Chapitre 1

- Cv $X_n \xrightarrow[n\to+\infty]{p.s.} X \Rightarrow X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X \Rightarrow X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$

$$\Uparrow$$

$$X_n \xrightarrow[n\to+\infty]{L^2} X$$

- $\boxed{X_n \xrightarrow{\mathcal{L}} X \text{ et } Z_n \xrightarrow{\mathbb{P}} a, \ a \text{ cste} \Rightarrow X_n Z_n \xrightarrow{\mathcal{L}} aX \text{ et } X_n + Z_n \xrightarrow{\mathcal{L}} X + a}$ Slutsky

- ⚠ $X_n \xrightarrow{\mathcal{L}} X$ et $Z_n \xrightarrow{\mathcal{L}} Z \not\Rightarrow X_n Z_n \xrightarrow{\mathcal{L}} XZ$, $X_n + Z_n \xrightarrow{\mathcal{L}} X + Z$

La cv en loi elle n'est pas stable par addition ni par mult.

- Inégalités Markov, Cauchy-Schwarz, ...

- 2 théorèmes. $(X_n)_{n\geq 1}$ suite de v.a. i.i.d.

	consistance	normalité asymptotique
$\frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n$	LGN	TCL
$g(\bar{X}_n)$, g continue	Théorème de continuité	Δ méthode

Chapitre 2

- EQM (erreur quad. moyenne) $\hat{\theta}$ estimateur de θ

$$\mathbb{E}\left[(\hat{\theta}-\theta)^2\right] = \underbrace{\text{Biais}^2(\hat{\theta})}_{=\mathbb{E}[\hat{\theta}-\theta]^2} + \underbrace{\text{Var}(\hat{\theta})}_{=\mathbb{E}[\hat{\theta}^2]-\mathbb{E}[\hat{\theta}]^2}$$

- $X_1, \dots, X_n$ v.a. iid de moyenne μ, variance σ^2.

- moyenne empirique. $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$,

(estimateur de μ)

(non biaisé
fortement consistant
asymptotiquement normal)

- variance empirique $\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$

(estimateur de σ^2)

(non biaisé
fortement consistant
asymptotiquement normal)

Méthode des moments

- On calcule $\mathbb{E}[X] = \varphi(\theta)$.
- Un estimateur naturel de $\varphi(\theta)$ est $\overline{X}_n$.
- (Si φ est inversible), un estimateur par la méthode des moments est $\varphi^{-1}(\overline{X}_n)$.

Estimateur du max de vraisemblance

- $\hat{\theta}^{MV} \in \underset{\theta \in \Theta}{\operatorname{argmax}} \; L_x(\theta)$

$$L_x(\theta) = \begin{cases} \prod_{i=1}^{n} f_{x_i}(\theta), & X \text{ continue.} \\ \prod_{i=1}^{n} \mathbb{P}(X = x_i), & X \text{ discrète.} \end{cases}$$

$\mathbb{1}_{\{x_i > 0\}}$ ✓ (on dérive pas par rapport à x)

1er cas

vous pouvez dériver la vraisemblance.

$$L_x'(\theta) = \frac{\partial L}{\partial \theta}(\theta).$$

$$L_x'(\hat{\theta}) = 0 \iff \hat{\theta} = \dots \text{ candidat pour l'EMV}$$

$\hat{\theta}$ maximise L_x : concavité de L_x

2ème cas

vous ne pouvez pas dériver la vrais.

$$L_x(\theta) \propto \mathbb{1}_{\{\dots\theta\}}$$

$\times$

raisonner sur l'indicatrice

↳ $\hat{\theta}^{MV} = \min X_i \;/\; \max X_i$

Chapitre 3 — Intervalle de confiance pour θ

θ : quantité inconnue que l'on souhaite estimer

↳ **asymptotique** : on cherche A_n, B_n des quantités qui ne dépendent de θ,

$$\lim_{n \to +\infty} \underbrace{\mathbb{P}(\theta \in [A_n, B_n])}_{= \mathbb{P}(A_n \leq \theta \leq B_n)} = 1 - \alpha$$

Si l'on cherche un IC de niveau $1-\alpha$

niveau de l'IC

AN :

IC de niveau 95%

$\alpha = 0{,}05$

↳ **non asymptotique** (il est vrai pour tout n ⇒ IC + fiable !)

