2.6.5) Test de la validité "globale" de la régression:

Une régression linéaire simple ($Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$) est valide, ou significative si $\beta_1 \neq 0$

Le tableau ANOVA obtenu en (2.5.3) peut être utilisé pour tester:

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0$$

avec la statistique de Fisher

$$F = \frac{MSR}{MSE} = \frac{SSR/1}{SSE/(n-2)}$$

Sous H_0, on a que $F \sim F(1, n-2)$

On rejette donc H_0 au niveau $100 \times (1-\alpha)\%$ si

$$\boxed{F > F_\alpha(1, n-2)}$$

<u>Remarque</u>: En régression linéaire simple SEULEMENT (= pas vrai en reg. multiple), le test F est équivalent au test t pour $\beta_1 = 0$:

$$F = \frac{SSR/1}{SSE/(n-2)} = \frac{SSR}{\hat{\sigma}^2} = \frac{SSR}{s^2} = \frac{\sum_{t=1}^{n}(\hat{Y}_t - \bar{Y})^2}{s^2}$$

$$= \frac{\sum_{t=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 X_t - \hat{\beta}_0 - \hat{\beta}_1 \bar{X})^2}{s^2} = \frac{\hat{\beta}_1^2 \times \sum_{t=1}^{n}(X_t - \bar{X})^2}{s^2}$$

$$\Rightarrow F = \frac{\hat{\beta}_1^2}{\frac{S^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}}$$

$$= \frac{(\hat{\beta}_1 - 0)^2}{\hat{Var}(\hat{\beta}_1)}$$

$$= t^2 .$$

<u>exemple:</u> Soit le tableau ANOVA suivant:

Source	SS	d.l	MS	F
Régression	48.845	1	48.845	9.249
Erreur	63.374	12	5.281	
Total	112.219	13		

Vérifier la validité de la régression à l'aide du test F

$\Rightarrow$ On a que $F = 9.249$

Or $F_{0.05}(1,12) = 4.75$

Puisque $F > F_{0.05}(1,12)$; on rejette H_0 ; la régression est significative au niveau de confiance 95%.

2.7) Prévisions et intervalles de confiance :

On peut utilisé la droite de régression pour faire de types de prévision de Y^* sachant X^*:

Type 1: Prévision pour la valeur moyenne (l'espérance de Y^*):

$$\mathbb{E}(Y^*) = \beta_0 + \beta_1 X^*$$

Type 2: Prévision pour la vraie valeur (réalisation de Y^*):

$$Y^* = \beta_0 + \beta_1 X^* + \varepsilon$$

Remarques:

1) Dans les deux types, la prévision est le point sur la droite de régression

$$\hat{\mathbb{E}}(Y^*) = \hat{Y}^* = \hat{\beta}_0 + \hat{\beta}_1 X^*$$

2) La prévision est sans biais:

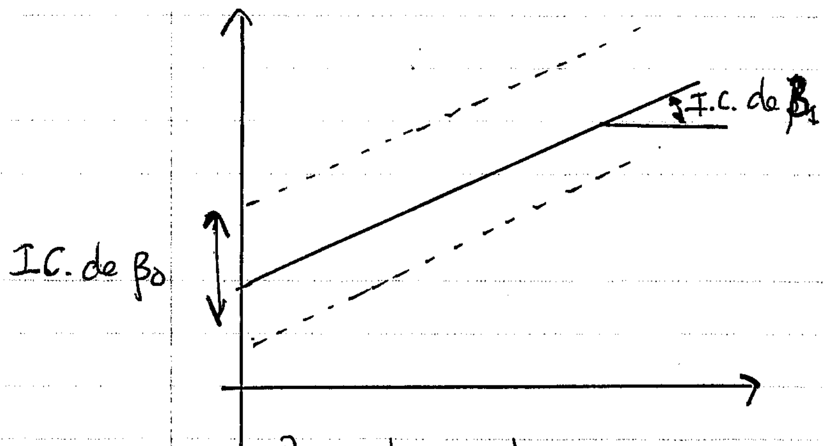
$$\mathbb{E}(\hat{\beta}_0 + \hat{\beta}_1 X^*) = \mathbb{E}(\hat{\beta}_0) + \mathbb{E}(\hat{\beta}_1) X^* = \beta_0 + \beta_1 X^*$$

