

Notes de Cours d'Inférence Statistique

YEHOR KOROTENKO

13 April 2026

AVANT-PROPOS

Ce document proviens des notes de cours d'Inférence Statistique prises par Yehor Korotenko aux lectures faites par Madame Marie-Anne Poursat. Le document contient toutes les définitions, théorèmes et les preuves mais peu d'explications, de plus, les notes sont assez seces en soit.

Ces notes doivent servir plutôt comme une référence ou un rappel rapide (CheatSheet). L'utilisation est conseillé avec les visites des lectures. Ces notes peuvent être retravaillées et être complétées en livre complet mais rien est promis.

Table des matieres

1 Introduction	3
1.1 Evaluation	3
1.2 Modèle Statistique	3
1.3 Estimateurs	3
1.4 Risque quadratique	4
1.4.1 Exemple : Modèle de Poisson	4
1.5 Consistance	5
1.5.1 Exemple : Retour au modèle de Poisson	5
1.5.2 Méthode « Plug-in »	6
2 Estimateurs	7
2.1 Cadre paramétrique	7
2.1.1 Modèle statistique paramétrique	7
2.2 Méthode des moments	8
2.3 Rendu sur le L.A.C.	9
2.3.1 Variance empirique	10
2.4 Méthode de maximum de vraisemblance	11
2.4.1 Modèle donné	11
2.4.2 En pratique	11
3 Information de Fisher, efficacité	13
3.1 Modèle régulier	13
3.2 Score et Information de Fisher	14
3.3 Information de Fisher et dérivée seconde	16
3.4 Inégalité de Cramer - Rao	17
4 Étude asymptotique des estimateurs	19
4.1 Convergences	19
4.2 Consistance des estimateurs	20

4.3	Normalité asymptotique	20
4.4	δ -méthode	22
5	Fonction de répartition empirique	23
5.1	Estimation empirique	24
5.2	Inverse généralisé	24
5.3	Quantile empirique	26
6	Intervalles de confiance	27
6.1	Définitions	27
6.2	Interprétation	27
6.3	Méthode pivotale	27
7	Compléments (avant partiel)	29
7.1	Propriétés asymptotiques d'une suite d'estimateurs $(\hat{\theta}_n)_{n \geq 1}$	29
7.2	Pivot (asymptotique) ou statistique pivotale	31
8	Estimation dans les échantillons gaussiens	34
8.1	TL&DR	34
8.2	Loi normale et lois dérivées	34
8.3	Loi des estimateurs empiriques	36
8.4	IC des paramètres	38
8.5	Exercice	38
9	Introduction aux tests statistiques	39
9.1	Exemple	39
9.1.1	Contrôle de qualité: industriel.	39
9.1.2	Modélisation	39
9.2	Principe d'un test	40
9.3	Risque d'erreur	40
9.4	Construction d'un test	41
10	Tests d'hypothèse (sur un paramètre)	44
10.1	Formalisme d'un test	44
10.1.1	Introduction	44
10.1.2	Risques d'erreur d'un test	44
10.2	Exemple	45
10.3	Construction d'un test	47
11	Tests d'un paramètre gaussien	48
11.1	Résumé de la construction	48
11.2	P -valeur	48
11.2.1	Généralisation(formule de calcul d'une p -valeur).	49
11.2.2	Remarques	50
11.2.3	Règle de décision avec la p -valeur	50
12	Test de Student (t-test)	53
12.1	Statistique de test	53
12.2	Région de rejet	54
12.3	Règle de décision	54

Introduction

§1

1.1 EVALUATION

- 0.4 CC +0.6 Examen.
- Répartition : 80% partiel, 20% Interro (prévue le 26/01).

1.2 MODÈLE STATISTIQUE

DEFINITION 1.1 (MODÈLE STATISTIQUE) – *Un modèle statistique est un espace de probabilité $(\Omega, \mathcal{A}, \mathcal{P})$ où \mathcal{P} est une famille de lois de probabilité $\{P_\theta; \theta \in \Theta\}$.*

- Si $\exists p \in \mathbb{N}^*, \Theta \subset \mathbb{R}^p$: modèle paramétrique.
- Sinon : modèle non paramétrique.

EXEMPLE 1.2 (FAMILLES DE LOIS) –

- *Lois de Poisson* : $\mathcal{P} = \{P(\lambda); \lambda > 0\}$.
- *Densité régulière* : $\mathcal{P} = \{\mathbb{P}; \mathbb{P} \text{ dont la densité admet une dérivée seconde bornée}\}$.

◇

DEFINITION 1.3 (OBSERVATION) – *Une observation est une variable aléatoire (v.a.) dont la loi appartient à $\{P_\theta, \theta \in \Theta\}$. Notre observation aura une structure de n -échantillons X_1, \dots, X_n i.i.d. (indépendants et identiquement distribués) de loi commune $\in \{P_\theta, \theta \in \Theta\}$.*

REMARQUE 1.4 – (X_1, \dots, X_n) est de loi $P_\theta^{\otimes n}$. L'échantillon contient toute l'information sur P_θ , donc sur θ . ◇

DEFINITION 1.5 (IDENTIFIABILITÉ) – *Un modèle est identifiable si et seulement si (ssi) l'application $\theta \mapsto P_\theta$ est injective.*

1.3 ESTIMATEURS

Hypothèse : On observe X_1, \dots, X_n i.i.d. de loi commune $\in \{P_\theta, \theta \in \Theta \subset \mathbb{R}^p\}$ (modèle paramétrique identifiable). Soit θ^* la vraie valeur inconnue telle que $P_{X_i} = P_{\theta^*}$.

DEFINITION 1.6 (ESTIMATEUR) – *Un estimateur de θ est une fonction de l'échantillon (X_1, \dots, X_n) mesurable et indépendante de θ (calculable à partir des données).*

Notation : $\hat{\theta} = \hat{\theta}_n = h(X_1, \dots, X_n)$. C'est une variable aléatoire.
Exemples : $\hat{\theta} = \bar{X}$, $\hat{\theta} = X_1 - X_3$, etc.

Questions fondamentales :

1. Comment définir un bon estimateur ?
2. Comment construire un bon estimateur ?

1.4 RISQUE QUADRATIQUE

Idée : En moyenne, $\hat{\theta}$ doit être proche de θ . On regarde $\mathbb{E}[\hat{\theta} - \theta]$.

DEFINITION 1.7 (BIAIS) – Le biais de $\hat{\theta}$ est défini par :

$$B(\hat{\theta}, \theta) = \mathbb{E}[\hat{\theta}] - \theta$$

On dit que $\hat{\theta}$ est *sans biais* si $B(\hat{\theta}, \theta) = 0$.

DEFINITION 1.8 (RISQUE QUADRATIQUE / MSE) –

$$R(\hat{\theta}, \theta) = \mathbb{E}[(\hat{\theta} - \theta)^2]$$

C'est la *Mean Squared Error (MSE)* en anglais.

On dit que $\hat{\theta}_1$ est meilleur que $\hat{\theta}_2$ ssi $R(\hat{\theta}_1, \theta) \leq R(\hat{\theta}_2, \theta)$.

1.4.1 EXEMPLE : MODÈLE DE POISSON

Soit X_1, \dots, X_n de loi P_θ de Poisson, $\theta > 0$. On cherche un estimateur de $\theta = \mathbb{E}[X_i]$.

Proposons : $\hat{\theta} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

Calcul du Biais :

$$\begin{aligned} B(\hat{\theta}, \theta) &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] - \theta \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] - \theta \quad (\text{par linéarité}) \\ &= \frac{1}{n} \cdot n \cdot \mathbb{E}[X_1] - \theta \\ &= \theta - \theta = 0 \end{aligned}$$

Donc $\mathbb{E}[\bar{X}] = \theta$, est l'estimateur sans biais.

Calcul du Risque :

$$\begin{aligned}
R(\hat{\theta}, \theta) &= \mathbb{E}[(\bar{X} - \theta)^2] = \mathbb{E}[(\bar{X} - \mathbb{E}[\bar{X}])^2] \\
&= \text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum X_i\right) \\
&= \frac{1}{n^2} \sum \text{Var}(X_i) \quad (\text{car i.i.d}) \\
&= \frac{1}{n^2} \cdot n \cdot \text{Var}(X_1) = \frac{\text{Var}(X_1)}{n} = \frac{\theta}{n}
\end{aligned}$$

THEOREME 1.9 (DÉCOMPOSITION BIAIS-VARIANCE DU RISQUE) –

$$R(\hat{\theta}, \theta) = (B(\hat{\theta}, \theta))^2 + \text{Var}(\hat{\theta})$$

Preuve.

$$\begin{aligned}
R(\hat{\theta}, \theta) &= \mathbb{E}[(\hat{\theta} - \theta)^2] \\
&= \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}] + \mathbb{E}[\hat{\theta}] - \theta)^2] \\
&= \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2] + \mathbb{E}[(\mathbb{E}[\hat{\theta}] - \theta)^2] + 2\mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])(\mathbb{E}[\hat{\theta}] - \theta)] \\
&= \text{Var}(\hat{\theta}) + (B(\hat{\theta}, \theta))^2 + 2(\mathbb{E}[\hat{\theta}] - \theta) \underbrace{\mathbb{E}[\hat{\theta} - \mathbb{E}[\hat{\theta}]]}_0 \\
&= \text{Var}(\hat{\theta}) + B(\hat{\theta}, \theta)^2
\end{aligned}$$

□

1.5 CONSISTANCE

Propriété asymptotique. On ne considère que des estimateurs consistants.

DEFINITION 1.10 (CONSISTANCE) – Soit (X_1, \dots, X_n) i.i.d. de loi P_θ . Soit $\hat{\theta}_n = h(X_1, \dots, X_n)$. $\hat{\theta}_n$ est un estimateur consistant (ou convergent) de θ ssi :

$$\hat{\theta}_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \theta$$

REMARQUE 1.11 – $\hat{\theta}_n$ est fortement consistant ssi $\hat{\theta}_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \theta$. ◇

1.5.1 EXEMPLE : RETOUR AU MODÈLE DE POISSON

$$\Theta = \mathbb{R}_+^*, \hat{\theta}_n = \bar{X}.$$

- On peut invoquer la Loi des Grands Nombres (LGN) : $\bar{X} \xrightarrow{\mathbb{P}} \mathbb{E}[X_i] = \theta$.
- Via le risque quadratique :

$$R(\hat{\theta}_n, \theta) = \text{Var}(\bar{X}) = \frac{\theta}{n} \xrightarrow[n \rightarrow +\infty]{} 0$$

D'après l'inégalité de Bienaymé-Tchebychev :

$$P(|\hat{\theta}_n - \theta| > \varepsilon) \leq \frac{\mathbb{E}[(\hat{\theta}_n - \theta)^2]}{\varepsilon^2} = \frac{R(\hat{\theta}_n, \theta)}{\varepsilon^2} \rightarrow 0$$

1.5.2 MÉTHODE « PLUG-IN »

Soit (X_1, \dots, X_n) i.i.d. $\text{Poisson}(\theta)$. On veut estimer $\beta = P(X_i = 0) = e^{-\theta}$.

$$\hat{\beta} = e^{-\hat{\theta}} = e^{-\bar{X}}$$

$\hat{\beta}$ est consistant pour estimer β .

LEMME 1.12 (LEMME DE L'APPLICATION CONTINUE) – Si $Z_n \xrightarrow{\mathbb{P}} Z$, alors $h(Z_n) \xrightarrow{\mathbb{P}} h(Z)$ pour toute fonction continue h . \diamond

Estimateurs

§2

2.1 CADRE PARAMÉTRIQUE

2.1.1 MODÈLE STATISTIQUE PARAMÉTRIQUE

On dispose d'une observation (X_1, \dots, X_n) , un échantillon de variable aléatoire i.i.d (indépendantes, identiquement distribuées) de loi commune P appartenant à une famille de lois de probabilités paramétrée $\{P_{\theta, \theta \in \Theta \subset \mathbb{R}^p}\}$.

REMARQUE 2.1 – Si $\Theta \subset$ espace de dimension infinie \rightarrow modèle non-paramétré. \diamond

Estimer P c'est estimer $\theta \in \mathbb{R}^p$.

EXEMPLE 2.2 – Bernoulli (θ) , Exp (θ) , $\mathcal{N}(\mu, \sigma^2)$, loi de densité $f_{\theta}(x) = \theta x^{\theta-1} \mathbf{1}_{x \in [0,1]}$ \diamond

NOTATION 2.3 – $E_{\theta_n}[h(X_1, \dots, X_n)]$, $\Theta[h(X_1, \dots, X_n)]$
Loi de $(X_1, \dots, X_n) \rightarrow P_{\theta}^{\otimes n}$ \diamond

DEFINITION 2.4 (ESTIMATEUR) –

$$\hat{\theta} = \hat{\theta}_n = h(X_1, \dots, X_n)$$

DEFINITION 2.5 (QUALITÉ) –

- Risque

$$R(\hat{\theta}, \theta) = E_{\theta}[(\hat{\theta} - \theta)^2]$$

- Consistance

$$\hat{\theta}_n \xrightarrow[n \rightarrow +\infty]{P} \theta$$

DEFINITION 2.6 (MODÈLE IDENTIFIABLE) –

$$\theta \rightarrow P_{\theta} \text{ injective}$$

2.2 MÉTHODE DES MOMENTS

DEFINITION 2.7 — On appelle *moment théorique* de la loi de X_i d'ordre k :

$$\mu_k = E[X_i^k], \quad k \geq 1$$

DEFINITION 2.8 — On appelle *moment empirique* de la loi des X_i d'ordre k :

$$\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n X_i^k$$

Par la loi des grands nombres $\hat{\mu}_k \xrightarrow[n \rightarrow +\infty]{P} \mu_k$.

La méthode des moments: si on peut écrire θ ou $g(\theta)$ paramètre d'intérêt comme une fonction des k premiers moments théoriques.

$$\theta = \mathcal{L}(\mu_1, \dots, \mu_k)$$

alors l'estimateur

$$\hat{\theta} = \mathcal{L}(\hat{\mu}_1, \dots, \mu_k)$$

est obtenu par la méthode.

EXEMPLE 2.9 (DES CALCULS DES ESTIMATEURS EN UTILISANT LA MÉTHODE DES MOMENTS) —

- $X_i \sim \text{Bernoulli}(\theta)$ à valeurs 0-1,

$$\theta = P(X_i = 1) = E[X_i] \rightarrow \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

- $X_i \sim \text{Exp}(\theta)$, $f_\theta(x) = \theta e^{-\theta x} 1_{x \geq 0}$, $E[X] = \frac{1}{\theta} \Leftrightarrow \theta = \frac{1}{\mu_1}$, par la méthode des moments,

$$\hat{\theta} = \frac{1}{\hat{\mu}_1} = \frac{1}{\bar{X}}$$

$$\begin{aligned} \Theta(X_i) = \frac{1}{\theta^2} &\Leftrightarrow \theta^2 = \frac{1}{E[X_i^2] - E[X_i]^2} \\ &\Leftrightarrow \theta = \frac{1}{\sqrt{\mu_2 - \mu_1^2}} \\ &\Rightarrow \hat{\theta}_2 = \frac{1}{\sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X})^2}} \end{aligned}$$

- X_1, \dots, X_n i.i.d. de la loi P_θ de densité

$$f_\theta(x) = \theta x^{\theta-1} 1_{x \in [0,1]}$$

$$E_\theta[X_i] = \theta \int_0^1 x^\theta dx = \frac{\theta}{\theta+1}$$

Méthode des moments:

$$\begin{aligned}
(\theta + 1)\mu_1 = \theta &\iff \theta(1 - \mu_1) = \mu_1 \iff \theta = \frac{E[X_i]}{1 - E[X_i]} \\
&\implies \hat{\theta}_M = \frac{\bar{X}}{1 - \bar{X}}, P_\theta(\bar{X} = 1) = P_\theta(X_1 = X_2 = \dots = X_n = 1) = 0
\end{aligned}$$

◇

2.3 RENDU SUR LE L.A.C.

(L.A.C = lemme des applications continues) $(X_n)_{n \geq 1}$ suite de variables aléatoires. Si X_n converge vers X , que peut-on dire de $g(X_n)_{n \geq 1}$? Si g continue, LAC.

