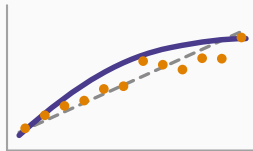


Économétrie avancée

Annexe C : Rappels de statistique mathématique



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Avant de commencer

Échantillonnage et estimateurs

Propriétés à échantillon fini

Propriétés asymptotiques

Méthodes générales d'estimation

Estimation par intervalle et tests d'hypothèses

Synthèse

Avant de commencer

Pourquoi ce chapitre ?

Le chapitre 11 a traité les variables aléatoires comme des objets de population. Ce chapitre franchit l'étape suivante : comment un **estimateur** — construit à partir d'un échantillon — hérite lui-même d'une distribution de probabilité, et quels critères permettent de juger qu'il est « bon ». C'est le socle exact des propriétés (non-biais, BLUE, convergence, normalité asymptotique) invoquées depuis le chapitre 2 sans être redémontrées.

Échantillonnage et estimateurs

Échantillon aléatoire (i.i.d.)

$\{Y_1, \dots, Y_n\}$ constitue un **échantillon aléatoire** tiré de $f(y; \theta)$ si les Y_i sont indépendantes et identiquement distribuées (i.i.d.) selon cette même densité. C'est l'hypothèse MLR.2 du chapitre 2, formulée dans le langage général de la statistique mathématique.

Estimateur vs estimation

Un **estimateur** $W = h(Y_1, \dots, Y_n)$ est une règle, fixée avant l'échantillonnage, qui associe à tout échantillon possible une valeur numérique — c'est une variable aléatoire, dotée d'une **distribution d'échantillonnage**. Une fois les données observées, $w = h(y_1, \dots, y_n)$ est une **estimation** : un simple nombre.

Propriétés à échantillon fini

Estimateur non biaisé

W est **non biaisé** pour θ si $E(W) = \theta$ pour toute valeur de θ . Le non-biais ne garantit **pas** qu'une estimation particulière soit proche de θ — seulement que la moyenne sur toutes les répétitions possibles de l'échantillonnage coïnciderait avec θ .

Le non-biais ne suffit pas

Estimer μ par la seule première observation $W = Y_1$ est non biaisé ($E(Y_1) = \mu$) mais gaspille presque toute l'information de l'échantillon : sa variance, $\text{Var}(Y_1) = \sigma^2$, ne diminue jamais avec n , alors que celle de la moyenne d'échantillon $\bar{Y}$ vaut σ^2/n . Le non-biais seul ne classe donc pas les estimateurs entre eux — il faut aussi comparer les variances.

Efficacité relative et erreur quadratique moyenne (EQM)

Entre deux estimateurs non biaisés, W_1 est **efficace relativement** à W_2 si $\text{Var}(W_1) \leq \text{Var}(W_2)$ (strictement pour au moins une valeur de θ). Pour comparer des estimateurs même biaisés, l'**erreur quadratique moyenne** $EQM(W) = E[(W - \theta)^2] = \text{Var}(W) + [\text{Biais}(W)]^2$ combine les deux sources d'erreur en un seul critère.

Propriétés asymptotiques

Convergence (consistance)

W_n est **convergent** pour θ si, pour tout $\varepsilon > 0$, $P(|W_n - \theta| > \varepsilon) \rightarrow 0$ quand $n \rightarrow \infty$ – noté $\text{plim}(W_n) = \theta$. Un estimateur non biaisé dont la variance tend vers zéro est automatiquement convergent (cas de $\bar{Y}$, via la **loi des grands nombres**).

Propriétés du plim

Pour g continue, $\text{plim } g(W_n) = g(\text{plim } W_n)$; et si $\text{plim}(T_n) = \alpha$, $\text{plim}(U_n) = \beta$, alors $\text{plim}(T_n + U_n) = \alpha + \beta$, $\text{plim}(T_n U_n) = \alpha\beta$, et $\text{plim}(T_n/U_n) = \alpha/\beta$ (si $\beta \neq 0$). Ces règles simples justifient la convergence de fonctions complexes d'estimateurs convergents – exactement la logique utilisée au chapitre 5 pour établir la convergence des MCO.

Théorème central limite (TCL)

Pour tout échantillon i.i.d. de moyenne μ et variance σ^2 finie,

$$Z_n = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1),$$

quelle que soit la distribution sous-jacente de Y . C'est un résultat remarquable : la normalité de $\bar{Y}$ en grand échantillon ne dépend d'aucune hypothèse sur la forme de la population.

