6 Le modèle linéaire simple à plusieurs variables explicatives

6.1 : Les estimateurs des Moindres Carrés Ordinaires (MCO)

Après avoir posé la spécification matricielle du modèle et les hypothèses de base du modèle nous verrons les estimateurs des MCO ainsi que les propriétés de ces estimateurs.

6.1.1. Spécification matricielle du modèle :

Y : variable expliquée reliée linéairement à k-1 variables exogènes ($X_2 \ldots X_k$) et à l'aléa ε

Les variables sont supposées observées sur un échantillon de taille n. A un instant t de l'échantillon, le modèle s'écrit :

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + \varepsilon_t \qquad t \in \{1,\ldots,n\}$$

$$\underset{(n,1)}{\begin{bmatrix} Y_1 \\ . \\ . \\ Y_t \\ . \\ . \\ Y_n \end{bmatrix}} = \underset{(n,k)}{\begin{bmatrix} 1 \; X_{21} \ldots\ldots\ldots\ldots X_{k1} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ 1 \ldots\ldots X_{jt} \ldots\ldots X_{kt} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ 1 \; X_{2n} \ldots\ldots\ldots\ldots X_{kn} \end{bmatrix}} \underset{(k,1)}{\begin{bmatrix} \beta_1 \\ . \\ . \\ . \\ . \\ . \\ \beta_k \end{bmatrix}} + \underset{(n,1)}{\begin{bmatrix} \varepsilon_1 \\ . \\ . \\ \varepsilon_k \\ . \\ . \\ \varepsilon_n \end{bmatrix}}$$

Soit : $\underset{(n,1)}{Y} = \underset{(n,k)}{X} \; \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$

L'introduction d'une variable X_1 toujours égale à un (première colonne de la matrice X) évite de recourir à un vecteur colonne de constantes. On remarque, par ailleurs, que les indices affectés aux éléments de la matrice X ont une interprétation contraire à celle retenue couramment. Le premier indice désigne la variable explicative (c'est l'indice de la colonne), le second indice est celui de l'observation (c'est l'indice de la ligne).

6.1.2. Les hypothèses de base du MLGS

Le problème statistique à résoudre est le calcul des paramètres β, sachant les observations de Y et de X.

Il existe deux méthodes permettant le calcul de β : la méthode des moindres carrés ordinaires (MCO) et la méthode du maximum de vraisemblance. L'utilisation de ces deux méthodes conduit à des valeurs calculées de β qui possèdent des propriétés statistiques remarquables. Cela implique qu'un certain nombre d'hypothèses de base soient vérifiées avant l'utilisation de ces méthodes.

➔H1 : Les valeurs des **variables explicatives** $X_2, \ldots X_k$ sont **non aléatoires**, contrôlées (indépendantes de l'aléa) : ce sont des données ($\operatorname{cov}(X, \varepsilon) = 0$). Les distributions des Y_t sont donc des distributions conditionnelles de $X_{2t}, \ldots X_{kt}$

➔H2 : Les valeurs des variables exogènes, les variances des variables exogènes, les covariances des exogènes restent finies lorsque n $\rightarrow +\infty$

➔H3 ε est indépendant des variables explicatives. Les valeurs des variables explicatives n'ont aucune influence sur la distribution de ε.

➔H4 ε est un vecteur colonne de n variables aléatoires ε_t et confère au vecteur **Y** un caractère **aléatoire**.

➔ H5 $E[\varepsilon_t] = 0 \quad \forall t \quad t \in \{1, \ldots, n\}$ **hypothèse de normalité.**

$$E[\varepsilon\ ] = \begin{bmatrix} E[\varepsilon_1] \\ . \\ . \\ E[\varepsilon_n] \end{bmatrix} = 0 \quad \forall\, t \quad t \in \{1,\ldots,n\}$$