3) Il y a 2 sources d'erreur dans les prévisions:

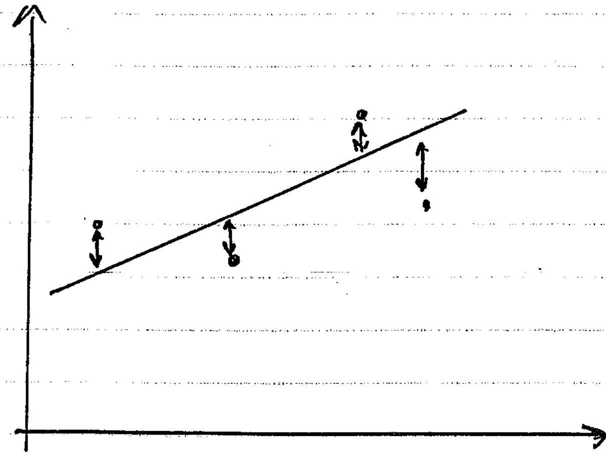
I) Parameter risk: Incertitude sur les estimateurs $(Var(\hat{\beta}_0), Var(\hat{\beta}_1))$

II) Process risk: Fluctuations autour de la droite de rég. $(Var(\varepsilon_i))$

Graphiquement:

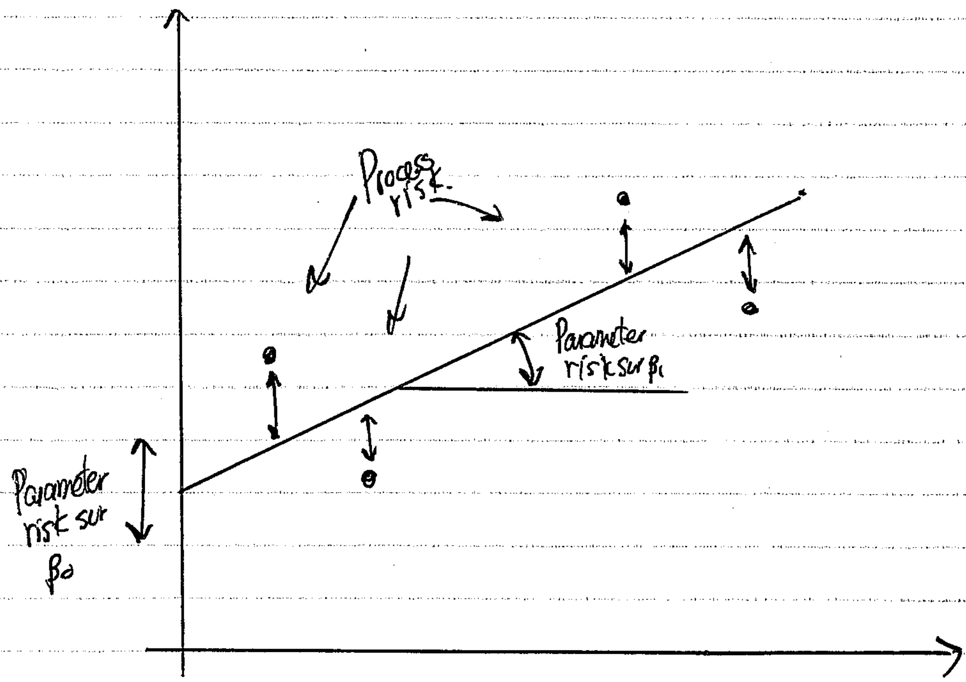


Parameter risk



Process risk

Combinaison des 2 effets :