Le pont vers le chapitre 5 de ce cours

Le TCL est exactement l'outil invoqué pour justifier la normalité asymptotique des MCO sans supposer u normal (chapitre 5) : la plupart des estimateurs économétriques s'écrivent comme fonctions de moyennes d'échantillon, et héritent donc de la normalité asymptotique du TCL par les propriétés du plim ci-dessus.

Méthodes générales d'estimation

Trois grandes familles

Méthode des moments, maximum de vraisemblance, moindres carrés

La **méthode des moments** remplace un moment de population (μ, σ_{XY}) par son analogue empirique — c'est ainsi qu'on obtient $\bar{Y}$, la covariance ou la corrélation échantillonnales. Le **maximum de vraisemblance (MV)** choisit la valeur de θ qui rend l'échantillon observé le plus probable, en maximisant $L(\theta) = \prod_i f(y_i; \theta)$; il est généralement le plus efficace asymptotiquement quand le modèle est correctement spécifié. Les **moindres carrés** minimisent une somme de carrés d'écarts — $\bar{Y}$ elle-même est l'estimateur des moindres carrés de μ . Ces trois principes coïncident souvent (normale, Bernoulli) mais pas toujours.

Application : où ces méthodes réapparaissent

Les MCO du chapitre 2 sont un estimateur des moindres carrés; l'estimateur du maximum de vraisemblance réapparaît explicitement dans les modèles GARCH (chapitre 4, économétrie financière), où la vraisemblance gaussienne est maximisée numériquement faute de formule fermée — un cas où les trois principes ne coïncident plus et où le choix de la méthode d'estimation devient un choix de modélisation à part entière.

Estimation par intervalle et tests d'hypothèses

Interprétation correcte d'un intervalle de confiance

Un IC à 95 % pour μ , $[\bar{y} \pm c \cdot se(\bar{y})]$, ne signifie **pas** “ μ a 95 % de chances d’être dans cet intervalle” une fois l’échantillon observé — μ et les bornes sont alors des nombres fixes. La bonne lecture : sur 95 % des échantillons aléatoires possibles, l’intervalle **construit selon cette procédure** contiendrait μ .

Règle pratique à retenir

Pour n grand, un IC à 95 % approché est simplement [estimation $\pm 2 \times$ erreur-type] — le facteur 1,96 (loi normale) ou la valeur t correspondante (petits échantillons, chapitre 4) se confondant avec 2 dès que n dépasse une vingtaine d’observations.

Hypothèse nulle et hypothèse alternative

H_0 est présumée vraie jusqu'à preuve statistique contraire "au-delà du doute raisonnable" — exactement le même principe que le test t du chapitre 4, ici formulé dans le cadre général de l'inférence statistique plutôt que spécifiquement pour un coefficient de régression.

Discrimination à l'embauche : un test par paires

Sur 241 paires de candidatures identiques hormis la couleur de peau, 22,4 % des candidats noirs reçoivent une offre contre 35,7 % des candidats blancs. L'IC approché à 95 % pour l'écart de probabilité, $-.133 \pm 1.96(.482/\sqrt{241}) \approx [-.164, -.102]$, exclut nettement zéro — même l'IC à 99 % exclut zéro. C'est un exemple archétypal de test par paires appariées (*matched pairs*), une méthode de contrôle des facteurs de confusion évoquée dès le chapitre 3.

Synthèse

L'essentiel du chapitre 12

- Un estimateur est une variable aléatoire; le non-biais seul ne suffit pas à le juger bon — la variance (ou l'EQM pour les estimateurs biaisés) doit être prise en compte.
- La convergence et la normalité asymptotique (via le théorème central limite) sont les propriétés qui justifient l'inférence en grand échantillon, y compris quand la population sous-jacente n'est pas normale.
- Méthode des moments, maximum de vraisemblance et moindres carrés sont trois grands principes d'estimation, qui coïncident souvent mais pas toujours.
- Un intervalle de confiance se lit comme une propriété de la **procédure** répétée, pas comme une probabilité sur le paramètre une fois l'échantillon observé.

Vers le chapitre 13

Chapitres 11 et 12 forment ensemble le socle probabiliste et statistique de tout ce cours. Les deux derniers chapitres reformulent la régression multiple elle-même en **notation matricielle** — un langage plus compact qui généralise immédiatement tout ce qui a été vu, du non-biais des MCO (chapitre 2) à l'inférence en grand échantillon (chapitre 5).