➔ H6 E[εε'] = σ_{ε}^{2} I_n

$$\text{Or } E[\ \varepsilon\varepsilon'] = E\left[\begin{bmatrix} \varepsilon_1 \\ . \\ . \\ \varepsilon_n \end{bmatrix} \begin{bmatrix} \varepsilon_1 \ldots\ldots\ldots \varepsilon_n \end{bmatrix}\right] = \begin{bmatrix} E[\varepsilon_1^2] & E[\varepsilon_1\varepsilon_2]\ldots\ldots\ldots\ldots E[\varepsilon_1\varepsilon_n] \\ E[\varepsilon_2\varepsilon_1] & E[\varepsilon_2^2]\ldots\ldots\ldots\ldots E[\varepsilon_2\varepsilon_n] \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots \\ E[\varepsilon_n\varepsilon_1] & E[\varepsilon_n\varepsilon_2.]\ldots\ldots\ldots\ldots E[\varepsilon_n^2] \end{bmatrix}.$$

De ce fait

$E[\varepsilon_t^2] = \sigma_{\varepsilon}^2 \quad \forall t\,, t \in \{1 \cdots n\}$ **hypothèse d'homoscédasticité.** Cela implique que la variance des ε_t est constante quel que soit le sous-échantillon tiré dans l'ensemble $\{1 \cdots n\}$.

➔ H7 $E[\varepsilon_t \varepsilon_{t'}] = 0 \quad \forall t \text{ et } t' \in \{1 \cdots n\} \quad t \neq t'$ **hypothèse de non autocorrélation** des résidus.

On peut regrouper les deux dernières hypothèses et écrire :

$$E[\varepsilon_t \varepsilon_{t'}] = \begin{cases} \sigma_{\varepsilon}^2 & \text{si } t = t' \\ 0 & \text{si } t \neq t' \end{cases}$$

Ces trois dernières hypothèses sont fondamentales pour l'application de la méthode des MCO. Pour utiliser la méthode du maximum de vraisemblance, on suppose en outre vérifiée l'hypothèse de normalité des aléas.

➔ H8 La matrice X du modèle est particulière puisqu'on a vu qu'il y avait (k-1) variables explicatives dans le modèle. Le nombre d'observations n est donc supérieur au nombre de paramètres à estimer

➔ H9 Rang[X]= $k < n$. Aucune variable explicative n'est donc linéairement dépendante des autres variables explicatives. La matrice X est régulière. On dit qu'il y a **absence de colinéarité** entre les variables.

Dans ces conditions le modèle à estimer s'écrit :

$$\underset{(n,1)}{Y} = \underset{(n,k)}{X}\ \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon} \qquad \text{avec :}$$

$E[\varepsilon\ ]=0 \qquad E[\varepsilon\varepsilon'] = \sigma_{\varepsilon}^{2}\, I_n \qquad E[\varepsilon_t \varepsilon_{t'}]=0 \quad \forall t \ et\ t' \in \{1 \cdots n\} \quad t \neq t' \quad$ Rang[X]= $k < n$

Plusieurs quantités sont donc inconnues dans ce modèle : les β et σ_{ε}^2. L'objectif des méthodes d'estimation est de trouver des estimateurs de ces paramètres inconnus qui possèdent les propriétés requises par la théorie statistique.

6.1.3. Les estimateurs des moindres carrés ordinaires

Il faut estimer les composantes du vecteur β de la relation linéaire :

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

La méthode des MCO consiste à minimiser la somme des carrés des écarts, écarts entre la valeur observée de la variable endogène (pour un point du nuage) et sa valeur calculée par le modèle. De ce fait, l'écart entre l'observée et le calculé est le résidu.

$$Y - \hat{Y} = e = Y - X\hat{\beta}$$

Soit : $Min \sum_{t=1}^{n} e_t^2 = Min\, e'e$ car $\underset{(n,1)}{e} = \begin{bmatrix} e_1 \\ . \\ . \\ e_n \end{bmatrix}$ e'e = $[e_1, \ldots\ldots e_n] \begin{bmatrix} e_1 \\ . \\ . \\ e_n \end{bmatrix}$ = $e_1^2 + e_2^2 + \ldots . e_n^2 = \sum_{t=1}^{n} e_t^2$

e'e = $(Y - X\hat{\beta})' (Y - X\hat{\beta})$

$$e'e = (Y - X\hat{\beta})'(Y - X\hat{\beta})$$
$$= (Y' - (X\hat{\beta})')(Y - X\hat{\beta}) = (Y' - \hat{\beta}'X')(Y - X\hat{\beta})$$
$$= Y'Y - \hat{\beta}'X'Y - Y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta}$$