- si $X_n \xrightarrow{P} X$ alors $g(X_n) \xrightarrow{P} g(X)$
- si $X_n \xrightarrow{\mathcal{L}} X$ alors $g(X_n) \xrightarrow{\mathcal{L}} g(X)$

REMARQUE 2.10 (CONDITION SUFFISANTE) –

$$D_g = \{\text{points de discontinuité de } g\}$$

si $P(X \in D_g) = 0$, le LAC est vrai.

◇

EXEMPLE 2.11 –

$$g(x) = \frac{x}{1-x}$$

- LGN: $\bar{X} \xrightarrow{P} E[X]$
- LAC: $g(\bar{X}) = \hat{\theta}_n \xrightarrow{P}_{n \rightarrow +\infty} g(E[X]) = \theta$

◇

LAC pour des couples de suites de variables aléatoires:

- si $(X_n, Y_n) \xrightarrow{P} (X, Y)$, alors $g(X_n, Y_n) \xrightarrow{P} g(X, Y)$, si $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ ou \mathbb{R}^2 continue
- si $(X_n, Y_n) \xrightarrow{\mathcal{L}} (X, Y)$, alors $g(X_n, Y_n) \xrightarrow{\mathcal{L}} g(X, Y)$

EXEMPLE 2.12 –

$$\hat{\theta}_2 = \frac{1}{\sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X})^2}} \quad \text{consistant?}$$

LGN:

- $\bar{X} \xrightarrow{P} \mu_1$
- $\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} \mu_2$

donc

$$\left(\begin{array}{c} \bar{X} \\ \frac{1}{n} \sum_{i=1}^n X_i^2 \end{array} \right) \xrightarrow{P} \left(\begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right)$$

$g(x, y) = \frac{1}{\sqrt{y-x^2}} \implies \hat{\theta}_2^M$ constant de θ , g continue sauf en $\{(x, y) \in \mathbb{R}^2, y = x^2\}$ de mesure nulle.

Mais c'est faux pour une convergence en loi.

◇

PROPOSITION 2.13 (CONVERGENCE DE COUPLES) –

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{P} \begin{pmatrix} X \\ Y \end{pmatrix} \text{ ssi } \begin{cases} X_n \xrightarrow{P} X \\ Y_n \xrightarrow{P} Y \end{cases}$$

Preuve.

- \Rightarrow alors LAC $g(x, y) = x$ continue donc $X_n \rightarrow X$ et $Y_n \rightarrow Y$
- \Leftarrow convergence du couple?

$$\forall \varepsilon > 0, P(|X_n - X| + |Y_n - Y| > \varepsilon) \leq \underbrace{P(|X_n - X| > \frac{\varepsilon}{2})}_{\rightarrow 0} + \underbrace{P(|Y_n - Y| > \frac{\varepsilon}{2})}_{\rightarrow 0}$$

Cette réciproque est fautive pour la convergence en loi!

□

2.3.1 VARIANCE EMPIRIQUE

Si la X_i admettent une espérance μ et une variance σ^2 , on appelle variance empirique

$$\begin{aligned} \hat{\sigma}_n^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 + \frac{1}{n} \sum_{i=1}^n \bar{X}^2 - \frac{2}{n} \sum_{i=1}^n X_i \bar{X} \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 + \bar{X}^2 - 2\bar{X}\bar{X} = \tilde{\sigma}^2 \end{aligned}$$

estimateur des moments:

$$\sigma^2 = E[X_i^2] - E[X_i]^2$$

On remplace les moments théoriques par les moments empiriques

$$\rightarrow \tilde{\sigma}^{2M} = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X})^2$$

Consistance: $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X})^2$,

$$\begin{cases} \bar{X} \xrightarrow{P} E[X] \\ \frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} E[X^2] \end{cases} \xrightarrow{\text{cv en proba}} \begin{pmatrix} \bar{X} \\ \frac{1}{n} \sum_{i=1}^n X_i^2 \end{pmatrix} \xrightarrow{\text{LAC}} \hat{\sigma}^2 \text{ consistant de } \text{Var}(X) = E[X^2] - E[X]^2$$

EXEMPLE 2.14 –

- calculer le biais de $\hat{\sigma}_n^2$
- calculer le risque de $\hat{\sigma}_n^2$

◇

2.4 MÉTHODE DE MAXIMUM DE VRAISEMBLANCE

2.4.1 MODÈLE DONNÉ

$(P_\theta)_{\theta \in \Theta}$ est donné s'il existe une mesure μ (positive σ définie $\rightarrow X_i$ à valeurs dans E , $E = \cup E_n$ avec $\mu(E_n)$ finie) telle que $\forall \theta, P_\theta$ admet une densité par rapport à μ .

2.4.2 EN PRATIQUE

- soit E au plus dénombrable: $\mu =$ mesure de comptage. Si $\exists, \{a_1, a_2, \dots\}$ tq $\sum_{k \geq 1} P_\theta(X_i = a_k) = 1$, alors $\mu = \sum_{k \geq 1} \delta_{a_k}$ avec $\delta_a(\{a\}) = 1$ mesure de dirac.

EXEMPLE 2.15 – Bernoulli (θ) , $X_i = 1$, probas $\theta \rightarrow \mu = \delta_0 + \delta_1$ On écrira

$$f_\theta(x) = \underbrace{P_\theta(\{x\})}_{=1-\theta} - P_\theta(X_i = x) \text{ avec } x \in \{a_1, a_2, \dots\}$$

◇

- soit $E = \mathbb{R}^p$, alors f_θ est la densité usuelle
 f_θ densité de P_θ

DEFINITION 2.16 – On appelle vraisemblance de l'échantillon (X_1, \dots, X_n) la fonction

$$\theta \rightarrow L_n(\theta) = \prod_{i=1}^n f_\theta(X_i) \text{ (variable aléatoire)}$$

DEFINITION 2.17 – Un estimateur du max de vraisemblance $\hat{\theta}_{MV}$ est définie par:

$$\forall \theta \in \Theta, L_n(\theta) \leq L_n(\hat{\theta})$$

On travaille souvent avec la **log-vraisemblance**

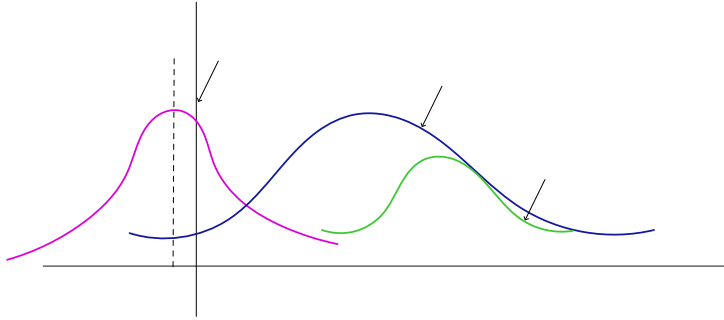
$$\log L_n(\theta) = \sum_{i=1}^n \ln f_\theta(X_i) \text{ somme de variables aléatoires}$$

$$\log L_n(\hat{\theta}) = \sup_{\theta \in \Theta} \log L_n(\theta)$$

REMARQUE 2.18 – $\hat{\theta}$ est une variable aléatoire

$$f_\theta(x) \quad \theta_1 \quad \theta_2 \quad \theta_3$$

x



◇

EXEMPLE 2.19 –

- Bernoulli(θ), $f_\theta(x) = \theta^x(1-\theta)^{1-x}$, X_i à valeurs 0-1

$$L_n(\theta) = \prod_{i=1}^n \theta^{X_i} (1-\theta)^{1-X_i} = \theta^{\sum_{i=1}^n X_i} (1-\theta)^{n-\sum_{i=1}^n X_i}$$

$$\log L_n(\theta) = \left(\sum_{i=1}^n X_i \right) \ln \theta + \left(n - \sum_{i=1}^n X_i \right) \ln(1-\theta)$$

$$(\log L_n)'(\theta) = \frac{\sum_{i=1}^n X_i}{\theta} - \frac{n - \sum_{i=1}^n X_i}{1-\theta} = \frac{\sum_{i=1}^n X_i - n\theta}{\theta(1-\theta)} (\bar{X} - \theta)$$

Equation de vraisemblance:

$$\begin{aligned} (\log L_n)'(\theta) = 0 &\Leftrightarrow (1-\theta) \sum_{i=1}^n X_i = \left(n - \sum_{i=1}^n X_i \right) \theta \\ &\Leftrightarrow \sum_{i=1}^n X_i = n\theta \Rightarrow \theta = \frac{\sum_{i=1}^n X_i}{n} \end{aligned}$$

le point critique, est-il un maximum?

La dérivée change de signe en $\bar{X} \rightarrow$ on a bien un max $\rightarrow \hat{\theta}^{MV} = \bar{X}$

Condition du 2nd ordre, si $(\log L_n)''(\theta) < 0$ pour tout $\theta \Rightarrow \log L_n$ est concave \Rightarrow max global

$$(\log L_n)''(\theta) = -\frac{\sum_{i=1}^n X_i}{\theta^2} - \frac{n - \sum_{i=1}^n X_i}{(1-\theta)^2} < 0, \forall \theta$$

◇

Information de Fisher, efficacité

§3

Soit $(P_\theta)_{\theta \in \Theta}$, $\Theta \subset \mathbb{R}^p$ (identifiable, donnée). On note f_θ densité de P_θ

$$\text{Supp } f_\theta = \{x \in E, f_{\theta(x)} > 0\}$$

Étant donné (X_1, \dots, X_n) , i.i.d. de loi P_θ et $\theta \mapsto L(\theta) = \prod_{i=1}^n f_{\theta(X_i)}$ la vraisemblance de l'échantillon. Sur $\text{Supp } f_\theta$ on peut calculer

$$\log L_n(\theta) = \sum_{i=1}^n \log f_{\theta(X_i)}$$

$$\hat{\theta} = \text{argmax}_{\theta \in \Theta} \log L_n(\theta)$$

PROPOSITION 3.1 – Si $\hat{\theta}$ EMV¹ de θ , $g(\hat{\theta})$ est un EMV de $g(\theta)$

Objectif: que peut-on avoir de « mieux » comme estimateur? → modèle régulier

3.1 MODÈLE RÉGULIER

DEFINITION 3.2 – Le modèle $(P_\theta)_{\theta \in \Theta}$ est dit régulier si

1. Θ est un ouvert et $\theta \mapsto f_{\theta(x)}$ est C^1
2. $\text{Supp } f_\theta$ ne dépend pas de θ : $S = \{x, f_{\theta(x)} > 0\}$
3. Pour tout θ , l'application

$$x \mapsto \frac{\frac{\partial f_\theta}{\partial \theta}(x)}{f_{\theta(x)}} \mathbb{1}_{f_{\theta(x)} > 0}$$

est intégrable (L, μ) et l'intégrale

$$I(\theta) = \int_S \frac{\frac{\partial f_\theta}{\partial \theta}(x)}{f_{\theta(x)}} \mathbb{1}_{f_{\theta(x)} > 0} dx$$

est continue sur Θ .

NOTATION 3.3 – On note la dérivée de $f_{\theta(x)}$ par rapport à θ : $\frac{\partial f_\theta}{\partial \theta}(x)$ La quantité $I(\theta)$ est appelée **Information de Fisher du modèle**. ◇

EXEMPLE 3.4 –

- $f_{\theta(x)} = \theta e^{-x\theta}$ densité par rapport à $\mu(dx) = \mathbb{1}_{x \geq 0} dx$

¹EMV = Estimateur de Maximum de Vraisemblance

$\theta \mapsto \theta e^{-x\theta}$ est C^∞ sur $\Theta =]0, +\infty[$, $\text{Supp } f_\theta = \mathbb{R}_+$

$$\frac{\partial f_\theta}{\partial \theta}(x) = (1 - x\theta)e^{-x\theta}$$

$$\frac{(1 - x\theta)^2 (e^{-x\theta})^2}{\theta e^{-x\theta}} = \frac{(1 - x\theta)^2}{\theta} e^{-x\theta}$$

$$\begin{aligned} I(\theta) &= \int_\theta^\infty \frac{(1 - x\theta)^2}{\theta^2} \theta e^{-x\theta} dx \\ &= \frac{1}{\theta^2} E_\theta(1 - X\theta)^2 \\ &= \frac{1}{\theta^2} [1 - 2\theta E(X) + \theta^2 E(X^2)] = \frac{1}{\theta^2} \end{aligned}$$

continue sur $]0, +\infty[$

◇

EXEMPLE 3.5 – Bernoulli(θ), $x = 0, 1$, $f_{\theta(0)} = 1 - \theta$, $f_{\theta(1)} = \theta$, densité par rapport à $\delta_0 + \delta_1$
Pour tout $x \in \{0, 1\}$, $\theta \mapsto f_{\theta(x)}$ est C^1

$$\frac{\left(\frac{\partial f_{\theta(0)}}{\partial \theta}\right)^2}{f_{\theta(0)}} = \frac{1}{1 - \theta}$$

$$\frac{\left(\frac{\partial f_{\theta(1)}}{\partial \theta}\right)^2}{f_{\theta(1)}} = \frac{1}{\theta} \Rightarrow I(\theta) = \frac{1}{1 - \theta} + \frac{1}{\theta} = \frac{1}{\theta(1 - \theta)}$$

continue sur $]0, 1[$

◇

EXEMPLE 3.6 – $f_{\theta(x)} = \frac{1}{\theta} \mathbb{1}_{[0, \theta]}(x) = \frac{1}{\theta} \mathbb{1}_{[x, +\infty[}(\theta)$ modèle non régulier

◇

3.2 SCORE ET INFORMATION DE FISHER

(X_1, \dots, X_n) i.i.d de loi de P_θ , f_θ

DEFINITION 3.7 – On appelle *score* ou *vecteur de score* la dérivée de la log vraisemblance
 $\frac{\partial}{\partial \theta} \log L_{n(\theta)} = S_{n(\theta)} = \sum_{i=1}^n \frac{\partial}{\partial \theta} \log f_\theta(X_i)$

EXEMPLE 3.8 – $X_i \sim \mathcal{E}(\theta)$, $L_{n(\theta)} = \theta^n e^{-\theta \sum_i X_i}$, $\log L_n(\theta) = n \log \theta - \theta \sum_i X_i$, donc $S - n(\theta) = \frac{n}{\theta} - \sum_{i=1}^n X - i$

◇

REMARQUE 3.9 –

$$E(S_n(\theta)) = E\left[n\left(\frac{1}{\theta} - \frac{\sum X_i}{n}\right)\right]$$

◇

Hyp supplémentaire de régularité: (H) pour tout estimateur $h(X)$ et tout θ , les intégrales suivantes existent et sont égales:

$$\frac{\partial}{\partial \theta} \int_S h(x) f_\theta(x) dx = \int_S h(x) \frac{\partial f_\theta}{\partial \theta}(x) dx$$

REMARQUE 3.10 – condition d'application du thm de dérivation de Lebesgue.

$$h \sup_{\theta \in V_\theta} \left| \frac{\partial f_\theta}{\partial \theta}(x) \right| \in L_1(\mu)$$

◇

PROPOSITION 3.11 – Sous (H), le score est centré (P_θ), $n = 1$

$$E_\theta \left[\frac{\partial}{\partial \theta} \log L_1(\theta) \right] = \int_S \frac{\partial}{\partial \theta} \log f_\theta(x) dx = \int_S \frac{\frac{\partial f_\theta}{\partial \theta}(x)}{f_\theta(x)} f_\theta(x) dx = \int_S \frac{\partial f_\theta}{\partial \theta}(x) dx \stackrel{(H)}{=} \frac{\partial}{\partial \theta} \int_S f_\theta(x) dx \stackrel{=1}{=} 0$$