2.7.1) IC. pour la prévision de type 1 (valeur moyenne):

* Aussi appelé IC. pour la droite de régression

On a que: (sous l'hyp. de la section (2.6.1))

$$\left(\hat{\mathbb{E}}(Y^*) = \hat{\beta}_0 + \hat{\beta}_1 X^*\right) \sim N\left(\beta_0 + \beta_1 X^* \,;\, Var(\hat{\beta}_0 + \hat{\beta}_1 X^*)\right)$$

Par conséquent:

$$\frac{(\hat{\beta}_0 + \hat{\beta}_1 X^*) - (\beta_0 + \beta_1 X^*)}{\sqrt{Var(\hat{\beta}_0 + \hat{\beta}_1 X^*)}} \sim N(0,1)$$

En remplacant σ^2 par S^2 dans $Var(\hat{\beta}_0 + \hat{\beta}_1 X^*)$; on a:

$$\frac{(\hat{\beta}_0 + \hat{\beta}_1 X^*) - (\beta_0 + \beta_1 X^*)}{\sqrt{\widehat{Var}(\hat{\beta}_0 + \hat{\beta}_1 X^*)}} \sim t(n-2).$$

Ainsi, un I.C. au niveau $100 \times (1-\alpha)\%$ pour la valeur moyenne est:

$$\boxed{(\hat{\beta}_0 + \hat{\beta}_1 X^*) \pm t_{\frac{\alpha}{2}}(n-2) \times \sqrt{\widehat{Var}(\hat{\beta}_0 + \hat{\beta}_1 X^*)}}$$

Or, $\quad Var(\hat{\beta}_0+\hat{\beta}_1 X^*) = Var\left(\bar{Y}-\bar{Y}+\hat{\beta}_0+\hat{\beta}_1 X^*\right)$

$$= Var\left(\bar{Y}-(\hat{\beta}_0+\hat{\beta}_1\bar{X})+\hat{\beta}_0+\hat{\beta}_1 X^*\right)$$

$$= Var\left(\bar{Y}+\hat{\beta}_1(X^*-\bar{X})\right)$$

$$= Var(\bar{Y})+(X^*-\bar{X})^2 Var(\hat{\beta}_1)$$

$$= \frac{\sigma^2}{n} + (X^*-\bar{X})^2 \times \frac{\sigma^2}{\sum_{t=1}^{n}(X_t-\bar{X})^2}$$

$$= \sigma^2\left(\frac{1}{n}+\frac{(X^*-\bar{X})^2}{\sum_{t=1}^{n}(X_t-\bar{X})^2}\right)$$

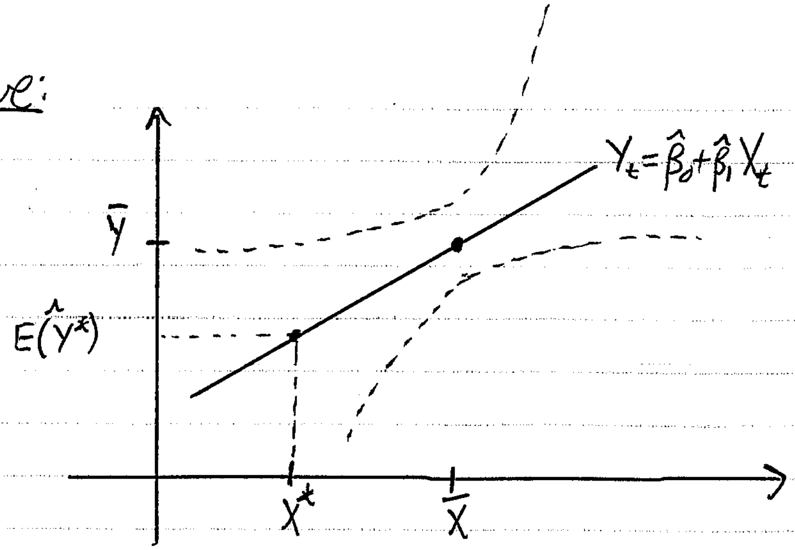
D'où:

$$\boxed{\widehat{Var}(\hat{\beta}_0+\hat{\beta}_1 X^*) = S^2\left(\frac{1}{n}+\frac{(X^*-\bar{X})^2}{\sum_{t=1}^{n}(X_t-\bar{X})^2}\right)}$$