on cherche A_n, B_n, ...

$$\forall n \in \mathbb{N}, \quad \mathbb{P}(\theta \in [A_n, B_n]) = 1 - \alpha.$$

• Méthode 1

→ partir d'une convergence asymptotique

$$\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

→ IC asymptotique

→ partir d'une loi

$$\forall n \in \mathbb{N},\ \sqrt{n}\,(\hat{\theta}_n - \theta) \overset{\mathcal{L}}{\sim} \mathcal{E}xp(\lambda)$$

→ IC non asymptotique

• se ramener à une $\mathcal{N}(0,1)$.

$$\frac{\sqrt{n}\,(\hat{\theta}_n - \theta)}{\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

• σ dépend de θ (exemple $\sigma = \theta^2$) ⇝ plug-in

remplacer θ au dénominateur par $\hat{\theta}$

⇝ étudier

Mq $$\frac{\sqrt{n}\,(\hat{\theta}_n - \theta)}{\hat{\theta}_n^2} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

$\boxed{\text{SLUTSKY}}$

avec $q_{1-\alpha/2}$: quantile d'ordre $1-\alpha/2$ d'une loi $\mathcal{N}(0,1)$.

• on encadre θ : $$\lim_{n\to+\infty} \mathbb{P}\left(-q_{1-\alpha/2} \leq \frac{\sqrt{n}\,(\hat{\theta}_n - \theta)}{\sigma} \leq q_{1-\alpha/2}\right) = 1-\alpha$$

• Méthode 2 (Exos 3, 4) Q12 IC de niveau $1-\alpha$:

$X_{(n)}$, $X_{(1)}$ (max, min X_i)

• On a $X_{(n)} \leq \theta$ p.s. (on une inégalité coûteux car elle est p.s vraie).

le plus souvent : non asymptotique

• Je cherche l'IC de la forme

$$[X_{(n)} ; \underbrace{X_{(n)} d_\alpha}] \quad \text{(forme finale)}$$

$\mathbb{P}(\theta \in [X_{(n)}, X_{(n)} + c_\alpha]) = 1-\alpha$ $\xrightarrow{\text{Forme initiale}}$ $\underbrace{[X_{(n)} ; X_{(n)} + c_\alpha]}$

• Ici, on cherche c_α tq :

$$\mathbb{P}(\theta \leq X_{(n)} + c_\alpha) = 1-\alpha$$

$\rightsquigarrow$ Utiliser la fdR de $X_{(n)}$

$$\mathbb{P}(\theta - c_\alpha \leq X_{(n)}) = 1-\alpha$$

$$1 - \mathbb{P}(X_{(n)} \leq \theta - c_\alpha) = 1-\alpha$$

TD3 → Exo 2, Exo 3 (ou 4), Exo 5

Exo 2

$X_1, \dots, X_p$ des v.a iid $\mathcal{N}(\mu_1, \sigma^2)$ $\quad X_1, \dots, X_p \perp\!\!\!\perp Y_1, \dots, Y_p$.

$Y_1, \dots, Y_q$ ———— $\mathcal{N}(\mu_2, \sigma^2)$.

Estimer $\mu_1 - \mu_2$

1) Un estimateur naturel pour μ_1 est $\overline{X}_p$

———— μ_2 est $\overline{Y}_q$

Donc ———— $\mu_1 - \mu_2$ est $\overline{X}_p - \overline{Y}_q = D$

2) Rappel combinaison linéaire de v.a. gaussienne est gaussienne.

$D = \overline{X}_p - \overline{Y}_q$ est une combinaison linéaire de v.a gaussienne.

$1 \times \overline{X}_p + (-1)\overline{Y}_q$.

$\Rightarrow D$ gaussienne

$D \sim \mathcal{N}(\mathbb{E}[D], Var(D))$

Pour caractériser la loi d'une v.a. Gaussienne

calculer $\mathbb{E}$, Var.

- $\mathbb{E}[D] = \mathbb{E}[\bar{X}_p - \bar{Y}_q] \underset{\substack{\uparrow \\ \text{linéarité}}}{=} \mathbb{E}[\bar{X}_p] - \mathbb{E}[\bar{Y}_q] = \mu_1 - \mu_2$.