DEFINITION 3.12 – L'information de Fisher associé à (X_1, \dots, X_n)

$$I_n(\theta) \stackrel{\text{def}}{=} E_\theta \left[\left(\frac{\partial}{\partial \theta} \log L_n(\theta) \right)^2 \right] \stackrel{\text{cor. de la prop 1}}{=} \text{Var}_\theta \left[\frac{\partial \log L_n(\theta)}{\partial \theta} \right]$$

$$(*) E_\theta \left[\frac{\partial}{\partial \theta} \log f_\theta(X_1) \right]^2 = \int_S \left(\frac{\frac{\partial f_\theta}{\partial \theta}(x)}{f_\theta(x)} \right)^2 f_\theta(x) dx = \int_S \frac{\left(\frac{\partial f_\theta}{\partial \theta}(x) \right)^2}{f_\theta(x)} = \text{"expression de la définition 1"}$$

EXEMPLE 3.13 – $(X_1, \dots, X_n) \sim \mathcal{E}(\theta)$, $\frac{\partial}{\partial \theta} \log L_n(\theta) = \frac{n}{\theta} - \sum_{i=1}^n X_i$

$$I_n(\theta) = E \left(\left(\frac{n}{\theta} - \sum X_i \right)^2 \right) = n^2 E \left[\left(\frac{1}{\theta} - \frac{\sum X_i}{n} \right)^2 \right] = n^2 \text{Var}(\bar{X}) = n^2 \frac{1}{n} \frac{1}{\theta^2} = \frac{n}{\theta^2}$$

◇

PROPOSITION 3.14 –

$$I_n(\theta) = nI(\theta)$$

en effet,

$$\begin{aligned}
 I_n(\theta) &= \text{Var} \left(\frac{\partial}{\partial \theta} \log L_n(\theta) \right) = \text{Var} \left(\sum_{i=1}^n \frac{\partial}{\partial \theta} \log f_\theta(X_i) \right) \stackrel{\text{independance}}{=} \sum_{i=1}^n \text{Var} \left(\frac{\partial}{\partial \theta} \log f_\theta(X_i) \right) = \\
 &= n \underbrace{\text{Var} \left(\frac{\partial}{\partial \theta} \log f_\theta(X_1) \right)} = nI(\theta)
 \end{aligned}$$

EXEMPLE 3.15 — (X_1, \dots, X_n) i.i.d $\mathcal{P}(\theta)$, $f_\theta(x) = e^{-\theta} \frac{\theta^x}{x!}$

$$\begin{aligned}
 \log L_n(\theta) &= -n\theta + \left(\sum X_i \right) \log \theta - \log \prod_{i=1}^n X_i! \\
 \frac{\partial}{\partial \theta} \log L_n(\theta) &= -n + \frac{\sum X_i}{\theta} \Rightarrow I_n(\theta) = \text{Var} \left(\frac{\sum X_i}{\theta} \right) = \frac{1}{\theta^2} n\theta = \frac{n}{\theta}
 \end{aligned}$$

◇

3.3 INFORMATION DE FISHER ET DERIVÉE SECONDE

PROPOSITION 3.16 — En ajoutant que $\theta \mapsto f_\theta(x)$ est C^2 et que (H) vrai pour $\frac{\theta^2}{\theta^2}$ alors l'info de Fisher s'écrit encore

$$I_n(\theta) = -E_\theta \left[\frac{\partial^2 \log L_n(\theta)}{\partial \theta^2} \right]$$

si $\hat{\theta}$ EMV, $I_n(\hat{\theta}) > 0$

$n = 1$

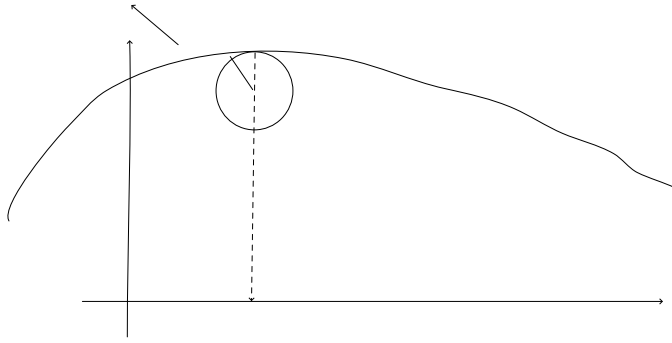
$$\frac{\partial^2}{\partial \theta^2} \log f_\theta(x) = \frac{\left(\frac{\partial^2 f_\theta(x)}{\partial \theta^2} \right)^2}{f_\theta(x)} - \frac{\left(\frac{\partial f_\theta(x)}{\partial \theta} \right)^2}{f_\theta^2(x)}$$

$$E \left[\frac{\partial^2}{\partial \theta^2} \log f_\theta(X_1) \right] = \int_S \frac{\frac{\partial^2 f_\theta(x)}{\partial \theta^2}}{f_\theta(x)} f_\theta(x) dx - \underbrace{\int_S \frac{\left(\frac{\partial f_\theta(x)}{\partial \theta} \right)^2}{f_\theta^2(x)} dx}_{I(\theta)}$$

$\log L(\theta)$

$\hat{\theta}$

θ



Si courbe très « piqué » en l'EMV (i.e. info. Fisher est grande) alors l'EMV est localisé de façon précise

3.4 INÉGALITÉ DE CRAMER - RAO

Soit $g(\theta)$ le paramètre d'intérêt où $g : \Theta \rightarrow \mathbb{R}$

PROPOSITION 3.17 – *Sous les hypothèses d'un modèle régulier, si pour tout θ , $I(\theta) > 0$, alors pour tout estimateur $T = T(X_1, \dots, X_n)$ sans biais, $E_\theta T^2 < +\infty$, on a*

$$\forall \theta \in \Theta, \underbrace{\text{Var}_\theta(T)} \geq \frac{(g'(\theta))^2}{I_n(\theta)} = \frac{(g'(\theta))^2}{nI(\theta)}$$

Preuve.

$$\begin{aligned} \forall \theta \quad E_{\theta(T)} &= g(\theta) \\ \Rightarrow \frac{\partial}{\partial \theta} E_\theta(T) &= g'(\theta) \\ T=T(X_1) \Leftrightarrow \frac{\partial}{\partial \theta} \int_S T(x) f_\theta(x) dx &= g'(\theta) \\ (H) \Leftrightarrow \int_S T(x) \frac{\frac{\partial}{\partial \theta} f_\theta(x)}{f_\theta(x)} f_\theta(x) dx &= g'(\theta) \\ \Leftrightarrow \int_S (T(x) - g(\theta)) \frac{\frac{\partial}{\partial \theta} f_\theta(x)}{f_\theta(x)} f_\theta(x) dx &= g'(\theta) \end{aligned}$$

□

Inégalité de Cauchy-Schwarz pour $\langle h_1, h_2 \rangle = \int h_1(x) h_2(x) f_\theta(x) dx$ avec $h_1(X)$ et $h_2(X)$ centrées

$$\left(\left\langle T(X) - g(\theta), \frac{\frac{\partial}{\partial \theta} f_\theta(x)}{f_\theta(x)} \right\rangle_\theta \right)^2 = (g'(\theta))^2 \underbrace{\int (T(x) - g(\theta))^2 f_\theta(x) dx}_{=\text{Var}_\theta(T)} \times \underbrace{\int \left(\frac{\frac{\partial}{\partial \theta} f_\theta(x)}{f_\theta(x)} \right)^2 f_\theta(x) dx}_{=I(\theta)}$$

DEFINITION 3.18 – *Si T réalise l'égalité, alors T est dit efficace.*

Étude asymptotique des estimateurs

§4

Dans un modèle paramétrique régulier, si $\hat{\theta}_n$ estimateur de θ , alors

$$\text{Var}(\hat{\theta}_n) \geq \frac{1}{I_n(\theta)} = \frac{1}{nI(\theta)}$$

si $\text{Var}(\hat{\theta}_n) = \frac{1}{nI(\theta)}$ sans biais, $\hat{\theta}_n$ est efficace efficace

Asymptotique: $n \rightarrow +\infty$,

$$n \text{Var}(\hat{\theta}_n) \xrightarrow{n \rightarrow +\infty} \frac{1}{I(\theta)}$$

4.1 CONVERGENCES

$(X_n)_{n \geq 0}$ suite de variables aléatoires réelles (\mathbb{R}^d)

- convergence en loi: $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ ssi $P(X_n \leq x) \rightarrow P(X \leq x)$ en tout point de continuité de x .

LEMME 4.1 (LEMME DE PORTMANTEAU) – Caractérisations équivalentes:

- Pour toute fonction continue bornée h ,

$$E[h(X_n)] \rightarrow E[h(X)]$$

\Rightarrow la convergence en loi est stable par passage aux fonctions continues (LAC) MAIS il est en général faux que si $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} Y$ alors $\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} X \\ Y \end{pmatrix}$

Cela est vrai dans 3 cas:

1. si $\begin{cases} \forall n, X_n \text{ et } Y_n \text{ sont indépendantes} \\ X \text{ et } Y \text{ sont indépendantes} \end{cases}$ alors $\begin{cases} \text{convergence en loi de } X_n \text{ et } Y_n \\ \text{convergence en loi du couple } \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \end{cases}$

2.

$$\text{si } \begin{cases} X_n \xrightarrow{P} X \\ Y_n \xrightarrow{P} Y \end{cases} \Rightarrow \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{P} \begin{pmatrix} X \\ Y \end{pmatrix} \Rightarrow \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} X \\ Y \end{pmatrix}$$

3. (Lemme de Slutsky) (le plus important)

$$\text{si } \begin{cases} X_n \xrightarrow{\mathcal{L}} X \\ Y_n \xrightarrow{\mathcal{L}} c \end{cases} \text{ alors } \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} X \\ c \end{pmatrix}$$

en appliquant le LAX,

$$\begin{aligned}
h(x, y) &= x + y & X_n + Y_n &\xrightarrow{\mathcal{L}} X + c \\
&= xy & X_n Y_n &\xrightarrow{\mathcal{L}} \xrightarrow{\mathcal{L}} cX \\
&= \frac{x}{y} & \frac{X_n}{Y_n} &\xrightarrow{\mathcal{L}} \frac{X}{c}
\end{aligned}$$

◇

4.2 CONSISTANCE DES ESTIMATEURS

DEFINITION 4.2 — $\hat{\theta}_n$ asymptotiquement sans biais si et seulement si

$$\text{Biais}(\hat{\theta}_n, \theta) = E[\hat{\theta}_n] - \theta \xrightarrow{n \rightarrow +\infty} 0$$

REMARQUE 4.3 — La convergence en proba n'implique pas la convergence des espérances.

Si $X_n \xrightarrow{P} X$, $|X_n| \leq Y \in L'$, alors par convergence dominée $X_n \rightarrow X$ dans L_1

◇

EXEMPLE 4.4 — $\hat{\tau}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum X_i^2 - (\bar{X})^2$ estimateur des moments de $\tau^2 = E[X^2] - (E[X])^2$

$$\text{Biais}(\hat{\tau}_n^2, \tau^2) = -\frac{1}{n} \tau^2 \text{ asymptotiquement sans biais}$$

Consistance de $\hat{\tau}_n^2$?

Outils pour montrer la consistance:

- LGN
- si $R(\hat{\theta}_n, \theta) \rightarrow 0$ alors $\hat{\theta}_n$ consistant car convergence $L^2 \Rightarrow$ convergence en probas
- revenir à la définition de convergence en probas

- si (X_i) sont i.i.d., alors (X_i^2) est i.i.d.

$$E[X_i^2] < +\infty$$

- LGN: $\frac{1}{n} \sum_i X_i^2 \xrightarrow{P} E[X^2] = \tau^2 + \mu^2$
- $\bar{X} \xrightarrow{P} \mu$ (LGN), LAC avec $h(x) = x^2$: $(\bar{X})^2 \xrightarrow{P} \mu^2$
- Donc $\left(\frac{\frac{1}{n} \sum_{i=1}^n X_i^2}{\frac{1}{n} \sum X_i} \right) \xrightarrow{P} \left(\frac{\tau^2 + \mu^2}{\mu} \right)$
- LAC $h(x, y) = x - y^2$

Donc $\frac{1}{n} \sum X_i^2 - (\bar{X})^2 \xrightarrow{P} \tau^2 + \mu^2 - \mu^2 = \tau^2$

◇

4.3 NORMALITÉ ASYMPTOTIQUE

$\hat{\theta}_n$ pour θ .

→ Question: quelle est la vitesse de convergence de $\hat{\theta}_n$ vers θ ?

(X_1, \dots, X_n) i.i.d., d'espérance θ , $\hat{\theta} = \bar{X}$ de variance $\tau^2(\theta)$
TLC $\sqrt{n}(\bar{X} - \theta) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, \tau^2(\theta))$ quelle que soit la loi des X_i

DEFINITION 4.5 — $\hat{\theta}_n$ est un estimateur asymptotiquement normal si et seulement si

- vitesse de convergence en \sqrt{n}
- convergence en loi
- loi limite est normale

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, \tau^2(\theta))$$

EXEMPLE 4.6 — $\hat{\tau}_n^2$ est-elle asymptotiquement normale ?

(X_1, \dots, X_n) i.i.d. d'espérance μ , de variance τ^2

$$\hat{\tau}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 + (\bar{X} - \mu)^2 + \underbrace{\frac{2}{n} \sum_{i=1}^n (X_i - \mu)(\mu - \bar{X})}_{=2(\mu - \bar{X})(\bar{X} - \mu)}$$

- TLC: si (X_i) i.i.d., alors les $(X_i - \mu)^2$ sont i.i.d. d'espérance τ^2 ,

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - \tau^2 \right) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, u_4 - \tau^4)$$

$$\text{Var}(X_i - \mu)^2 = E[(X_i - \mu)^4] - \mu^4 = \mu_4 - \tau^4$$

- TLC: $\sqrt{n}(\bar{X} - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2)$
- $\sqrt{n}(\hat{\tau}_n^2 - \tau^2) = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - \tau^2 \right) - \underbrace{\sqrt{n}(\bar{X} - \mu)^2}_{\substack{\sqrt{n}(\bar{X} - \mu) \times (\bar{X} - \mu) \\ \xrightarrow{\mathcal{L}, \mathcal{N}(0, \tau^2)} \quad \xrightarrow{\mathcal{L}, P} 0}}$

$$\left. \begin{array}{l} \bar{X} - \mu \xrightarrow{\mathcal{L}} 0 \\ \sqrt{n}(\bar{X} - \mu) \xrightarrow{\mathcal{L}} U \sim \mathcal{N}(0, 1) \end{array} \right\} \text{lemme de Slutsky} \implies \sqrt{n}(\bar{X} - \mu)^2 \xrightarrow{\mathcal{L}} \frac{\mathcal{L}}{P} 0$$

$$\sqrt{n}(\hat{\tau}_n^2 - \tau^2) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z + 0$$

Donc $\hat{\tau}_n^2$ est un estimateur asymptotiquement normal ◇

REMARQUE 4.7 — $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \tau^2) \iff \frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\tau} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$ Application du lemme de Slutsky: si $\hat{\tau}^2$ est un estimateur consistant de τ^2 , alors on a encore

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

◇

Preuve.

$$\frac{\sqrt{n}(\hat{\theta} - \theta)}{\hat{\tau}} = \underbrace{\left(\frac{\sqrt{n}(\hat{\theta} - \theta)}{\tau} \right)}_{\xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0,1)} \times \underbrace{\left(\frac{\tau}{\hat{\tau}} \right)}_{\xrightarrow{P} 1}$$

$\xrightarrow{\mathcal{L}} 1 \times Z$ par Slutsky et consistance de $\hat{\tau}$

□

4.4 δ -MÉTHODE

$\hat{\theta}$ estimateur asymptotiquement normal: quelle est la loi asymptotique de $g(\theta)$?