Comme $\underset{(1,k)(k,n)(n,1)}{\hat{\beta}'X'Y} = \underset{(1,1)}{\gamma} \in \mathbb{R}$ et

$$\underset{(1,n)(n,k)(k,1)}{Y'X\hat{\beta}} = \underset{(1,1)}{\tilde{\gamma}} \in \mathbb{R}$$

et que la transposée d'une matrice (1,1) (ou d'un scalaire) est la matrice (1,1) (ou le scalaire), on peut écrire :

$$(\hat{\beta}'X'Y)' = Y'X\hat{\beta} = \hat{\beta}'X'Y$$

D'où $e'e = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta}$

De plus, la méthode des MCO impose :

$$\frac{\delta e'e}{\delta \hat{\beta}} = 0$$

Donc $\dfrac{\delta e'e}{\delta \hat{\beta}} = -2\dfrac{\delta(\hat{\beta}'X'Y)}{\delta\hat{\beta}} + \dfrac{\delta(\hat{\beta}'X'X\hat{\beta})}{\delta\hat{\beta}} = 0$

$$\Rightarrow \frac{\delta(\hat{\beta}'X'X\hat{\beta})}{\delta\hat{\beta}} = 2\frac{\delta(\hat{\beta}'X'Y)}{\delta\hat{\beta}} \qquad (1)$$

- L'expression $(\hat{\beta}'X'Y)$ est une forme linéaire en $\hat{\beta}'$ puisque $\hat{\beta}'$ est un vecteur $(1,k)$ et $(X'Y)$ est un vecteur de dimension $(k,1)$

$$\underset{(k,1)}{\hat\beta} = \begin{bmatrix} \hat\beta_1 \\ \vdots \\ \hat\beta_k \end{bmatrix} \qquad \underset{(1,k)}{\hat\beta'} = \begin{bmatrix} \hat\beta_1, \dots & \hat\beta_k \end{bmatrix}$$

supposons

$$\underset{(k,1)}{X'Y} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix} \qquad \text{alors}$$

$$\hat\beta'(X'Y) = \hat\beta_1\alpha_1 + \dots + \hat\beta_k\alpha_k \qquad \text{de sorte que}$$

$$\frac{\delta \hat\beta'(X'Y)}{\delta \hat\beta_1} = \alpha_1 \quad \dots \quad \frac{\delta \hat\beta'(X'Y)}{\delta \hat\beta_k} = \alpha_k$$

- L'expression $\left(\hat\beta'(X'X)\hat\beta\right)$ est une forme quadratique en $\hat\beta$ où $(X'X)$ est une matrice de dimension $(k\ k)$ symétrique, définie positive (puisque le rang de $\hat\beta$ ou de $\hat\beta'$ ne peut qu'être égal au min de $(k,1)$ soit 1). De ce fait :

$$\frac{\delta \hat\beta'(X'X)\hat\beta}{\delta \hat\beta} = 2(X'X)\hat\beta = 2\hat\beta'(X'X)$$

- En reportant ces 2 résultats dans l'expression (1), on obtient
$$2(X'X)\hat\beta = 2X'Y \iff$$
$$(X'X)\hat\beta = X'Y$$

Si la matrice X est de rang k, la matrice $(X'X)$ est de rang k et l'inverse de $(X'X)$ existe. Donc

$$\boxed{\hat\beta = (X'X)^{-1}X'Y}$$

On obtient $\hat\beta = (X'X)^{-1}(X'Y)$ X'X doit être régulière.

