed{\Omega_{\hat{\alpha}} = \sigma^2 (X'X)^{-1}}$

- Quelques éléments de démonstration :

✓ Puisque, $\hat{\alpha} = (X'X)^{-1} X'Y$ et en posant $k = (X'X)^{-1} X'$; on obtient l'écriture : $\hat{\alpha} = kY$ combinaison des Y_i

✓ $E(\hat{\alpha}) = kE(Y) = k(X\alpha) = \underbrace{(X'X)^{-1} X'X}_{I} \alpha = \alpha$

✓ $\Omega_{\hat{\alpha}} = kV(Y)k' = k(\sigma^2 I_n)k' = \sigma^2 kk'$

$= \sigma^2 (X'X)^{-1} X'\left[X(X'X)^{-1}\right]$

$= \sigma^2 (X'X)^{-1} [X'X] (X'X)^{-1} = \sigma^2 (X'X)^{-1}$

✓ Rem : *on peut utiliser la définition :*

$$\Omega_{\hat{\alpha}} = E\left[(\alpha - \hat{\alpha})(\alpha - \hat{\alpha})'\right]$$

PROPRIETES ET DISTRIBUTION DE L'ESTIMATEUR

- Matrice de Var-Cov de W et son estimateur

$$\Omega_W = \Omega = E\left[WW'\right] = \begin{pmatrix} E(W_1^2) & E(W_1W_2) & \cdots & E(W_1W_n) \\ E(W_2W_1) & E(W_2^2) & & \\ \vdots & & \ddots & \\ E(W_nW_1) & & & E(W_n^2) \end{pmatrix}$$

$$= \begin{pmatrix} E(W_1^2) & 0 & \cdots & 0 \\ 0 & E(W_2^2) & & \\ \vdots & & \ddots & \\ 0 & & & E(W_n^2) \end{pmatrix} = \sigma^2 I_n$$

PROPRIETES ET DISTRIBUTION DE L'ESTIMATEUR

- Matrice de Var-Cov de W et son estimateur

$$e = Y - \hat{Y} = (X\alpha + W) - (X\hat{\alpha})$$
$$= X\alpha + W - X\left(\alpha + (X'X)^{-1}X'W\right)$$
$$= \left(\underbrace{I - X(X'X)^{-1}X'}_{\Gamma}\right)W \quad (\text{où } \Gamma \text{ est symétrique et idempotente})$$
$$\Rightarrow \sum e_i^2 = e'e = W'\Gamma W \Rightarrow E(e'e) = \sigma^2 Tr(\Gamma) \quad (\text{à développer})$$
$$= \sigma^2\left(n-(p+1)\right) \Rightarrow \sigma^2 = \frac{E(e'e)}{n-p-1} \Rightarrow \hat{\sigma}^2 = \frac{\sum e_i^2}{n-p-1}$$
$$= \frac{SCR}{n-p-1} \text{ est donc un estimateur sans Biais de } \sigma^2$$

PROPRIETES ET DISTRIBUTION DE L'ESTIMATEUR

- Estimation de la Matrice de Var-Cov de $\hat{\alpha}$:
- On a déjà vu que :

$$\Omega_{\hat{\alpha}} = \sigma^2 (X'X)^{-1} \Rightarrow \hat{\Omega}_{\hat{\alpha}} = \hat{\sigma}^2 (X'X)^{-1}$$

Tableau d'Analyse de la Variance (*ANOVA*) d'un MRLM

- **Équation fondamentale de l' ANOVA:**
  - La formule de décomposition de la variance permet de connaître la part de variation de Y expliquée par celle des X_i :

$$\underbrace{\sum (Y_i - \bar{Y})^2}_{SCT} = \underbrace{\sum (\hat{Y}_i - \bar{Y})^2}_{SCE} + \underbrace{\sum (Y_i - \hat{Y}_i)^2}_{SCR = \sum e_i^2}$$

Tableau de l' *ANOVA*

Source de var°	Somme des carrés	d.d.l	Carrés moyens	Fisher
x	SCE= $\sum (\hat{Y}_i - \bar{Y})^2$	p	SCE/p	F= (SCE/ p)/ (SCR/(n-p-1))
Résidu	SCR= $\sum e_i^2$	n-p-1	SCR/(n-p-1)	
Totale	SCT= $\sum (Y_i - \bar{Y})^2$	n-1		

- Mesure de la Qualité de l'ajustement
- L'évaluation globale de la régression est donnée par R^2 le coefficient de détermination, qui exprime la part de variabilité totale expliquée par le modèle:

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

- Remarque:
- R^2 doit être utilisé avec précaution.
- On ne peut utiliser R^2 dans un modèle sans constante.
- Si p augmente, R^2 augmente aussi, même s'il y a des variables qui n'ont rien à voir avec le phénomène; pour ce on corrige R^2 :

$$\bar{R}^2 = 1 - \frac{(n-1)}{(n-p-1)}\left(1-R^2\right) = 1 - \frac{SCR/_{n-p-1}}{SCT/_{n-1}} < R^2$$

Tests et Intervalles de Confiance des Coefficients du Modèle

- Test de Significativité Globale de la Régression
- On vient d'évaluer la qualité d'un ajustement par le calcul de R^2 (où R est appelé coefficient de corrélation multiple), mais sur un échantillon; pour que ça soit sur toute la population; on doit procéder à une inférence statistique sur le ρ^2 paramètre théorique:
- C'est un F-Test :

$$\begin{cases} H_0 : \rho^2 = 0 \\ \# \\ H_1 : \rho^2 \neq 0 \end{cases}$$

- 1- Test de Significativité Globale de la Régression
- D'une façon analogue au MRLS, a statistique du test est , sous H_0

$$F = \frac{SCE/p}{SCR/(n-p-1)} = \frac{R^2/p}{(1-R^2)/(n-p-1)} \rightsquigarrow \mathcal{F}\,(p\,;n-p-1)$$

- **R.D.** si $F>F_\alpha(p;n-p-1)$, On Rejette H_0

- 1- Test de Significativité Globale de la Régression
- Le F-Test précédent est équivalent au F-Test sur le vecteur coefficient $\boldsymbol{\alpha}$:
- F.H.

$$\begin{cases} H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0 \\ \# & \forall\, \alpha_0 \\ H_1 : \exists\ j\ tq\ \ \alpha_j \neq 0 \end{cases}$$

2- Test de Significativité individuel des coefficients

Est-ce que la Variable X_i joue significativement dans l'explication de Y ? On effectue alors un T-test

F.H.

$$\begin{cases} H_0 : \alpha_i = 0 \\ \# \\ H_1 : \alpha_i \neq 0 \end{cases}$$

- S.U.
- $$T_{\hat\alpha_i} = \frac{\hat\alpha_i}{\hat\sigma_{\hat\alpha_i}} \rightsquigarrow \text{t(n-p-1)}$$

- Calcul de $\hat\sigma_{\hat\alpha_i}$:

✓ On a vu que

$$\hat\Omega_{\hat\alpha} = \begin{pmatrix} \hat\sigma^2_{\hat\alpha_0} & & & \\ & \hat\sigma^2_{\hat\alpha_1} & & \\ & & \ddots & \\ & & & \hat\sigma^2_{\hat\alpha_p} \end{pmatrix} = \hat\sigma^2\,(X'X)^{-1}$$

$$= \frac{\sum e_i^2}{n-p-1}(X'X)^{-1}$$

si on pose d_{ii} les élément diagonaux de $(X'X)^{-1}$

alors : $\hat\sigma^2_{\hat\alpha_i} = \dfrac{\sum e_i^2}{n-p-1} \times d_{ii}$

Tests et Intervalles de Confiance des Coefficients du Modèle

- Test Modèle réduit VS Modèle Complet

VOIR TD

PREVISION ET INTERVALLE DE PREVISION

- Si on ajoute une observation k =n+1 pour chacune des variables explicatives, on obtient une prévision ponctuelle :

$$\hat y_k = \hat\alpha_0 + \hat\alpha_1 x_{k1} + \hat\alpha_2 x_{k2} + \cdots + \hat\alpha_p x_{kp}$$
$$= X_k\hat\alpha \ ;\ où\quad X_k = \begin{pmatrix}1 & x_{k1} & \cdots & x_{kp}\end{pmatrix}$$

- Et on montre que :

$$\hat\sigma^2_{e_k} = \hat\sigma^2\left[1 + X_k\,(X'X)^{-1} X_k'\right]$$

- $$\Rightarrow \frac{e_k}{\hat\sigma_{e_k}} = \frac{\hat y_k - y_k}{\hat\sigma_{e_k}} \rightsquigarrow \text{t(n-p-1)}$$

- INTERVALLE DE PREVISION de y_k

$$I_p\left(y_k\right)=\left[\hat{y}_k - t_{n-p-1;\alpha/2}\hat{\sigma}_{e_k} ; \hat{y}_k + t_{n-p-1;\alpha/2}\hat{\sigma}_{e_k}\right]$$

TOUIJAR 187