LEMME 4.8 (MÉTHODE DÉLTA) – Soit Z_n suite de variables aléatoires réelles t.q.

$$\sqrt{n}(Z_n - \mu) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, \tau^2)$$

Soit g une fonction dérivable, $g'(\mu) \neq 0$. Sous ces hypothèses, on a

$$\sqrt{n}[g(Z_n) - g(\mu)] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \tilde{Z} \sim \mathcal{N}(0, (g'(\mu))^2 \tau^2)$$

$$g(x) = g(\mu) + g'(\mu)(x - \mu) + (x - \mu)R(x - \mu) \text{ où } R(y) \xrightarrow[y \rightarrow 0]{} 0$$

$$\sqrt{n}(g(Z_n) - g(\mu)) = \underbrace{g'(\mu)\sqrt{n}(Z_n - \mu)}_{\xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, \tau^2)} + \underbrace{(\sqrt{n})(Z_n - \mu)R(Z_n - \mu)}_{\xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2) \quad \xrightarrow{P} 0?}$$

$\rightarrow \mathcal{N}(0, (g'(\mu))^2 \tau^2)$

A-t-on $Z_n \xrightarrow{P} \mu$?

$$\begin{aligned} P(|X_n - \mu| > \varepsilon) &= P\left(\frac{\sqrt{n}|Z_n - \mu|}{\tau} > \frac{\sqrt{n}\varepsilon}{\tau}\right) \\ &= P\left(\frac{\sqrt{n}(Z_n - \mu)}{\tau} > \frac{\sqrt{n}\varepsilon}{\tau}\right) + P\left(\frac{\sqrt{n}(Z_n - \mu)}{\tau} < -\frac{\sqrt{n}\varepsilon}{\tau}\right) \\ &\sim 1 - \Phi_n\left(\frac{\sqrt{n}\varepsilon}{\tau}\right) + \Phi_n\left(-\frac{\sqrt{n}\varepsilon}{\tau}\right) = 2\left(1 - \Phi\left(\frac{\sqrt{n}\varepsilon}{\tau}\right)\right) \end{aligned}$$

◇

Fonction de répartition empirique

§5

(X_1, \dots, X_n) échantillon i.i.d. à valeurs réelles de loi F inconnue.

$$\forall x \in \mathbb{R}, F(x) = P(X_1 \leq x) = E[\mathbb{1}_{X_1 \leq x}]$$

DEFINITION 5.1 – La fonction de répartition empirique associée à (X_1, \dots, X_n) est définie par:

$$\begin{aligned} \hat{F}_n : \mathbb{R} &\longrightarrow [0, 1] \\ x &\mapsto \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{X_i \leq x} \end{aligned}$$

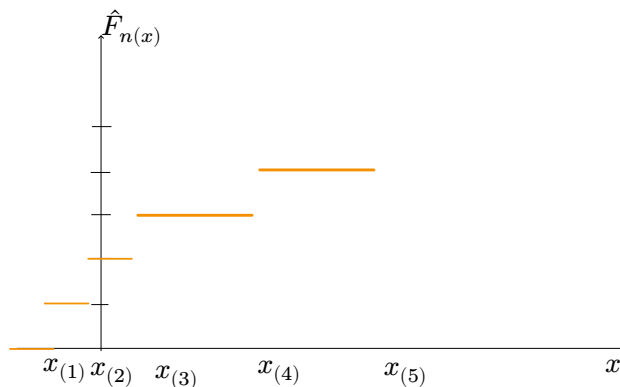
$\forall x \in \mathbb{R}, \hat{F}_n(x)$ est une variable aléatoire, estimateur de $F(x)$.

DEFINITION 5.2 – Loi empirique $P_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ est une loi discrète uniforme sur $\{X_1, \dots, X_n\}$.

Représentation graphique

Conditionnelement $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$

$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ valeurs ordonnées



PROPOSITION 5.3 (PROPRIÉTÉS IMMÉDIATS) –

- $n\hat{F}_n(x) = \sum_{i=1}^n \mathbb{1}_{X_i \leq x}$ suit la loi binomiale $(n, F(x))$

- $R(\hat{F}_n(x), F(x)) = 0 + \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n \mathbb{1}_{X_i \leq x}\right) \underset{\text{indep}}{=} \frac{1}{n} F(x)(1 - F(x)) \xrightarrow{n \rightarrow +\infty} 0$ donc $\forall x \in \mathbb{R}, \hat{F}_n(x) \xrightarrow{P} F(x)$
- ou bien LGN: $\hat{F}_n(x)$ estimateur consistant de $F(x)$.
- On a un résultat de convergence uniforme :

$$\sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| \xrightarrow{n \rightarrow +\infty} 0 \quad (\text{Théorème de Glivenko-Cantelli})$$

- $\hat{F}_n(x)$ est-il asymptotiquement normal?

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{X_i \leq x}$$

TLC: les X_i sont i.i.d., donc les $\{\mathbb{1}_{X_i \leq x} = Y_i\}$ sont i.i.d.

$$\forall x, F(x) \in]0, 1[, \quad \sqrt{n}(\hat{F}_n(x) - F(x)) \xrightarrow{n \rightarrow +\infty} \mathcal{N}(0, F(x)(1 - F(x)))$$

$$\Leftrightarrow \frac{\hat{F}_n(x) - F(x)}{\sqrt{\frac{F(x)(1 - F(x))}{n}}} \xrightarrow{n \rightarrow +\infty} \mathcal{N}(0, 1)$$

5.1 ESTIMATION EMPIRIQUE

« plug-in » ou méthode de substitution, paramètre d'intérêt $\theta = c(F)$, la méthode empirique définit $\hat{\theta}$, estimateur empirique en remplaçant F par $\hat{F}_n \rightarrow \hat{\theta}_n = c(\hat{F}_n)$.

EXEMPLE 5.4 — $\theta = E_F(X) \rightarrow \hat{\theta}_n = E_{\hat{F}_n}(X) = \sum_{i=1}^n X_i \times \frac{1}{n} = \bar{X}$ si X_i distinctes

$$\theta = \text{Var}_F(X) \rightarrow \hat{\theta}_n = \text{Var}_{\hat{F}_n}(X) = \frac{1}{n} \sum (X_i - \bar{X})^2$$

◇

5.2 INVERSE GÉNÉRALISÉ

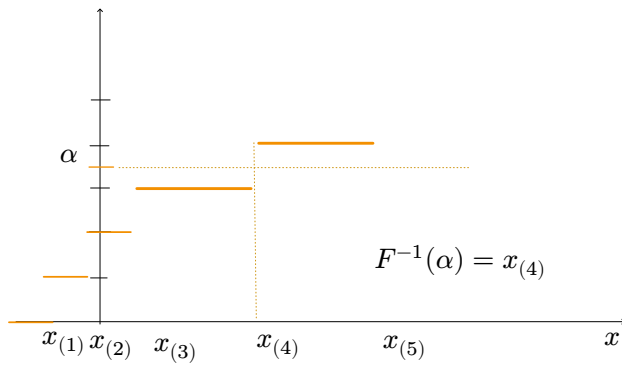
DEFINITION 5.5 — On définit l'inverse généralisé de F par:

$$F^{-1} : [0, 1] \rightarrow \mathbb{R}$$

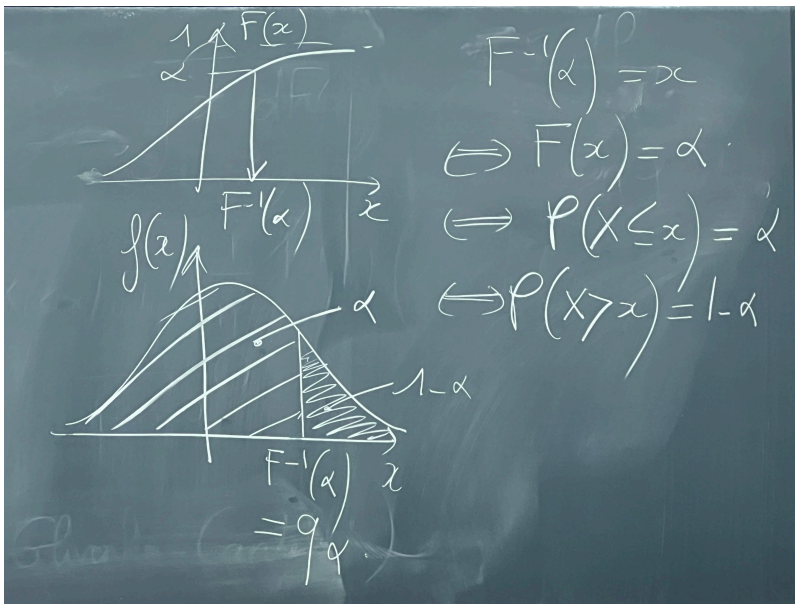
$$\forall \alpha \in [0, 1], F^{-1}(\alpha) = \inf\{x \in \mathbb{R}, F(x) \geq \alpha\}$$

Si F est strictement croissante, $\inf x$ tel que $F(x) \geq a \Leftrightarrow x \geq F^{-1}(\alpha)$, si F est la fonction d'une loi discrète.

²Thm de Glivenko-Cantelli: https://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8me_de_Glivenko-Cantelli

**EXEMPLE 5.6 –**

$$F^{-1}(\alpha) = x \Leftrightarrow F(x) = \alpha \Leftrightarrow P(X \leq x) = \alpha \Leftrightarrow P(X > x) = 1 - \alpha$$



◇

Vocab:

- F^{-1} s'appelle aussi la fonction quantile
- $F^{-1}(\alpha)$ = quantile d'ordre α , de la loi F
- $F^{-1}(\frac{1}{4})$ = 1er quantile
- $F^{-1}(\frac{1}{2})$ = médiane
- $F^{-1}(\frac{3}{4})$ = 3eme quantile

LEMME 5.7 – U variable aléatoire sur $[0, 1]$, F f.r., alors $F^{-1}(U)$ est une variable aléatoire de loi F

◇

- Si F bijective:

$$P(F^{-1}(U) \leq x) \underset{F \text{ bijective}}{\equiv} P(U \leq F(x)) \underset{\text{car } P(U \leq x) = x \text{ sur } [0,1]}{\equiv} F(x)$$

- Si F discrète: F^{-1} inverse généralisé: $F^{-1}(y) \leq x \Leftrightarrow y \leq F(x)$

5.3 QUANTILE EMPIRIQUE

DEFINITION 5.8 – On définit le quantile empirique (sample quantile) d'ordre α , comme étant le quantile de \hat{F}_n :

$$\hat{q}_{n,\alpha} = \hat{F}_n^{-1}(\alpha) = \inf\{x, \hat{F}_n(x) \geq \alpha\}$$

PROPOSITION 5.9 –

- On peut montrer que $\hat{q}_{n,\alpha} = X_{([n\alpha])}$ où $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ est l'échantillon ordonné des $(X_i)_{1 \leq i < n}$

$$[u] = \text{le plus petit entier } \geq u$$

EXEMPLE 5.10 – $\alpha = \frac{1}{2}$, $[\frac{n}{2}]$,

$$\begin{cases} \text{si } n = 2k & \text{medianne} = \hat{q}_{n,\frac{1}{2}} = X_{(k)} \\ \text{si } n = 2k + 1 & \text{medianne} = \hat{q}_{n,\frac{1}{2}} = X_{(k+1)} \end{cases}$$

- *Consistance*

si $\alpha \in]0, 1[$, si F est strictement croissante au voisinage de α

◇

Intervalles de confiance

§6

6.1 DÉFINITIONS

(X_1, \dots, X_n) i.i.d. de loi $P \in \{P_\theta, \theta \in \Theta \subset \mathbb{R}^p\}$, on s'intéresse à $\theta \in \mathbb{R}$ ou $g(\theta) : \mathbb{R}^p \rightarrow \mathbb{R}$.

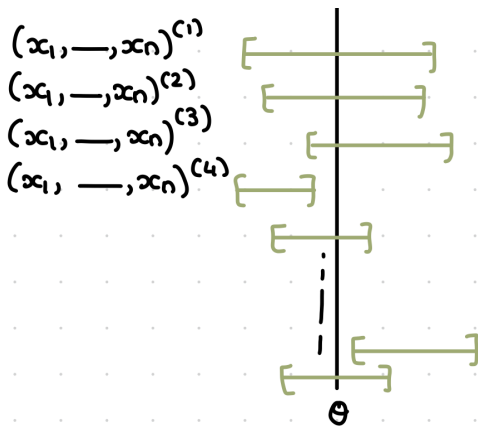
Un intervalle de confiance pour θ , de niveau de confiance $1 - \alpha, \alpha \in]0, 1[$ est un intervalle dont les bornes sont aléatoires, fonctions de l'échantillon et ne dépend PAS des paramètres inconnus du modèle et tel que

$$P([B \text{ inf}(X_1, \dots, X_n); B \text{ sup}(X_1, \dots, X_n)] \ni \theta) \geq 1 - \alpha$$

3

- Un IC est calculable à partir des données
- si l'inégalité est une égalité = niveau de confiance est exact.
- si on a $P(\theta \in [B \text{ inf}, B \text{ sup}]) \xrightarrow{n \rightarrow +\infty} 1 - \alpha$, niveau est asymptotique.
- en général $\alpha = 1\%, 5\%$

6.2 INTERPRÉTATION



$IC = [B \text{ inf}(X_1, \dots, X_n), B \text{ sup}(X_1, \dots, X_n)]$ formule mathématique qui garantit le niveau $1 - \alpha$. On observe $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$, une réalisation de l'échantillon aléatoire. On calcule $IC = [2.3; 5.1]$ de niveau de confiance 95% ($\alpha = 5\%$).

En moyenne, sur 100 intervalles calculés (avec la même formule), il y a 5 intervalles qui ne contiennent pas θ .

$$P(\theta \in [B \text{ inf}, B \text{ sup}]) = 1 - \alpha$$

$$P(\theta \in [2.3, 5.1]) = 95\% \text{ car } \theta \text{ un nombre}$$

6.3 MÉTHODE PIVOTALE

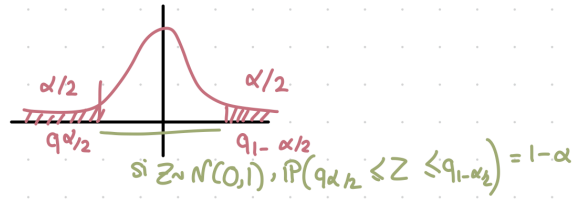
(X_1, \dots, X_n) i.i.d. d'espérance $\theta \in \mathbb{R}$, de variance $\sigma^2(\theta)$. Soit $\hat{\theta}$, asymptotiquement normal:

$$\begin{aligned} \sqrt{n}(\hat{\theta}_n - \theta) &\xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(\theta)) \\ \Leftrightarrow \frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\sigma(\theta)} &\xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1) \end{aligned}$$

Par définition des quantiles gaussiens, $q_\alpha = \Phi^{-1}(\alpha)$ où Φ f.r. de $\mathcal{N}(0, 1)$

³ $B \text{ inf}$ pour borne inférieure et $B \text{ sup}$ pour borne supérieure

$$P\left(q_{\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma(\theta)} \leq q_{1-\frac{\alpha}{2}}\right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha \quad (1)$$



- **pivot ou statistique pivotale** = $\frac{\sqrt{n}(\hat{\theta} - \theta)}{\hat{\sigma}}$ statistique centrée réduite issue de $\hat{\theta}$, où $\sigma^2(\theta)$ estimé par $\hat{\sigma}^2$, consistant pour estimer $\sigma^2(\theta)$.