- $\mathrm{Var}(D) = \mathrm{Var}(\bar{X}_p - \bar{Y}_q)$ ⊥⊥ des échantillons $(X_1, \dots, X_p)$ $(Y_1, \dots, Y_q)$

$$= \mathrm{Var}(\bar{X}_p) + \mathrm{Var}(-\bar{Y}_q)$$

$$= \mathrm{Var}(\bar{X}_p) + (-1)^2\, \mathrm{Var}(\bar{Y}_q)$$

$$= \mathrm{Var}(\bar{X}_p) + \mathrm{Var}(\bar{Y}_q)$$

$$= \mathrm{Var}\left(\frac{1}{p}\sum_{i=1}^{p} X_i\right) + \mathrm{Var}\left(\frac{1}{q}\sum_{i=1}^{q} Y_q\right)$$

$X_1, \dots, X_p$ ⊥⊥
$Y_1, \dots, Y_q$ ⊥⊥

$$= \frac{1}{p^2}\sum_{i=1}^{p} \mathrm{Var}(X_i) + \frac{1}{q^2}\sum_{i=1}^{q} \mathrm{Var}(Y_q)$$

m̂ loi $X_1 \dots$
$Y_1 \dots$

$$= \left(\frac{1}{p} + \frac{1}{q}\right)\sigma^2$$

$$D \sim \mathcal{N}\left(\mu_1 - \mu_2,\ \left(\frac{1}{p} + \frac{1}{q}\right)\sigma^2\right).$$

3) Estimateur de σ^2

Il y a deux estimateurs naturels (non biaisés)

$$S_X^2 = \frac{1}{p-1}\sum_{i=1}^{p}(X_i - \bar{X}_p)^2$$

$$S_Y^2 = \frac{1}{q-1}\sum_{i=1}^{q}(Y_i - \bar{Y}_q)^2$$

on a envie de poser un estimateur pour σ^2 qui prenne S_X^2 et S_Y^2 en compte.

Soit $\tilde{S} = \cancel{\dfrac{S_X^2 + S_Y^2}{2}}$? $\rightsquigarrow$ biaisé $\times$

pas les bonnes propriétés.

Rappel (cours) $\dfrac{(p-1)\,S_X^2}{\sigma^2} \sim \chi^2_{p-1}$ et $\dfrac{q-1\ S_Y^2}{\sigma^2} \sim \chi^2_{q-1}$

Par le théorème de Cochran (S_X^2 et S_Y^2 sont indépendants car les échantillons sont $\perp\!\!\!\perp$)

$$\frac{(p-1)\,S_X^2}{\sigma^2} + \frac{(q-1)\,S_Y^2}{\sigma^2} \sim \chi^2_{p+q-2} \quad (*)$$

Je cherche S^2 tq : →

On veut que
$$\boxed{\frac{k\,S^2}{\sigma^2} \overset{\mathcal{L}}{\sim} \chi^2_k .}$$

(bonne propriété de l'estimateur de la variance)

Je cherche S^2 ?

Posons
$$S^2 = \frac{1}{p+q-2}\Big((p-1)\,S_x^2 + (q-1)\,S_y^2\Big).$$

J'ai bien
$$\frac{(p+q-2)\;S^2}{\sigma^2} \sim \chi^2_{p+q-2}$$
(en utilisant (*))

4) On veut mq $\boxed{\dfrac{D - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{p} + \frac{1}{q}}}} \sim T_{p+q-2}$

On va réécrire la statistique :

$$\frac{D-(\mu_1-\mu_2)}{S\sqrt{\frac{1}{p}+\frac{1}{q}}} = \frac{D-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{p}+\frac{1}{q}}} \quad \frac{1}{\frac{S}{\sigma}} = \frac{D-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{p}+\frac{1}{q}}}\;\frac{1}{\sqrt{\frac{S^2}{\sigma^2}(p+q-2)}}\sqrt{p+q-2}$$

- $D \sim \mathcal{N}\left(\mu_1-\mu_2, \sigma^2\left(\frac{1}{p}+\frac{1}{q}\right)\right)$

$$\Rightarrow \frac{D-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{p}+\frac{1}{q}}} \sim \mathcal{N}(0,1)$$

- $(p+q-2)\dfrac{S^2}{\sigma^2} \sim \chi^2_{p+q-2}$

(question 3)