6.1.4. Les propriétés des estimateurs des moindres carrés ordinaires

On peut démontrer que les estimateurs des MCO du MLGS possèdent des propriétés identiques à celles du modèle linéaire entre 2 variables

→ **Les estimateurs des MCO sont des fonctions linéaires des Y_t de Y.**

→ **$\hat{\beta}$ est un estimateur (linéaire) sans biais de β**

→ **$\hat{\beta}$ est de variance minimale.**

Les estimateurs des moindres carrés $\hat{\beta}$ sont parmi tous les estimateurs linéaires sans biais les meilleurs au sens de la variance minimale. Cette propriété permet d'établir l'écriture de la matrice des variances-covariances de $\hat{\beta}$, notée : $\Omega_{\hat{\beta}}$

$$\Omega_{\hat{\beta}} = \begin{bmatrix} V[\hat{\beta}_1] \ \mathrm{Cov}[\hat{\beta}_1, \hat{\beta}_2] \ldots\ldots \mathrm{Cov}[\hat{\beta}_1, \hat{\beta}_k] \\ \\ \mathrm{sym} \\ \\ \qquad\qquad\qquad\qquad V[\hat{\beta}_k] \end{bmatrix}$$

qui s'écrit $\Omega_{\hat{\beta}} = E[\ [\hat{\beta} - \beta]\ [\hat{\beta} - \beta]'\ ]$ et qui est égale à : $\Omega_{\hat{\beta}} = \sigma_\varepsilon^2\ (X'X)^{-1}$ où σ_ε^2 est la variance de l'erreur ε

L'estimateur $\hat{\beta}$ de β est un estimateur linéaire, sans biais et de variance minimale. Il est dit **estimateur BLUE** (Best Linear Unbiased Estimator).

→ **$\hat{\beta}$ est un estimateur convergent.**

→ **Estimateur de la variance de l'erreur σ_ε^2**

$\hat{\sigma}_\varepsilon^2 = e'e/(n-k)$ avec $e = Y - X\hat{\beta}$ C'est un estimateur sans biais de σ_ε^2. Il permet de calculer la variance de $\hat{\beta}$.

$V[\hat{\beta}] = \sigma_\varepsilon^2\ \mathrm{DIAG}(X'X)^{-1} \quad \rightarrow \quad V[\hat{\beta}] = \hat{\sigma}_\varepsilon^2\ \mathbf{DIAG(X'X)^{-1}}$

6.2 : Test des estimateurs

Si on suppose que les aléas obéissent à des lois normales centrées et de variance σ_ε^2 :

$$\varepsilon_t \equiv N(0, \sigma_\varepsilon) \quad \forall t, \; t = 1 \ldots n$$

alors ε obéit à une loi normale à n dimensions

$$\varepsilon \equiv N(0, \sqrt{\sigma_\varepsilon^2 I_n})$$

Or on démontre que $\hat{\beta} = \beta + (X'X)^{-1} X'\varepsilon$

Donc $\hat{\beta}$ dépend de ε. D'où $\hat{\beta}$ obéit à une loi normale

$$\hat{\beta} \equiv N(\beta, \sqrt{\sigma_\varepsilon^2 \, DIAG(X'X)^{-1}})$$

Or σ_ε^2 inconnu $\rightarrow$ studentisation

On démontre que $\dfrac{e'e}{\sigma_\varepsilon^2} \equiv \chi^2(n-k) \Leftrightarrow \boxed{(n-k)\dfrac{\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} \equiv \chi^2(n-k)}$

D'où $\boxed{T(n-k) \equiv \dfrac{\hat{\beta} - \beta}{\sqrt{DIAG(X'X)^{-1}} \cdot \hat{\sigma}_\varepsilon}}$

$$H_0 : \beta_j = 0 \qquad j = 1 \ldots k \qquad H_1 \; \beta_j \neq 0$$

$$1 - \alpha = Prob\left[t_{\alpha/2} < T(n-k) < t_{1-\alpha/2} \right]$$
$$= Prob\left[\hat{\beta}_j \in \left[\pm t_{\alpha/2} \hat{\sigma}_\varepsilon \sqrt{DIAG(X'X)^{-1}} \right] \right]$$

Règle de décision :

- si $\hat{\beta}_j \in IA$ H_0 acceptée au risque de 1ère espèce α
- si $\hat{\beta}_j \notin IA$ H_0 rejetée au risque de 1ère espèce α. $\rightarrow$ Validité du modèle

α = risque de 1ère espèce.

$$\hat{\beta} = (x'x)^{-1} x'y$$
$$E[\hat{\beta}] = \beta$$
$$V[\hat{\beta}] = \sigma_\varepsilon^2 (x'x)^{-1}$$

→ En tenant compte des résultats des MCO dans le calcul de $e'e$ où $e = y - x\hat{\beta}$ on trouve :

$$y'y = e'e + \hat{\beta}' x' y$$

qui est l'expression de la décomposition de la variance totale de la variable endogène y centrée.