$\Rightarrow$ L'I.C est donc:

$$\boxed{(\hat{\beta}_0+\hat{\beta}_1 X^*) \pm t_{\frac{\alpha}{2}}(n-2) \times \sqrt{S^2 \times \left(\frac{1}{n}+\frac{(X^*-\bar{X})^2}{\sum_{t=1}^{n}(X_t-\bar{X})^2}\right)}}$$

Remarque:



1) Plus X^* s'éloigne de $\bar{X}$, plus l'I.C est large (parce que l'incertitude augmente)

2) Les limites de l'intervalle sont des "hyperboles" centrées en $(\bar{X}, \bar{Y})$.

3) Cet I.C peut être appelé:
- I.C. pour la valeur moyenne
- I.C pour la droite de régression.
- I.C pour la tendance.

(*) 4) Dans ce type de I.C., on tient seulement compte du parameter risk

2.7.2) I.C. pour la prévision de type 2 (vraie valeur):

* Aussi appelé I.C. pour les points ... les réalisations de Y^*

Pour obtenir un I.C pour la vraie valeur de Y^*; il faut tenir compte du ~~parameter risk~~ (...$Var(\hat{\beta}_i)$) ET du process risk (...$var(\varepsilon_t)$). / Process risk

On considère donc: (de manière équivalente à 2.7.1)

$$\frac{Y^* - \hat{Y}^*}{\sqrt{Var(Y^* - \hat{Y}^*)}} \sim N(0,1)$$

En remplaçant σ^2 par S^2 dans $Var(Y^* - \hat{Y}^*)$; on a:

$$\frac{Y^* - \hat{Y}^*}{\sqrt{\widehat{Var}(Y^* - \hat{Y}^*)}} \sim t(n-2)$$

Ainsi; un I.C au niveau $100 \times (1-\alpha)\%$ pour la vraie valeur de Y^* est:

$$\hat{Y}^* \pm t_{\frac{\alpha}{2}}(n-2) \times \sqrt{\widehat{Var}(Y^* - \hat{Y}^*)}$$

Or, $$Var(Y^* - \hat{Y}^*) = Var(Y^*) + Var(\hat{Y}^*) \quad \text{...par hypothèse}$$

$$= \underbrace{\sigma^2}_{\text{Process risk.}} + \underbrace{\sigma^2 \left(\frac{1}{n} + \frac{(X^* - \bar{X})^2}{\sum_{t=1}^{n} (X_t - \bar{X})^2} \right)}_{\text{Parameter risk}}$$

D'où : $$\boxed{\hat{Var}(Y^* - \hat{Y}^*) = S^2\left(1 + \frac{1}{n} + \frac{(X^* - \bar{X})^2}{\sum_{t=1}^{n}(X_t - \bar{X})^2}\right)}$$

L'I.C. est donc:

$$\boxed{(\hat{\beta}_0 + \hat{\beta}_1 X^*) \pm t_{\frac{\alpha}{2}}(n-2) \times \sqrt{S^2 \times \left(1 + \frac{1}{n} + \frac{(X^* - \bar{X})^2}{\sum_{t=1}^{n}(X_t - \bar{X})^2}\right)}}$$

<u>exemple (Logiciel R):</u>

Il est possible d'obtenir le résultats des formules de (2.7.1) et (2.7.2) directement en R.

```
fit ← lm(Y~x)
```