Si c'est le cas,

$$\underbrace{\frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma(\theta)}}_{\xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \text{ } \hat{\theta} \text{ as. normal}} \times \underbrace{\frac{\sigma^2(\theta)}{\hat{\sigma}^2}}_{\xrightarrow{P} 1 \text{ } \text{estimateur consistant}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0,1) \text{ par lemme de Slutsky}$$

- on en déduit

$$P\left(q_{\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\hat{\theta} - \theta)}{\hat{\sigma}}(\theta) \leq q_{1-\frac{\alpha}{2}}\right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha$$

$$P\left(\hat{\theta} - \frac{1}{\sqrt{n}}\hat{\sigma}q_{1-\frac{\alpha}{2}} \leq \theta \leq \hat{\theta} - \frac{1}{\sqrt{n}}\hat{\sigma}q_{\frac{\alpha}{2}}\right) \rightarrow 1 - \alpha$$

REMARQUE 6.1 (POURQUOI $\frac{\alpha}{2}$?) – On peut observer que les quantiles dans Équation 1 sont d'ordre $\frac{\alpha}{2}$ et $1 - \frac{\alpha}{2}$. Pour comprendre pourquoi, il suffit d'effectuer un calcul simple. D'abord, on note $\frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma(\theta)} =: Z \sim \mathcal{N}(0,1)$.

$$P\left(q_{\frac{\alpha}{2}} \leq Z \leq q_{1-\frac{\alpha}{2}}\right) = P\left(Z \leq q_{1-\frac{\alpha}{2}}\right) - P\left(Z \leq q_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$$

◇

Compléments (avant partiel)

§7

1. Retour sur normalité asymptotique
2. Exemple
3. Pivot asymptotique
4. Exemple 2

7.1 PROPRIÉTÉS ASYMPTOTIQUES D'UNE SUITE D'ESTIMATEURS $(\hat{\theta}_n)_{n \geq 1}$

- Consistance $\hat{\theta}_n \xrightarrow{P} \theta$
- Normalité asymptotique, s'il existe $\sigma^2 > 0$

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

De façon générale, s'il existe $v_n \xrightarrow[n \rightarrow +\infty]{} +\infty$

$$v_n(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} Y$$

On dit que $\hat{\theta}_n$ converge à la vitesse $\frac{1}{v_n}$

REMARQUE 7.1 – Si $\hat{\theta}_n$ asymptotiquement normal $\Rightarrow \hat{\theta}_n$ consistant

$$\hat{\theta}_n - \theta = \underbrace{\frac{1}{\sqrt{n}}}_{\rightarrow 0} \underbrace{\sqrt{n}(\hat{\theta}_n - \theta)}_{\xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)} \xrightarrow[\text{Slutsky}]{\mathcal{L} \text{ ou } P} 0$$

$$U_n = \frac{1}{\sqrt{n}} \rightarrow 0$$

◇

δ -méthode

$$\begin{aligned} \sqrt{n}(X_n - 1) &\xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \\ \sqrt{n}(X_n - 1) &\stackrel{\mathcal{L}}{\approx} Z \sim \mathcal{N}(0, 1) \\ X_n &\stackrel{\mathcal{L}}{\approx} 1 + \frac{1}{\sqrt{n}}Z \end{aligned}$$

Si g dérivable en 1,

$$\begin{aligned} g(1+h) &= g(1) + hg(1) \\ g(X_n) &\approx g(1) + \frac{1}{\sqrt{n}}g'(1)Z \\ \sqrt{n}(g(X_n) - g(1)) &\approx g'(1)Z \end{aligned}$$

δ -méthode

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

g dérivable en θ

$$g(x) = g(\theta) + g'(\theta)[(x - \theta) + r(x)] \text{ où } r(x) \xrightarrow{x \rightarrow 0} 0$$

$$\hat{\theta}_n \xrightarrow{P} \theta \text{ donc (LAC) } r(\hat{\theta}_n) \rightarrow r(\theta) = 0$$

$$g(\hat{\theta}_n) = g(\theta) + (\hat{\theta}_n - \theta)[g'(\theta) + r(\hat{\theta}_n)]$$

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) = \underbrace{\sqrt{n}(\hat{\theta}_n - \theta)}_{\xrightarrow{\mathcal{L}} Z} \left[\underbrace{g'(\theta) + r(\hat{\theta}_n)}_{\xrightarrow{P} g'(\theta)} \right] \stackrel{\text{Slutsky}}{\Rightarrow} \sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{\mathcal{L}} g'(\theta)Z \sim \mathcal{N}(0, (g'(\theta))^2)$$

EXEMPLE 7.2 — X_1, \dots, X_n de loi de densité $f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}, x \geq 0, \mu = E[X_i] > 0$

μ estimé par $\hat{\mu} = \bar{X}$ efficace? $\log L_n(\mu) = -n \log \mu - \frac{1}{\mu} \sum_{i=1}^n X_i$

$$\text{Var}(\hat{\mu}) = \frac{1}{n^2} \text{Var}\left(\sum_i X_i\right) \stackrel{\text{indép}}{=} \frac{1}{n^2} \sum_i \text{Var}(X_i) \stackrel{\text{i.i.d.}}{=} \frac{1}{n} \text{Var}(X_i) = \frac{\mu^2}{n}, E[\hat{\mu}] = \mu$$

$$\frac{\partial}{\partial \mu} (\log L_n)(\mu) = -\frac{n}{\mu} + \frac{1}{\mu^2} \sum (X_i)$$

$$\begin{aligned} I_{n(\mu)} &= \text{Var}\left(-\frac{n}{\mu} + \frac{1}{\mu^2} \sum X_i\right) \\ &= \frac{1}{\mu^4} \text{Var}\left(\sum X_i\right) \\ &= \frac{n}{\mu^4} \text{Var}(X_i) \end{aligned}$$

$$I_{n(\mu)} = \frac{n}{\mu^2}$$

$\hat{\mu}$ sans biais et $\text{Var}(\hat{\mu}) = \frac{1}{I_{n(\mu)}}$. Donc $\hat{\mu}$ est efficace.

TLC: $\sqrt{n}(\hat{\mu}_n - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mu^2) \Rightarrow \text{Var}(\hat{\mu}_n) = \frac{\mu^2}{n} = \text{variance de la loi gaussienne asymptotique}$
 $\frac{\sqrt{n}(\hat{\mu}_n - \mu)}{\mu}$ a pour loi asymptotique $\mathcal{N}(0, 1)$

- autre paramétrisation: (X_1, \dots, X_n) i.i.d. $f(x) = \theta e^{-\theta x}, x \geq 0$

$$EX_i = \frac{1}{\theta}, \text{Var } X_i = \frac{1}{\theta^2}$$

$$\begin{aligned}\log L_n(\theta) &= n \log \theta - \theta \sum_{i=1}^n X_i \\ \frac{\partial}{\partial \theta}(\log L_n)(\theta) &= \frac{n}{\theta} - \sum X_i \hookrightarrow \hat{\theta}^{\text{MV}} = \frac{1}{\bar{X}} \\ \hookrightarrow I_{n(\theta)} &= \text{Var}\left(\frac{n}{\theta} - \sum X_i\right) = \text{Var}\left(\sum X_i\right) = \frac{n}{\theta^2}\end{aligned}$$

REMARQUE 7.3 – cd TD1: $n\bar{X} \sim \Gamma(n, \theta)$

$$E\left[\frac{1}{n\bar{X}}\right] = \frac{\theta}{n-1} \text{ et } \text{Var}\left(\frac{1}{(n\bar{X})^2}\right) = \frac{\theta^2}{(n-1)(n-2)}$$

◇

$$E\left[\frac{1}{\bar{X}}\right] = n \frac{\theta}{n-1}$$

$$\hookrightarrow \tilde{\theta} = \frac{n-1}{n} \hat{\theta} \text{ non biaisé}$$

$$\begin{aligned}\text{Var}(\tilde{\theta}) &= \left(\frac{n-1}{n}\right)^2 \text{Var}\left(\frac{1}{\bar{X}}\right) = \frac{(n-1)^2}{n^2} \left[E\left[\frac{1}{(\bar{X})^2}\right] - \left(E\left[\frac{1}{\bar{X}}\right]\right)^2 \right] \\ &= \frac{\cancel{(n-1)^2}}{n^2} \times \frac{n^2 \theta^2}{\cancel{(n-1)}(n-2)} - \frac{(n-1)^2}{n^2} \frac{n^2}{(n-1)^2} \theta^2 \\ &= \theta^2 \frac{n-1}{n-2} - \theta^2 = \frac{\theta^2}{n-2} \underset{\text{BCR}}{\geq} \frac{1}{I_{n(\theta)}} \text{ non efficace}\end{aligned}$$

$$\sqrt{n}\left(\bar{X} - \frac{1}{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\theta^2}\right)$$

$\hat{\theta}$ est asymptotiquement efficace

\bar{X} asymptotiquement normal (TLC). $g(x) = \frac{1}{x}$ sur $]0, +\infty[$, $g'(x) = -\frac{1}{x^2} \neq 0$, méthode delta:

$$\sqrt{n}\left(\frac{1}{\bar{X}} - \theta\right) \xrightarrow{\mathcal{L}} \underbrace{g'\left(\frac{1}{\theta}\right)}_{=\theta^2} \mathcal{N}\left(0, \frac{1}{\theta^2}\right) = \mathcal{N}\left(0, \frac{\theta^4}{\theta^2} = \theta^2\right)$$

◇

7.2 PIVOT (ASYMPTOTIQUE) OU STATISTIQUE PIVOTALE

DEFINITION 7.4 – Statistique dont la loi ne dépend pas de paramètres inconnus

EXEMPLE 7.5 – X_1, \dots, X_n i.i.d. Bernoulli(θ) avec $\theta \in]0, 1[$:

$$\begin{aligned} \sqrt{n}(\bar{X} - \theta) &\xrightarrow[\text{TLC}]{\mathcal{L}} \mathcal{N}(0, \theta(1 - \theta)) \\ \Leftrightarrow \underbrace{\sqrt{n} \frac{\bar{X} - \theta}{\sqrt{\theta(1 - \theta)}}}_{\text{pivot ou stat. pivotale}} &\xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \end{aligned}$$

méthode pivotale pour IC: On estime $\sqrt{\theta(1 - \theta)}$ par $\sqrt{\hat{\theta}(1 - \hat{\theta})}$ « plug-in » par le LAC $g(x) = \sqrt{x(1 - x)}$ avec $x \in]0, 1[$, $\sqrt{\hat{\theta}(1 - \hat{\theta})}$ estimateur consistant de $\sqrt{\theta(1 - \theta)}$

$$\sqrt{n} \frac{\hat{\theta} - \theta}{\sqrt{\hat{\theta}(1 - \hat{\theta})}} = \underbrace{\sqrt{n} \frac{\hat{\theta} - \theta}{\sqrt{\theta(1 - \theta)}}}_{\xrightarrow[\text{TLC}]{\mathcal{L}} \mathcal{N}(0, 1)} \times \underbrace{\frac{\sqrt{\theta(1 - \theta)}}{\sqrt{\hat{\theta}(1 - \hat{\theta})}}}_{\xrightarrow[\text{consistant}]{P} 1}$$

◇

EXEMPLE 7.6 — (X_1, \dots, X_n) de densité $\theta > 0$. $f_\theta(x) = \frac{3}{\theta} x^2 \exp(-\frac{x^3}{\theta}) \mathbb{1}_{x \geq 0}$

EMV?

$$\log L_n(\theta) = n(\log^3 - \log \theta) + \sum_{i=1}^n \log(X_i^2) - \frac{1}{\theta} \sum_{i=1}^n X_i^3$$

$$(\log L_n)'(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum X_i^3 \Rightarrow \hat{\theta} = \frac{\sum X_i^3}{n}$$

$$(\log L_n)''(\theta) = \frac{n}{\theta^2} - \frac{2}{\theta^3} \sum X_i^3; (\log L_n)''(\hat{\theta}) = \frac{n}{\hat{\theta}^2} - \frac{2}{\hat{\theta}^3} n\hat{\theta} = -\frac{n}{\hat{\theta}^2} < 0 + \text{unicité}$$

\Rightarrow max global

TLC:

$$\begin{aligned} \sqrt{n}(\hat{\theta}_n - \theta) &\xrightarrow{\mathcal{L}} \mathcal{N}(0, \theta^2) \\ \Leftrightarrow \underbrace{\frac{\sqrt{n}(\hat{\theta} - \theta)}{\theta}}_{\text{pivot asymptotique}} &\xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \\ \Rightarrow \underbrace{\frac{\sqrt{n}(\hat{\theta} - \theta)}{\hat{\theta}}}_{\text{Slutsky}} &\xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \end{aligned}$$

$q_{\frac{\alpha}{2}}$ et $q_{1-\frac{\alpha}{2}}$ quantiles de $\mathcal{N}(0, 1)$

$$P\left(q_{\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\hat{\theta} - \theta)}{\hat{\theta}} \leq q_{1-\frac{\alpha}{2}}\right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha$$

$$P\left(q_{\frac{\alpha}{2}} \frac{\hat{\theta}}{\sqrt{n}} \leq \hat{\theta} - \theta \leq q_{1-\frac{\alpha}{2}} \frac{\hat{\theta}}{\sqrt{n}}\right) \rightarrow 1 - \alpha$$

$$P\left(\underbrace{\hat{\theta} - q_{1-\frac{\alpha}{2}} \frac{\hat{\theta}}{\sqrt{n}} \leq \theta \leq \hat{\theta} - q_{\frac{\alpha}{2}} \frac{\hat{\theta}}{\sqrt{n}}}_{\Rightarrow IC(\theta) \text{ de niveau asymptotique } (1-\alpha)}\right) \rightarrow 1 - \alpha$$

◇

Estimation dans les échantillons gaussiens

§8

1. Loi normale et lois dérivées
2. Loi des estimateurs empiriques
3. IC des paramètres
4. Exercice

8.1 TL&DR

(X_1, \dots, X_d) des variables aléatoires i.i.d. qui suivent $\mathcal{N}(0, 1)$ et $X \sim \mathcal{N}(0, 1)$,

$$Y = X_1^2 + \dots + X_d^2 \sim \chi^2(d)$$

$$\frac{X}{\sqrt{\frac{Y}{d}}} \sim \text{Student}(d)$$

8.2 LOI NORMALE ET LOIS DÉRIVÉES

DEFINITION 8.1 – Z est dite gaussienne (normale) centrée réduite si sa loi admet pour densité

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$$

On note $Z \sim \mathcal{N}(0, 1)$.