En effet : • $y'y = \sum y_t^2 = \sum (Y_t - \bar{Y})^2$ est (à n près) la variance totale de y

• $e'e = \sum e_t^2$: variance résiduelle

• $\hat{\beta}' x' y = y'y - e'e$: variance expliquée par $x_1 \dots x_k$

$$\Rightarrow \boxed{VT = VR + VE}$$

→ Par définition, le <u>coefficient de détermination</u> (noté R^2) est le rapport entre la variance expliquée et la variance totale.

$$R^2 = \frac{\hat{\beta}' x' y}{y'y} \qquad \boxed{R^2 = \frac{VE}{VT} = 1 - \frac{VR}{VT}} \qquad 0 \le R^2 \le 1$$

La racine carrée donne le coefficient de corrélation linéaire multiple : r

Autre écriture du test (à privilégier car dans tous les logiciels)

$$t_{c_j} = \frac{\hat{\beta}_j}{\hat{\sigma}_{\hat{\beta}_j}} \sim T(n-k) \quad \text{avec } \hat{\sigma}_{\hat{\beta}_j} = \hat{\sigma}_\varepsilon \sqrt{DIAG(X'X)^{-1}}$$

Règle de décision :

$$\begin{cases} \text{si } |t_{c_j}| \geq T_{1-\alpha/2}(n-k) & H_0 \text{ rejetée au risque de } 1^{\text{ère}} \text{ espèce } \alpha \\ \text{si } |t_{c_j}| < T_{1-\alpha/2}(n-k) & H_0 \text{ acceptée au risque de } 1^{\text{ère}} \text{ espèce } \alpha . \end{cases}$$

6.3. L'analyse de la variance et Test du coefficient de détermination

→ écriture du MLGS avec des variables centrées :

Soit $\begin{cases} x_{jt} = X_{jt} - \bar{X}_j \\ y_t = Y_t - \bar{Y} \end{cases}$ Le MLGS s'écrit sous forme matricielle :

$$y = x\beta + (\varepsilon - \bar{\varepsilon})$$

Dans ce cas, le vecteur β ne comporte pas de terme β_1 et la matrice x n'a plus de première colonne formée de 1. Si on pose $\hat{\varepsilon} = \varepsilon - \bar{\varepsilon}$, on retrouve la forme matricielle du MLGS et de ce fait la méthode des MCO conduit à des résultats identiques à savoir pour l'essentiel :

La matrice des coefficients de corrélation est :

$$\underset{(k,k)}{\hat{R}} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1k} \\ & 1 & & \\ & & \ddots & \\ sym & & & 1 \end{bmatrix}$$

→ <u>Test du coefficient de détermination</u>

$$\frac{\hat{\beta}' x' y / k-1}{e'e / n-k} \sim F(k-1, n-k)$$

or $R^2 = \frac{\hat{\beta}' x' y}{y'y} \Leftrightarrow \hat{\beta}' x' y = R^2 . y'y$

$R^2 = 1 - \frac{e'e}{y'y} \Leftrightarrow e'e = (1-R^2) y'y$

$$\Rightarrow \boxed{F_c = \frac{R^2}{1-R^2} \cdot \frac{n-k}{k-1} \sim F_{1-\alpha}(k-1, n-k)}$$

$H_0 \quad e^2 = 0$

<u>Règle de Décision :</u>

$$\begin{cases} \text{si } F_c < F_{1-\alpha}(k-1, n-k) & H_0 \text{ acceptée au risque de 1}^{\text{re}} \text{ espèce } \alpha \\ \text{si } F_c \geq F_{1-\alpha}(k-1, n-k) & H_0 \text{ rejetée au risque de 1}^{\text{re}} \text{ espèce } \alpha \end{cases}$$