<u>Pour obtenir l'I.C de type 1:</u>

```
predict(fit, interval="confidence")
```
: I.C pour tous les X dans les observations

```
predict(fit, interval="confidence", newdata=x_star)
```
: I.C pour un vecteur $X^* = x\_star$

<u>Pour obtenir l'I.C. de type 2:</u>

```
predict(fit, interval="prediction", newdata=x_star)
```
: I.C. pour un vecteur $X^* = x\_star$

Chapitre 3 : Régression multiple :

Il n'est pas rare que plus d'une variable soit nécessaire pour expliquer un phénomène

ex1:
$$\underbrace{Y}_{\text{Revenu (salaire)}} = \beta_0 + \beta_1 \underbrace{X_1}_{\text{Scolarité}} + \beta_2 \underbrace{X_2}_{\text{Expérience}} + \varepsilon$$

De manière générale, la régression multiple considère le modèle :

$$Y_t = \beta_0 + \beta_1 \times X_{t,1} + \beta_2 \times X_{t,2} + \dots + \beta_p \times X_{t,p} + \varepsilon_t$$

; par $t = 1, \dots, n$.

- n observations
- p variables exogènes $(X_1, \dots, X_p)$
- $(p+1)$ paramètres à estimer $(\beta_0, \beta_1, \dots, \beta_p)$

3.0) Quelques éléments d'algèbre matricielle - Vecteurs et matrices aléatoires :

Soient $X_1, \dots, X_n$ des v.a. On définit le vecteur aléatoire X suivant

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}_{n \times 1}$$

On définit le vecteur espérance de la façon suivante:

$$\mathbb{E}(X) = \begin{bmatrix} \mathbb{E}(X_1) \\ \mathbb{E}(X_2) \\ \vdots \\ \mathbb{E}(X_n) \end{bmatrix}_{n \times 1}$$

et la matrice de variance-covariance

$$\mathrm{Var}(X) = \underbrace{\mathbb{E}\left((X - \mathbb{E}(X))(X - \mathbb{E}(X))^{\overset{\text{transposée}}{'}} \right)}_{\text{Produit matriciel!}} = \begin{bmatrix} \mathrm{Var}(X_1) & \cdots & \mathrm{Cov}(X_1, X_n) \\ \vdots & & \\ \mathrm{Cov}(X_n, X_1) & \cdots & \mathrm{Var}(X_n) \end{bmatrix}_{n \times n}$$

Théorème: Soit X, un vecteur aléatoire et A une matrice de constantes tels que:

$$X = X_{n \times 1} \quad \text{et} \quad A = A_{p \times n}$$

alors

$$\mathbb{E}(AX) = A\,\mathbb{E}(X)$$

et

$$\mathrm{Var}(AX) = A\,\mathrm{Var}(X)\,A'$$

<u>exemple:</u>

- $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}'_{2\times1} = [1, 1]_{1\times2}$

- $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}_{2\times1}$

<u>Intuitivement:</u>

$\Rightarrow AX = X_1 + X_2 \Rightarrow$ • $E(AX) = E(X_1+X_2) = E(X_1) + E(X_2)$

• $Var(AX) = Var(X_1+X_2)$
$= Var(X_1) + Var(X_2) + 2\,Cov(X_1, X_2)$

<u>Matriciellement:</u>

- $E(AX) = A \times E(X) = [1 \quad 1] \begin{bmatrix} E(X_1) \\ E(X_2) \end{bmatrix} = E(X_1) + E(X_2) \Rightarrow$ ok

- $Var(AX) = A\ Var(X)\ A' = [1 \quad 1] \begin{bmatrix} Var(X_1) & Cov(X_1,X_2) \\ Cov(X_1,X_2) & Var(X_2) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$= \left[\left(Var(X_1) + Cov(X_1,X_2)\right), \left(Cov(X_1,X_2) + Var(X_2)\right) \right] \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$= Var(X_1) + Var(X_2) + 2\,Cov(X_1, X_2) \Rightarrow$ ok

Rendu ici
28-09-2010