Dans le cours on a :

$$\frac{\dfrac{D-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{p}+\frac{1}{q}}}}{\sqrt{p+q-2\left(\frac{S^2}{\sigma^2}\right)}} \times \sqrt{p+q-2}$$

$\sim T_{(p+q-2)}$

5) $$\boxed{\frac{D-(\mu_1-\mu_2)}{S\left(\frac{1}{p}+\frac{1}{q}\right)^{1/2}} \sim T_{p+q-2}} \quad (*)$$

estimateur donc connu !

$$\frac{D-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{p}+\frac{1}{q}}} \sim \mathcal{N}(0,1)$$

Si je pars de là, je vais me retrouver bloqué car σ et $\mu_1-\mu_2$ sont inconnus !

$\mu \triangleq \mu_1-\mu_2$

- On note $t_{1-\alpha/2}$ le quantile de T_{p+q-2} d'ordre $1-\alpha/2$.

- $(*)$ implique

$$\mathbb{P}\left(-t_{1-\alpha/2} \leq \frac{D-\mu}{S\sqrt{\frac{1}{p}+\frac{1}{q}}} \leq t_{1-\alpha/2}\right) = 1-\alpha$$

$$= \mathbb{P}\left(-t_{1-\alpha/2}S\sqrt{\frac{1}{p}+\frac{1}{q}} - D \leq -\mu \leq t_{1-\alpha/2}S\sqrt{\frac{1}{p}+\frac{1}{q}} - D\right) = 1-\alpha$$

$$= \mathbb{P}\left(D - t_{1-\alpha/2}S\sqrt{\frac{1}{p}+\frac{1}{q}} \leq \mu \leq D + t_{1-\alpha/2}S\sqrt{\frac{1}{p}+\frac{1}{q}}\right) = 1-\alpha$$

$$IC_\mu = \left[D \pm t_{1-\alpha/2} S \sqrt{\frac{1}{q} + \frac{1}{p}} \right].$$

6) H_0 : "$\mu_1 = \mu_2$" contre H_1 : "$\mu_1 \neq \mu_2$" (test)

"$\mu_1 - \mu_2 = \mu = 0$" "$\mu_1 - \mu_2 = \mu \neq 0$"

↳ $$S^2 = \frac{1}{p+q-2} \left(\underbrace{(p-1) S_x^2}_{} + (q-1) S_y^2 \right).$$

$$S_x^2 = \frac{1}{p-1} \sum (x_i - \bar{X}_p)^2 = \frac{1}{p-1} \left(\sum x_i^2 - p \bar{X}_p^2 \right).$$

$p = 15$
$q = 9$

$$= \frac{1}{22} \left(\sum_{i=1}^{15} x_i^2 - 15 \underbrace{\bar{X}_{15}^2}_{\left(\frac{1}{15}\sum_{i=1}^{15} x_i\right)^2} \right) + \left(\sum_{j=1}^{9} y_j^2 - 9 \underbrace{\bar{Y}_9^2}_{\left(\frac{1}{9}\sum_{j=1}^{9} y_j\right)^2} \right)$$

AN $S^2 = 1{,}54$.

$$T(\mu) = \frac{(\bar{X}_p - \bar{Y}_q) - \mu}{S\sqrt{\frac{1}{p} + \frac{1}{q}}}$$

↳ Sous H_0, la statistique de test est

$$T(0) = \frac{\bar{X}_{15} - \bar{Y}_9}{S\left(\frac{1}{15} + \frac{1}{9}\right)^{1/2}} \sim T_{22}.$$

- Zone de rejet :

Test de risque 1%.

$1-\alpha = 0,99$.

$1-\frac{\alpha}{2} = 0,995$

$$ZR = \left\{ |T(0)| \geq t_{p+q-2,\,1-\alpha/2} \right\}.$$

$$= \left\{ |T(0)| \geq t_{22\,,\,0,995} \right\} \qquad t_{22,0,995} = 2,819$$

$$|T(0)| = 0,174.$$

On a $|T(0)| < 2,819$. $\Rightarrow$ on ne rejette pas H_0.

Vocabulaire

Test $\rightarrow$ on n'écrit pas

« on accepte H_1 ».

$\rightarrow$ on rejette H_0.

on ne rejette pas H_0.