6.4 Utilisation du modèle en prévision

L'utilisation du MLGS pour la prévision revêt les 2 aspects décrits pour le modèle linéaire entre 2 variables (cf cours de statistique)

→ <u>Prévision de la valeur moyenne de la variable endogène pour des valeurs données des variables explicatives</u>

Soit $Y_0 = \beta_1 + \beta_2 X_{20} + \dots + \beta_k X_{k0} + \varepsilon_0$

Soit le vecteur c des variables explicatives :

$$c = \begin{bmatrix} 1 \\ X_{20} \\ \vdots \\ X_{k0} \end{bmatrix}$$

On peut retenir $c'\hat{\beta}$ comme prévision et écrire :

$$\hat{Y}_0 = c'\hat{\beta} = \hat{\beta}_1 + \hat{\beta}_2 X_{20} + \dots + \hat{\beta}_k X_{k0}$$

$c'\hat{\beta}$: estimateur linéaire sans biais de $c'\beta$

↳ Détermination de la loi du prédicteur $c'\hat{\beta}$

$$\begin{cases} E[c'\hat{\beta}] = c'\beta \quad (ESB) \\ V[c'\hat{\beta}] = \sigma^2_\varepsilon \; c'(X'X)^{-1}c \end{cases} \Rightarrow$$

$$c'\hat{\beta} \equiv N\left(\underbrace{c'\beta}_{=Y_0} ; \; \sigma_\varepsilon \sqrt{c'(X'X)^{-1}c}\right)$$

σ_ε → studentisation

<u>Rappel</u> : $$\frac{(n-k)\,\hat{\sigma}^2_\varepsilon}{\sigma^2_\varepsilon} \equiv \chi^2(n-k)$$

$$\Rightarrow \boxed{T(n-k) \equiv \frac{(c'\hat{\beta} - c'\beta)}{\hat{\sigma}_\varepsilon \sqrt{c'(X'X)^{-1}c}}}$$

On peut alors construire un intervalle de confiance de cette prévision :

$$1-\alpha = \mathrm{Prob}[\, t_{\alpha/2} < T(n-k) < t_{1-\alpha/2} \,]$$

$$= \mathrm{Prob}\Big[\, \underbrace{c'\beta}_{=Y_0} \in \Big[\underbrace{c'\hat{\beta}}_{=\hat{Y}_0} \pm t_{\alpha/2}\, \hat{\sigma}_\varepsilon \sqrt{c'(X'X)^{-1}c} \,\Big]\Big]$$

→ Comparaison d'une prévision ponctuelle à la droite des MCO

Soit $[Y_0, x_{10} \ldots x_{k0}]$ une prévision dont on se demande si elle est compatible avec le résultat de l'estimation par la méthode des MCO.

$$\hat{Y}_0 = c'\hat{\beta} \to E[c'\hat{\beta}] = c'\beta$$

$$V[c'\hat{\beta}] = \sigma_\varepsilon^2 \sqrt{1 + c'(X'X)^{-1}c}$$

$$\underbrace{c'\hat{\beta}}_{=\hat{Y}_0} \equiv N\Big(\underbrace{c'\beta}_{=Y_0} ;\ \sigma_\varepsilon \sqrt{1 + c'(X'X)^{-1}c}\Big)$$

utilisation ici aussi de la loi de Student

$$\boxed{T(n-k) \equiv \frac{c'\hat{\beta} - c'\beta}{\hat{\sigma}_\varepsilon \sqrt{1 + c'(X'X)^{-1}c}}}$$

H_0 : compatibilité entre ajustement et prévision

$Y_0 = c^{te}$

$$1-\alpha = \text{Prob}\left[\hat{Y}_0 \in \left[Y_0 \pm t_{\alpha/2}\, \hat{\sigma}_\varepsilon \sqrt{1 + c'(X'X)^{-1}c} \right] \right]$$

Règle de décision :

si $\hat{Y}_0 \in IA$ H_0 acceptée au risque de 1ère espèce α

si $\hat{Y}_0 \notin IA$ H_0 rejetée au risque de 1ère espèce α