X est dite de loi normale de paramètres $\mu \in \mathbb{R}$ et $\sigma^2 > 0$ ssi

$$X = \mu + \sigma Z$$

notée

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

Autres caractérisations de la loi normale:

- par densité

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

- par la fonction génératrice des moments

$$M(t) = E[e^{tX}] = e^{t\mu + \frac{1}{2}\sigma^2 t^2}, \forall t \in \mathbb{R}$$

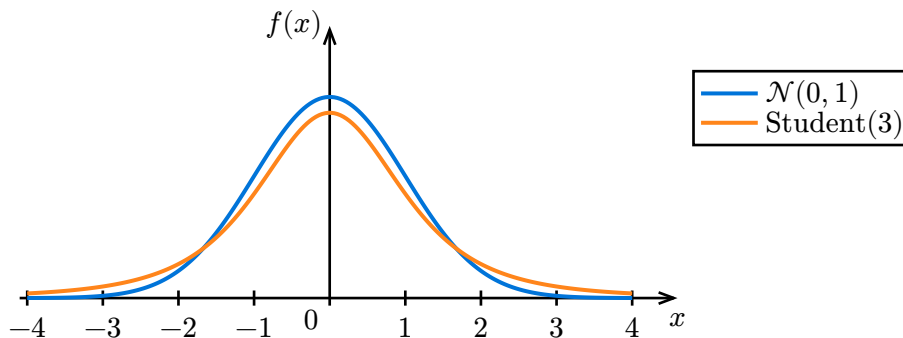
REMARQUE 8.2 –

- $\sigma^2 = 0 \rightarrow X = \mu$ presque sûrement

- si $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ et $\lambda \in \mathbb{R}$, alors $\lambda X_1 + X_2 \sim \mathcal{N}(\lambda\mu_1 + \mu_2, \lambda^2\sigma_1^2 + \sigma_2^2)$

◇

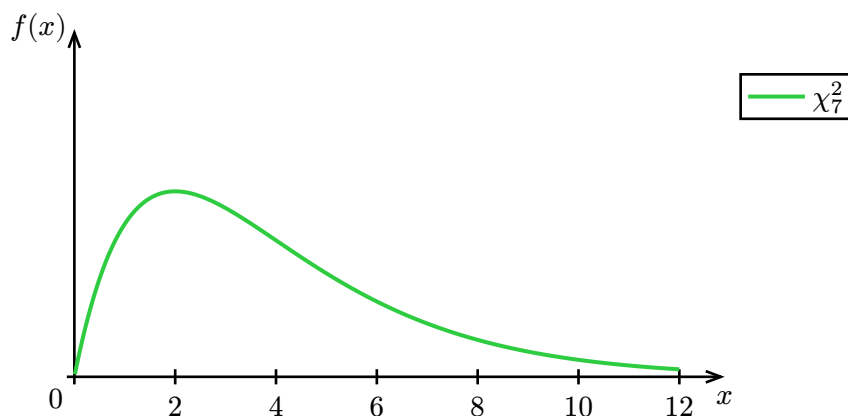
Moments centrés: densité symétrique par rapport à μ



moments centrés: $E[(X - \mu)^k]$

- tous les moments centrés d'ordre impaire sont nuls
- $\mu_{2k} = \frac{(2k)!}{2^k k!} \sigma^{2k}$
 - $E[(X - \mu)^4] = 3\sigma^4$
 - $\text{Var}(X) = E[(X - \mu)^2] = \sigma^2$

DEFINITION 8.3 — (X_1, \dots, X_d) échantillon *i.i.d.* $\mathcal{N}(0, 1)$. La loi de $X_1^2 + X_2^2 + \dots + X_d^2$ est appelée loi du χ^2 (chi 2) à d degrés de liberté (*ddl*) (*degrees of freedom (df)*).



COROLLAIRE 8.4 —

- si Y de loi $\chi^2(d)$, $E[Y] = d$, $\text{Var}(Y) = 2d$

$$\text{Var}(X_1^2 + \dots + X_d^2) \stackrel{\text{indep}}{=} d \underbrace{\text{Var}(X_i^2)}_{EX_i^4 - E[X_i^2]^2 = 3 - 1 = 2}$$

- support \mathbb{R}_+
- $M(t) = (1 - 2t)^{-\frac{d}{2}}$, $(t < \frac{1}{2})$

◇

DEFINITION 8.5 — si $X \sim \mathcal{N}(0, 1)$ et $Y \sim \chi^2(d)$ indépendantes, la loi de $Z = \frac{X}{\sqrt{\frac{Y}{d}}}$ est appelée loi de Student à d ddl.

REMARQUE 8.6 — si $d \rightarrow +\infty$, la loi de Student converge vers la loi $\mathcal{N}(0, 1)$

$$\frac{Y}{d} \stackrel{\mathcal{L}}{\underset{\text{def}}{=} } \frac{1}{d} \sum_{i=1}^d U_i^2 \text{ où } U_i \sim \mathcal{N}(0, 1) \text{ indép entre elles de } X$$

$$\xrightarrow[\text{LGN}]{P} E(U_i^2) = 1$$

donc (LAC)

$$g(x) = \sqrt{x} \frac{1}{\sqrt{\frac{Y}{d}}} \xrightarrow{P} 1$$

par le Lemme de Slutsky $Z \xrightarrow{\mathcal{L}} 1 \cdot X \sim \mathcal{N}(0, 1)$ ◇

On introduit (X_1, \dots, X_n) i.i.d. $\mathcal{N}(\mu, \sigma^2)$ où μ et σ^2 paramètres inconnus.

- $\hookrightarrow \mu = E[X_i] \rightsquigarrow \hat{\mu} = \bar{X}$
- $\hookrightarrow \sigma^2 = \text{Var}(X_i) \rightsquigarrow \hat{\sigma}^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$

Soit $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ non biaisé

8.3 LOI DES ESTIMATEURS EMPIRIQUES

THEOREME 8.7 (LOI DE $\hat{\mu}$ ET $\hat{\sigma}^2$) —

- \bar{X} et $\sum_{i=1}^n (X_i - \bar{X})^2$ sont des variables aléatoires indépendantes
- $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$
- $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1) \Rightarrow \frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1)$ et $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$
- $\frac{\bar{X} - \mu}{\frac{S_n}{\sqrt{n}}} \sim \text{Student}(n-1)$
- \bar{X} et $\overbrace{(\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X})}^T$ sont indépendantes

Preuve.

$$\begin{aligned}
M(u, t_1, \dots, t_n) &= E \left[e^{u\bar{X} + t_1(X_1 - \bar{X}) + \dots + t_n(X_n - \bar{X})} \right] \\
&= E \left[e^{\left(\frac{u}{n} + \frac{t_1 + t_2 + \dots + t_n}{n} \right) X_1} \dots e^{\left(\frac{u}{n} + \bar{t} \right) X_n} \right] \\
&= E \left[\prod_{i=1}^n e^{\left(\frac{u}{n} + t_i - \bar{t} \right) X_i} \right] \\
X_i \text{ indép} &= \prod_{i=1}^n \underbrace{E \left[e^{\left(\frac{u}{n} + t_i - \bar{t} \right) X_i} \right]}_{M(u_n + t_i - \bar{t})} \\
&= \prod_{i=1}^n e^{\mu \left(\frac{u}{n} + t_i - \bar{t} \right) + \frac{\sigma^2}{2} \left(u_n + t_i - \bar{t} \right)^2} \\
&= e^{\sum_{i=1}^n \mu \left(\frac{u}{n} + t_i - \bar{t} \right) + \frac{\sigma^2}{2} \left(u_n + t_i - \bar{t} \right)^2} \\
&= e^{\mu u + \mu \overbrace{\sum_i (t_i - \bar{t})}^0 + \frac{\sigma^2}{2} \sum_i \left(\frac{u^2}{n^2} + (t_i - \bar{t})^2 + 2 \frac{u}{n} (t_i - \bar{t}) \right)} \\
&= e^{\mu u + \frac{\sigma^2}{2} \left(\frac{u^2}{n} + \sum_i (t_i - \bar{t})^2 \right)} \\
&= \underbrace{e^{\mu u + \frac{\sigma^2 u^2}{2n}}}_{M_{\bar{X}}(u)} \underbrace{e^{\frac{\sigma^2}{2} \sum_i (t_i - \bar{t})^2}}_{M_T(t_1, \dots, t_n)}
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 &= \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{n}{\sigma^2} (\bar{X} - \mu)^2 + \frac{2}{\sigma^2} \sum_i (X_i - \bar{X}) (\bar{X} - \mu) \\
&= \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{n}{\sigma^2} (\bar{X} - \mu)^2 + \frac{2}{\sigma^2} (\bar{X} - \mu) \underbrace{\sum_i (X_i - \bar{X})}_{=0} \\
&= \underbrace{\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2}_{=\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n)} + \underbrace{\frac{n}{\sigma^2} (\bar{X} - \mu)^2}_{=\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \right)^2 \sim \chi^2(1)}
\end{aligned}$$

$$\text{par indép} \Rightarrow M_{\chi^2(n)}(t) = M_{\chi^2(1)}(t) M_{\chi^2(n-1)}(t) \Rightarrow M_{\chi^2(n)}(t) = \frac{(1-2t)^{-\frac{n}{2}}}{(1-2t)^{-\frac{1}{2}}} = (1-2t)^{-\frac{(n-1)}{2}}$$

qui caractérise la loi $\chi^2(n-1)$

$$\frac{\bar{X} - \mu}{\frac{S_n}{\sqrt{n}}} = \frac{\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}{\frac{S_n}{\frac{\sigma}{\sqrt{n}}} \times \frac{\sqrt{n}}{\sigma}} = \frac{\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{S_n^2}{\sigma^2}}}$$

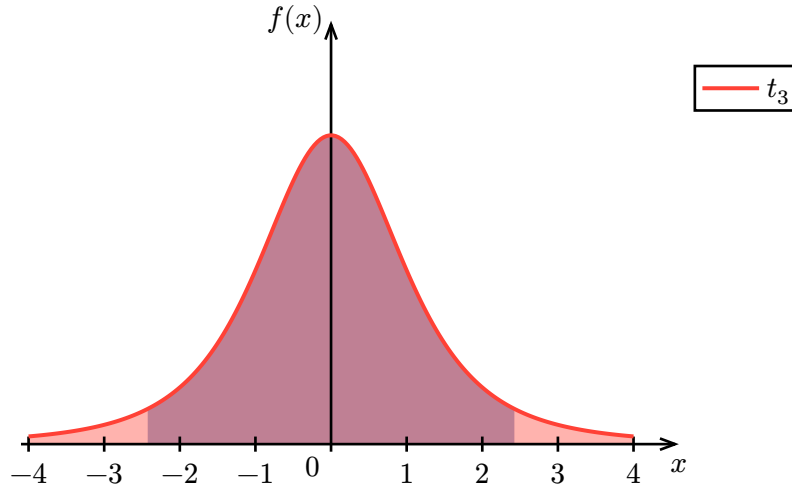
$$\frac{S_n^2}{\sigma^2} = \left(\frac{\sum (X_i - \bar{X})^2}{\sigma^2} \right) \sim \chi^2(n-1)$$

donc $\bar{X} + S_n^2$ indépendantes \Rightarrow Student($n-1$) □

def Student

8.4 IC DES PARAMÈTRES

Pivot. $\frac{\bar{X} - \mu}{\frac{S_n}{\sqrt{n}}} \underset{\text{loi exacte}}{\sim} \text{Student}(n-1)$



$$P\left(q_{\frac{\alpha}{2}} t(n-1) \leq \frac{\bar{X} - \mu}{\frac{S_n}{\sqrt{n}}} \leq q_{1-\frac{\alpha}{2}} t(n-1)\right)$$

$$\Leftrightarrow P\left(\bar{X} - \frac{S_n}{\sqrt{n}} q_{1-\frac{\alpha}{2}} t(n-1) \leq \mu \leq \bar{X} + \frac{S_n}{\sqrt{n}} q_{1-\frac{\alpha}{2}} t(n-1)\right) = 1 - \alpha$$

IC (σ^2), $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1)$

$$P\left(q_{\frac{\alpha}{2}} \chi^2(n-1) \leq \frac{n\hat{\sigma}^2}{\sigma^2} \leq q_{1-\frac{\alpha}{2}} \chi^2(n-1)\right) = 1 - \alpha$$

$$= P\left(\frac{n\hat{\sigma}^2}{q_{1-\frac{\alpha}{2}} \chi^2(n-1)} \leq \sigma^2 \leq \frac{n\hat{\sigma}^2}{q_{\frac{\alpha}{2}} \chi^2(n-1)}\right) = 1 - \alpha$$

$$\leadsto \text{IC} = \left[\frac{n\hat{\sigma}^2}{q_{1-\frac{\alpha}{2}} \chi^2(n-1)}, \frac{n\hat{\sigma}^2}{q_{\frac{\alpha}{2}} \chi^2(n-1)} \right]$$

REMARQUE 8.8 — $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1)$ et $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$

$$\frac{\bar{X} - \mu}{\frac{S_n}{\sqrt{n}}} \sim \text{Student}(n-1)$$

◇

8.5 EXERCICE

- Montrez que $(\hat{\mu}, \hat{\sigma}^2)$ sont les EMV de μ et σ^2
- $R(S_n^2, \sigma^2) > R(\hat{\sigma}_n^2, \sigma^2)$ où R représente une risque

Introduction aux tests statistiques

§9

9.1 EXEMPLE

9.1.1 CONTRÔLE DE QUALITÉ: INDUSTRIEL.

Produit des « pièces »

- ↪ de bonne qualité
- ↪ défectueuses

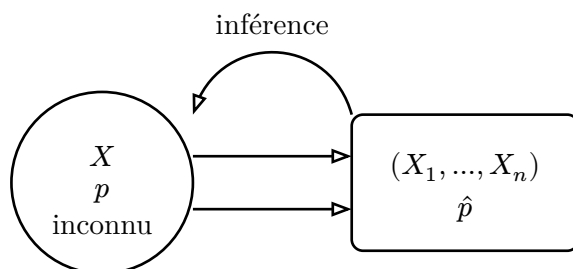
Pour l'industriel, on suppose acceptable une proportion de 20% de pièces défectueuses.

Pour contrôler: prélever « au hasard » n pièces, vérifiées ($p \leq 20\%$)

9.1.2 MODÉLISATION

$i^{\text{ème}}$ pièce $X_i = \begin{cases} 0 & \text{si bonne qualité} \\ 1 & \text{si défectueuse} \end{cases} \quad p = P(X_i = 1)$

↪ on prélève n pièces et on observe un échantillon (X_1, \dots, X_n) dont les valeurs observées sont (x_1, \dots, x_n) .



Que vaut p ? → on estime

- proportion empirique
- $X_i \underset{\text{indep}}{\sim} \text{Bernoulli}(p) \rightarrow \hat{p} = \bar{X}$

On observe $\bar{x} = 0.22$, $n = 100$

On procède avec un intervalle de confiance pour p . On définit $\hat{p} = \bar{X}$, TLC:

$$\frac{\bar{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

on estime l'écart-type par $\frac{\hat{p}(1-\hat{p})}{n}$ (consistant). Lemme de Slutsky:

LGN: $\bar{X} \xrightarrow{P} p$, LAC: $g(x) = \sqrt{\frac{x(1-x)}{n}}$

$$\frac{\bar{X} - p}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} = \frac{\bar{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \times \frac{\overbrace{\sqrt{\frac{p(1-p)}{n}}}^{P \rightarrow 1}}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} 1 \cdot \mathcal{N}(0, 1)$$

$$P \left(\text{qnorm}_{\frac{\alpha}{2}} \leq \frac{\bar{X} - p}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \leq \text{qnorm}_{1-\frac{\alpha}{2}} \right) \xrightarrow[n \rightarrow +\infty]{} 1 - \alpha$$

$$\Leftrightarrow P \left(\bar{X} - q_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \leq p \leq \bar{X} - q_{\frac{\alpha}{2}} \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \right) \xrightarrow[n \rightarrow +\infty]{} 1 - \alpha$$

IC(p) = $\bar{X} \pm \text{qnorm}_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$ de niveau asymptotique $1 - \alpha$.

ex: $\bar{x} = 0.22$, $\alpha = 5\%$, $n = 100$, IC = [0.14, 0.30]

Question: est-ce que $p \leq 0.2$ ou bien $p > 0.2$?

9.2 PRINCIPE D'UN TEST

$\Theta \subset]0, 1[$. On veut tester si $p \leq 0.2$ ou $p > 0.2$.

$\Theta = \underbrace{\Theta_0}_{]0, 0.2]} \cup \underbrace{\Theta_1}_{]0.2, 1[}$ sous-ensembles disjoints.

On teste $H_0: p \in \Theta_0, p \leq 0.2$ contre $H_1: p \in \Theta_1, p > 0.2$

Conclusion:

- Soit on conserve H_0 : ($p \leq 0.2$)
- Soit on rejette H_0 (on conclut $p > 0.2$)

DEFINITION 9.1 – Un test de H_0 contre H_1 est défini par la construction d'une région de rejet de H_0 , \mathcal{R}

- si $(X_1, \dots, X_n) \in \mathcal{R}$, on rejette H_0 (au profit de H_1)
- si $(X_1, \dots, X_n) \notin \mathcal{R}$, on conserve H_0

Souvent $\mathcal{R} = \{(X_1, \dots, X_n), T(X_1, \dots, X_n) > c\}$

- T : statistique de test (à valeur réelle)
- c : seuil du test

REMARQUE 9.2 – la décision d'un test est aléatoire (dépend de T aléatoire) ◇

Comment relier \mathcal{R} aux hypothèses testées ?

9.3 RISQUE D'ERREUR

DEFINITION 9.3 – Erreur de 1^{ère} espèce ou risque de type I est la fonction définie sur

$$\Theta_0 \longrightarrow [0, 1]$$

$$p \mapsto P_p((X_1, \dots, X_n) \in \mathcal{R}) = P_p(\text{on rejette } H_0)$$

Le test est dit de niveau α si

$$\sup_{p \in \Theta_0} P_p(\text{rejet de } H_0) \leq \alpha$$

R_q erreur de première espèce = $P(\text{rejet de } H_0 \text{ à tort})$

réalité / décision	H_0 vraie	H_1 vraie
H_0 vraie	ok	erreur de première espèce
H_1 vraie	erreur de seconde espèce	ok

DEFINITION 9.4 – L'erreur de seconde espèce est la fonction définie sur risque de type II

$$\Theta_1 \longrightarrow [0, 1]$$

$$\beta : p \mapsto P_p((X_1, \dots, X_n) \notin \mathcal{R}) = P_p(\text{on conserve } H_0)$$

REMARQUE 9.5 – erreur de seconde espèce est $P(\text{conserver } H_0 \text{ à tort})$ ◇

puissance du test: = 1-erreur 2^{de} espèce

$$\beta : p \in \Theta_1 \longrightarrow P_p((X_1, \dots, X_n) \in \mathcal{R})$$

Choix: les 2 erreurs ne peuvent pas être minimiser simultanément. En général α augmente quand β diminue.

Test: On choisit de contrôler l'erreur de 1^{ère} espèce (\Rightarrow l'erreur de seconde espèce est inconnue en général)

9.4 CONSTRUCTION D'UN TEST

Principe: déterminer \mathcal{R} tel que erreur de première espèce $\leq \alpha$ (si on a plusieurs tests, on choisira (point de vue théorique) celui dont l'erreur de seconde espèce est la plus petite (ou de puissance la plus grande)). Basé sur une dissymétrie de H_0 et H_1 dans la construction.

EXEMPLE 9.6 – $H_0 : p \leq 0.2$ contre $H_1 : p > 0.2$ ($\bar{x} = 0.22$)

- p inconnu donc on l'estime $\hat{p} = \bar{X}$
- idée: sous H_1 , \hat{p} prend de plus grandes valeurs que sous H_0

$\hookrightarrow \mathcal{R}$ du type $\hat{p} > c$ avec c tel que $P_p(\hat{p} > c) \leq \alpha$? (calcul? loi limite du paramètre p ?)

$$\hat{p} = \bar{X} \longrightarrow \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

a pour loi approché $\mathcal{N}(0, 1)$

$$P(\hat{p} > c) = P\left(\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} > \frac{c-p}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

On veut que

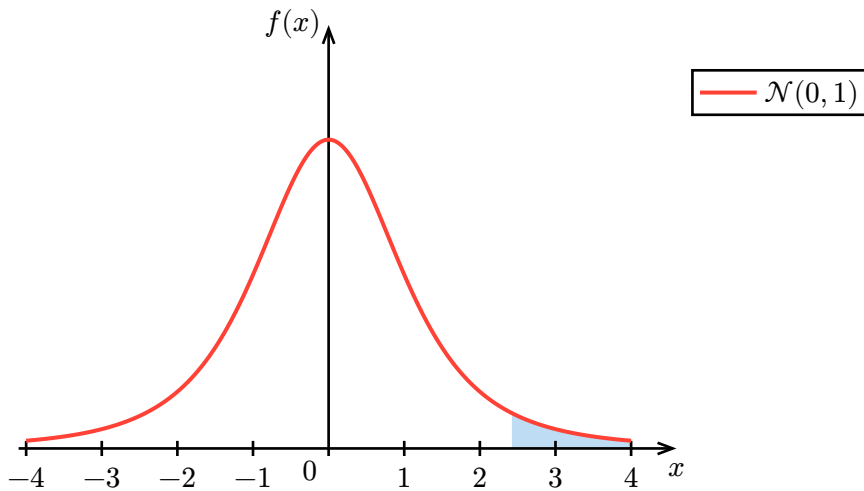
$$\sup_{\substack{p \in \Theta_0 \\ p \leq 0.2}} P\left(\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} > c\right) \leq \alpha$$

↳ le sup est atteint en $p = 0.2$

$$\mathcal{R} = \left\{ (X_1, \dots, X_n), \frac{\hat{p} - 0.2}{\sqrt{\frac{0.2(1-0.2)}{n}}} > c \right\}$$

Trouver c tel que $P((X_1, \dots, X_n) \in \mathcal{R}) \xrightarrow{n \rightarrow +\infty} \alpha$

$$P\left(\frac{\hat{p} - 0.2}{\sqrt{\frac{0.2(1-0.2)}{n}}} > c\right) \xrightarrow{n \rightarrow +\infty} \alpha \text{ ssi } c = \text{qnorm}_{1-\alpha}$$



• rejet de H_0 ssi

$$T = \underbrace{\frac{\hat{p} - 0.2}{\sqrt{\frac{0.2(1-0.2)}{n}}}}_{\text{statistique de test}} > \text{qnorm}_{1-\alpha}$$

A.N. $\alpha = 5\%$, $\text{qnorm}_{1-\alpha} = 1.645$, $n = 100$, $\hat{p} = \bar{x} = 0.22$

↳

$$T = \frac{\bar{x} - 0.2}{\sqrt{\frac{0.2(1-0.2)}{100}}} = \frac{0.02}{\sqrt{\frac{0.2(1-0.2)}{100}}} = \frac{0.2}{\sqrt{0.2 \cdot 0.8}} = \frac{0.2}{0.4} = \frac{1}{2} < 1.645$$

Conclusion: on conserve H_0 (on ne connaît pas le risque associé)

$$\begin{aligned}\text{rejet } H_0 &\Leftrightarrow \bar{x} > 0.2 + 1.645 \sqrt{\frac{0.2(1-0.2)}{100}} \\ &\Leftrightarrow \bar{x} > 0.266\end{aligned}$$

◇

Tests d'hypothèse (sur un paramètre)

§10

10.1 FORMALISME D'UN TEST

10.1.1 INTRODUCTION

DEFINITION 10.1 (TEST STATISTIQUE) – *Un test d'hypothèse est une fonction (mesurable) de l'échantillon (X_1, \dots, X_n) à valeurs dans $\{0, 1\}$.*

- H_0 est acceptée si $\varphi(X_1, \dots, X_n) = 0$
- H_0 est rejetée si $\varphi(X_1, \dots, X_n) = 1$

Le domaine $\{(X_1, \dots, X_n), \varphi(X_1, \dots, X_n) = 1\} =: \mathcal{R}$ est la région de rejet du test, \mathcal{R}^c est la région d'acceptation. On peut écrire: $\varphi(X_1, \dots, X_n) = \mathbb{1}_{\mathcal{R}}(X_1, \dots, X_n)$

Très souvent, \mathcal{R} est construite à partir de $T = T(X_1, \dots, X_n)$ statistique de test **Definition 10.1**, elle-même basée sur un estimateur $\hat{\theta}_n$ de θ , paramètre d'intérêt.

↳ La question est: comment construire \mathcal{R} ?

10.1.2 RISQUES D'ERREUR D'UN TEST

Risque de 1^{ère} espèce.

De manière générale, on testera

$$H_0 : \theta = a_{a \in \Theta} \text{ contre } H_1 : \theta \neq a$$

$$H_0 : \theta \leq a \text{ contre } H_1 : \theta > a \text{ (ex: contrôle de qualité)}$$

Si on considère une partition $\Theta_0 \cup \Theta_1 = \Theta_{\text{espace des paramètres}}$, $\Theta_0 \cap \Theta_1 = \emptyset$, alors hypothèses sont: $H_0: \theta \in \Theta_0$ contre $H_1: \theta \in \Theta_1$

REMARQUE 10.2 (VOCABULAIRE) –

- *Test bilatère*

$\Theta_0 = \{a\}$ H_0 est une hypothèse simple.

$\Theta_1 = \Theta \setminus \{a\}$, H_1 est une hypothèse bilatère

- *Test unilatère*

si $\Theta_0 =]-\infty, a]$ et $\Theta_1 =]a, +\infty[$ H_1 et H_0 sont unilatères

$H_0 = \theta = a$ contre $H_1: \theta > a \rightarrow$ Test unilatère

◇

DEFINITION 10.3 (ERREUR DE 1^{ère} ESPÈCE) – – celle que l'on veut contrôler

$$\begin{aligned}\alpha : \Theta_0 &\longrightarrow [0, 1] \\ \theta &\mapsto P_\theta((X_1, \dots, X_n) \in \mathcal{R}) = E_\theta[\varphi(X)] \\ &= P_{H_0}(\text{rejet de } H_0) \\ &\quad \text{vrai}\end{aligned}$$

- niveau α ssi

$$\sup_{\theta \in \Theta_0} P_\theta((X_1, \dots, X_n) \in \mathcal{R}) \text{ op } \alpha$$

où pour

$$\text{op} = \begin{cases} \leq & \text{pour lois discrètes} \\ = & \text{pour lois continues exactes} \\ \rightarrow & \text{pour lois asymptotiques} \end{cases}$$

DEFINITION 10.4 (ERREUR DE 2^{nde} ESPÈCE) –

$$\begin{aligned}\beta : \Theta_1 &\longrightarrow [0, 1] \\ \theta &\mapsto P_\theta((X_1, \dots, X_n) \in \mathcal{R}^c) = P_{H_1}(\text{conserver } H_0) \\ &\quad \text{vrai}\end{aligned}$$

DEFINITION 10.5 (FONCTIONS DE PUISSANCE) –

$$\begin{aligned}\Pi : \Theta &\longrightarrow [0, 1] \\ \theta &\mapsto P_\theta((X_1, \dots, X_n) \in \mathcal{R})\end{aligned}$$

- si $\theta \in \Theta_0$: $\Pi(\theta) = \alpha(\theta)$
- si $\theta \in \Theta_1$: $\Pi(\theta) = P_\theta((X_1, \dots, X_n) \in \mathcal{R}) = 1 - P_{H_1}((X_1, \dots, X_n) \in \mathcal{R}^c) = 1 - \beta(\theta)$

10.2 EXEMPLE

(X_1, \dots, X_n) i.i.d. de loi $\mathcal{N}(\theta, 1)$. Hypothèses à tester:

$$H_0 : \theta \leq 0 \text{ contre } H_1 : \theta > 0$$

Comme $E[X_i] = \theta$ est inconnue, on l'estime avec $\hat{\theta} = \bar{X}$.

1. Première idée: rejet de H_0 si $\hat{\theta} > 0$

$$\mathcal{R} = \{(X_1, \dots, X_n), \hat{\theta}(X_1, \dots, X_n) > 0\}$$

Soit $\theta \leq 0$,

$$\alpha(\theta) = P_\theta(\hat{\theta} > 0) = P_\theta(\bar{X} > 0)$$

Quelle est la loi de \bar{X} ?

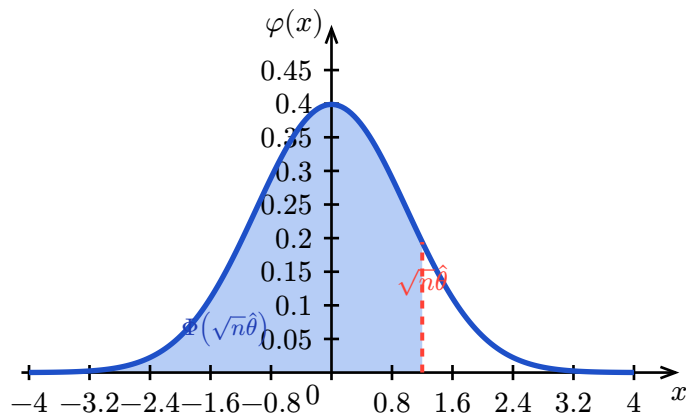
$\bar{X} \underset{\text{loi exacte}}{\sim} \mathcal{N}()$ car toute combinaison linéaire de variables aléatoires gaussiennes est une gaussienne.

$$E[\bar{X}] = E[X_i] = \theta, \text{Var}(\bar{X}) = \frac{1}{n} \text{Var}(X_i) = \frac{1}{n}$$

reflexe: centrer et réduire la loi normale:

$$\alpha(\theta) = P_{\theta} \left(\frac{\bar{X} - \theta}{\sqrt{\frac{1}{n}}} > -\sqrt{n}\theta \right) = P(\mathcal{N}(0,1) > -\sqrt{n}\theta) = 1 - \Phi(-\sqrt{n}\theta) = \Phi(\sqrt{n}\theta)$$

Où Φ est une fonction de répartition de la loi $\mathcal{N}(0,1)$



$$\text{niveau} = \sup_{\theta \leq 0} \Phi(\sqrt{n}\theta) = \Phi(0) = \frac{1}{2} = 50\%$$

Alors, on a une chance sur 2 de se tromper – ce qui n'est pas acceptable !

→ on souhaite α petit: $\alpha = 5\%$:

- $\mathcal{R} = \{\hat{\theta} > 0\} \rightarrow \mathcal{R} = \{\hat{\theta} > c\}$ ($c > 0$)
- valeur de $c = c(\alpha)$ telle que $\sup_{\theta \leq 0} \alpha(\theta) \leq \alpha$

$$\begin{aligned} \alpha(\theta) &= P_{\theta \leq 0}(\bar{X} > c) = P_{\theta \leq 0} \left(\frac{\bar{X} - \theta}{\sqrt{\frac{1}{n}}} > \frac{c - \theta}{\sqrt{\frac{1}{n}}} \right) \\ &= P(\mathcal{N}(0,1) > \sqrt{n}(c - \theta)) \end{aligned}$$

Condition de niveau:

Trouver c telle que

$$\begin{aligned} \sup_{\theta \leq 0} P(\mathcal{N}(0,1) > \sqrt{n}(c - \theta)) &\stackrel{\substack{= \\ \text{loi} \\ \text{continue}}}{=} \alpha \\ \Leftrightarrow P(\mathcal{N}(0,1) > \sqrt{nc}) &= \alpha \\ \Leftrightarrow 1 - \Phi(\sqrt{nc}) & \\ \Leftrightarrow \Phi(\sqrt{nc}) &= 1 - \alpha \\ \Leftrightarrow \sqrt{nc} = \Phi^{-1}(1 - \alpha) &\Rightarrow c_{\alpha} = \frac{1}{\sqrt{n}} \text{qnorm}_{1-\alpha} \end{aligned}$$

On a construit un test de niveau α avec

$$\mathcal{R} = \left\{ (X_1, \dots, X_n), \bar{X} > \frac{\text{qnorm}_{1-\alpha}}{\sqrt{n}} \right\}$$

application numérique: $\alpha = 5\% \Rightarrow \text{qnorm}_{1-\alpha} = 1.645, n = 100 \rightarrow c_\alpha = 0.1645$

expérience $\rightarrow \bar{X}^{\text{obs}} =$ réalisation de \bar{X} sur mes données

- si $\bar{X}^{\text{obs}} = 0.1 < c_\alpha$ on ne rejette pas H_0
- si $\bar{X}^{\text{obs}} = 0.3 > c_\alpha \Rightarrow$ rejet de H_0

10.3 CONSTRUCTION D'UN TEST

1.
 - Définir les hypothèse H_0 et H_1
 - Identifier le paramètre d'intérêt
2.
 - définir la forme de \mathcal{R} , forme de $H_1 \Rightarrow$ forme de $\mathcal{R} = \{T > c\}$ ou bien $\{T < c\}$
 - trouver une statistique de test
 - $T =$ version normalisée de

$$\hat{\theta} = \frac{\hat{\theta} - \theta}{\sqrt{\text{Var}(\hat{\theta})}}$$

3. Trouver le seuil c pour voir un test de niveau α

Tests d'un paramètre gaussien

§11

11.1 Résumé de la construction	48
11.2 P -valeur	48
11.2.1 Généralisation(formule de calcul d'une p -valeur)	49
11.2.2 Remarques	50
11.2.3 Règle de décision avec la p -valeur	50

11.1 RÉSUMÉ DE LA CONSTRUCTION

$X = (X_1, \dots, X_n)$ i.i.d. de loi P_θ

1. Préciser les hypothèses testées:

$$H_0: \theta \leq \theta_0 \text{ contre } H_1: \theta > \theta_0$$

2. Statistique de test: $T(X)$: sous H_0 $T(X)$ calculable. La loi de T sous H_0 permet de distinguer H_0 et H_1 .

↳ $\mathcal{R} = \{T(X) > c\}$ (sous H_1 , si la ... de T s'écarte de H_0 vers la droite), (si $H_1: \theta < \theta_0 \rightarrow \mathcal{R} = \{T(X) < c\}$, si test bilatère $H_1: \theta \neq \theta_0 \rightarrow \mathcal{R} = \{|T(X)| > c\} = \{T(X) > c \text{ ou } T(X) < -c\}$)

3. Règle de décision α niveau fixé,

- Condition de niveau:

$$\sup_{\theta \leq \theta_0} P_{H_0}(T(X) > c) = \alpha \text{ (si loi de } T \text{ sous } H_0 \text{ est continue)}$$

$$\leq \alpha \text{ (si loi de } T \text{ discrète)}$$

$$\xrightarrow{n \rightarrow +\infty} \text{ (si loi de } T \text{ est asymptotique)}$$

4. Application numérique:

- calcul du seuil
- calcul de la réalisation de $T = T^{\text{obs}} = T(X)$ si $x = (x_1, \dots, x_n)$ réalisation de (X_1, \dots, X_n) dans notre expérience
 - si $T^{\text{obs}} > c_\alpha$ alors on rejette H_0 , avec un risque de se tromper de α .
 - si $T^{\text{obs}} \leq c_\alpha$, on conserve H_0 , avec un risque de se tromper de inconnu (en général)

REMARQUE 11.1 – Le test de $H_0: \theta \geq \theta_0$ contre $H_1: \theta < \theta_0$, est le même que le test de $H_0: \theta = \theta_0$ contre $H_1: \theta < \theta_0$, $\mathcal{R} = \{T < c\}$ ◇

11.2 P -VALEUR

EXEMPLE 11.2 – (X_1, \dots, X_n) i.i.d. de loi $\mathcal{N}(0, 1)$; test $H_0: \theta = 0$, contre $H_1: \theta > 1$

$$\hookrightarrow T = \frac{\hat{\theta} - 0}{\frac{1}{\sqrt{n}}} = \sqrt{n}\hat{\theta} \underset{H_0}{\sim} \mathcal{N}(0, 1) \text{ où } \hat{\theta} = \bar{X}; \mathcal{R} = \{T > c\} \text{ condition de niveau } P_{\theta=0} \left(\underbrace{T}_{\rightarrow \sim \mathcal{N}(0,1)} > c \right) = \alpha \Rightarrow c_\alpha = \text{qnorm}_{1-\alpha} = \Phi^{-1}(1 - \alpha) \Rightarrow \mathcal{R} = \{\sqrt{n}\hat{\theta} > \Phi^{-1}(1 - \alpha)\}$$

rejet de $H_0 \Leftrightarrow \alpha > 1 - \Phi(\sqrt{n}\hat{\theta})$

$\alpha = 5\%$ rejette-t-on à 5%? 10%? 1%?

A.N $\hat{\theta} = 0.3$, $n = 100$, $\sqrt{n}\hat{\theta} = 3$, $1 - \Phi(3) \approx 10^{-3}$ \diamond

DEFINITION 11.3 – Si (X_1, \dots, X_n) i.i.d., $\mathcal{R} = \{T(X) > c_\alpha\}$. Pour une réalisation $x = (x_1, \dots, x_n)$ de $X = (X_1, \dots, X_n)$, on appelle p -valeur du test de région de \mathcal{R} :

$$\begin{aligned} \text{pval} &= \inf\{\alpha \in [0, 1], T(X) > c_\alpha\} \\ &= \inf\{\alpha, H_0 \text{ est rejetée au niveau } \alpha\} \end{aligned}$$

p value $(\varphi(X) = \mathbb{1}_{\mathcal{R}}(X))$ - niveau de significativité probabilité critique

La **Definition 11.3** peut sembler assez abstraite mais elle a une application assez intuitive. $T(X)$ dépend de notre échantillon observée et c_α dépend de la loi (sous H_0) et de α (important: indépendant des données observée).

Une propriété importante est que c_α est croissante en fonction de α . Finalement, on calcule $T(x_1, \dots, x_n) \in \mathbb{R}$, on cherche le plus petit alpha (équivalent à chercher le plus grand c_α tel que $T(x_1, \dots, x_n) > c_\alpha$). Puis $P(T(x_1, \dots, x_n) > c_\alpha) = \alpha = \text{pval}$ d'après **Definition 11.3**.

INTUITION 11.4 – P valeur nous dit: quelle est la probabilité d'avoir telles données aussi loin de notre région où on conserve H_0 . Plus α est petit, moins des valeurs extrêmes ($T(x_1, \dots, x_n)$ observées), donc plus est la tendance à rejeter H_0 . \diamond

EXEMPLE 11.5 –

$$\begin{aligned} \text{pval} &= 1 - \Phi(\sqrt{n}\hat{\theta}^{\text{obs}}) \\ &= 1 - \Phi(T^{\text{obs}}) \\ &= 1 - P(\mathcal{N}(0, 1) \leq T^{\text{obs}}) \\ \text{pval} &= P\left(\underbrace{\mathcal{N}(0, 1)}_{\text{loi de } T} > T^{\text{obs}}\right) \\ &= P\left(T > \underbrace{T^{\text{obs}}}_{\in \mathbb{R}}\right) \end{aligned}$$

\diamond

11.2.1 GÉNÉRALISATION (FORMULE DE CALCUL D'UNE P-VALEUR).

$T(X)$ statistique de test

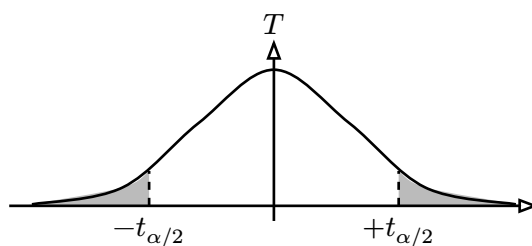
- $\mathcal{R} = \{T(X) > c\}$, alors p -valeur = $P_{H_0}(T(X) > T^{\text{obs}})$
- $\mathcal{R} = \{T(X) < c\}$, alors p -valeur = $P_{H_0}(T(X) < T^{\text{obs}})$
- $\mathcal{R} = \{|T(X)| > c\}$, p -valeur = $P_{H_0}(|T(X)| > |T^{\text{obs}}|)$

11.2.2 REMARQUES

REMARQUE 11.6 – Sur l'exemple

$$\begin{aligned} P_{H_0}(|T| > T^{\text{obs}}) &= P_{H_0}(T > T^{\text{obs}} \text{ ou } T < -T^{\text{obs}}) \\ &= P_{H_0}(T > T^{\text{obs}}) + P(T < -T^{\text{obs}}) \\ &= 1 - \Phi(T^{\text{obs}}) + \Phi(-T^{\text{obs}}) \\ (\text{symétrie}) &= 2(1 - \Phi(T^{\text{obs}})) \end{aligned}$$

la p -valeur du test bilatère est le double de la p -valeur du test unilatère.



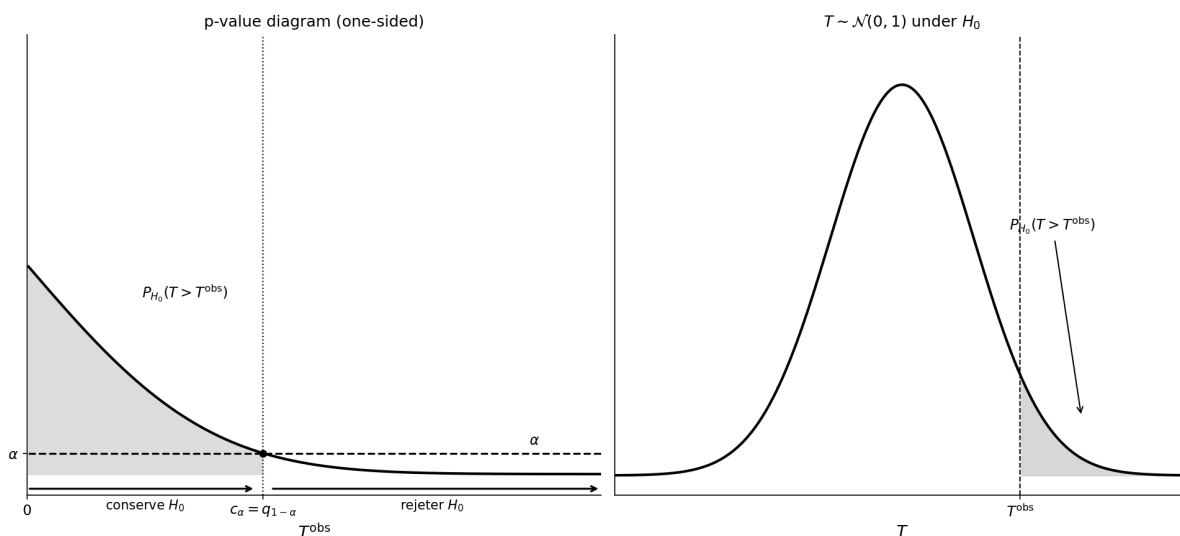
◇

REMARQUE 11.7 – Si la loi de T sous H_0 est discrète.

◇

11.2.3 RÈGLE DE DÉCISION AVEC LA p -VALEUR

EXEMPLE 11.8 – $\theta = 0$, contre $\theta > 0$, $T = \sqrt{n}\hat{\theta} \sim \mathcal{N}(0, 1)$



◇

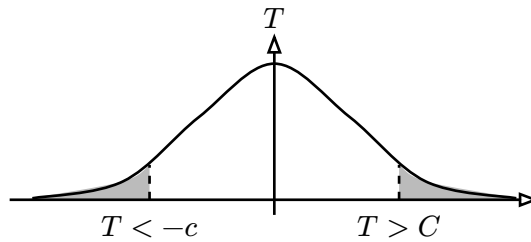
1. $H_0: \mu = \mu_0$ contre $H_1: \mu \neq \mu_0$, $\hat{\mu} = \bar{X}$

$$T = \frac{\bar{X} - \mu_0}{\frac{S_n}{\sqrt{n}}} \underset{\tilde{H}_0}{\text{loi exacte}} \text{Student}(n-1)$$

$H_1 \Rightarrow \mathcal{R} = \{|T| > c\}$, quelle est la règle de décision?
bilatère

\hookrightarrow calcul de $c = c_\alpha$ avec la condition de niveau $\Rightarrow c_\alpha = \text{qt}_{1-\frac{\alpha}{2}}(n-1)$ quantile Student($n-1$)

$$\begin{aligned} \text{pvaleur} &= P_{H_0}(|T| > |T^{\text{obs}}|) \\ (\text{symétrie de loi de Student}) &= 2P(T > |T^{\text{obs}}|) \\ &= 2(1 - F_{\text{Student}}(|T^{\text{obs}}|)) \end{aligned}$$

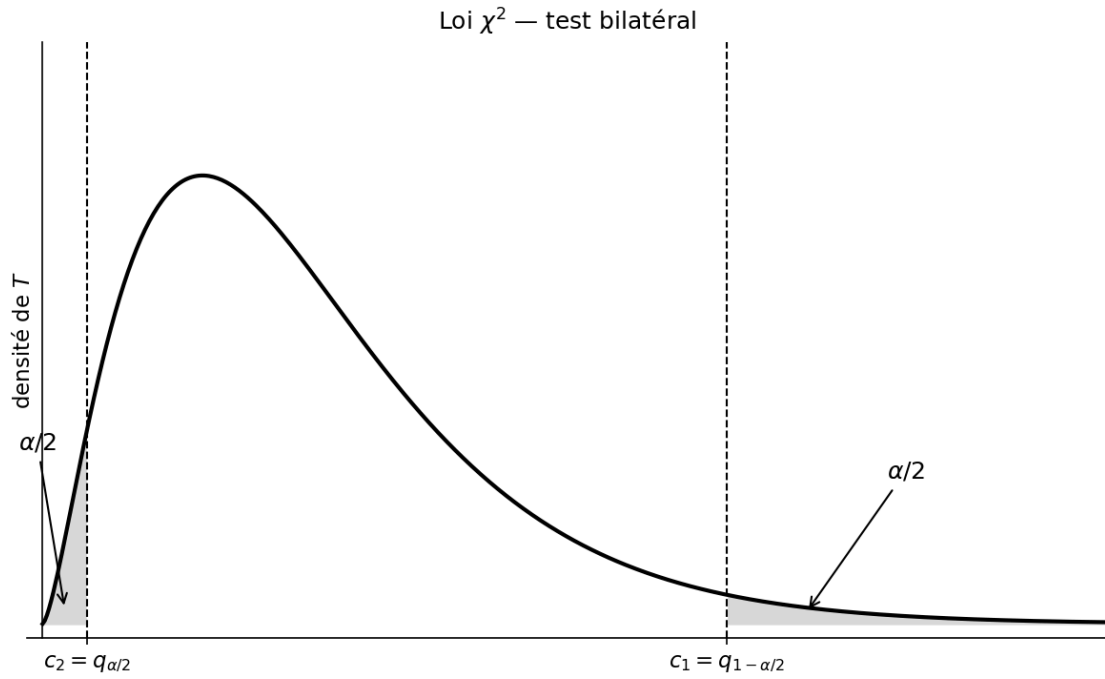


- Si $p\text{-valeur} < \alpha$ - on rejette H_0
 - Si $p\text{-valeur} \geq \alpha$ - on conserve H_0
2. $H_0: \sigma^2 = \sigma_0^2$ contre $H_1: \sigma^2 \neq \sigma_0^2$, σ^2 inconnu donc on l'estime
- S_n^2 sans biais
 - $\hat{\sigma}^2$ EMV

par le théorème de la loi des estimateurs dans le modèle gaussien

$$0 \leq T = \frac{n\hat{\sigma}^2}{\sigma^2} = \frac{(n-1)S_n^2}{\sigma^2} = \frac{\sum_i^n (X_i - \bar{X})^2}{\sigma^2} \underset{\tilde{H}_0}{\sim} \chi^2(n-1)$$

$$\mathcal{R} = \left\{ T > q_{1-\frac{\alpha}{2}}\chi^2(n-1) \text{ ou } T < q_{\frac{\alpha}{2}}\chi^2(n-1) \right\}$$



On calcule T^{obs} : on rejette H_0 ssi $T_{\text{obs}} > q_{1-\frac{\alpha}{2}}\chi^2$ ou $T_{\text{obs}} < q\chi^2_{\frac{\alpha}{2}}$

Quelle est la p -valeur?

$$\text{pvaleur} \underset{\substack{\equiv \\ \text{par convention}}}{=} 2P(T > T^{\text{obs}}) \text{ ou } 2P(T < T^{\text{obs}})$$

p -valeur $< 5\%$

Test de Student (t-test)

§12

Soient (X_1, \dots, X_{n_1}) i.i.d. $\mathcal{N}(\mu_1, \sigma_1^2)$, (Y_1, \dots, Y_{n_2}) i.i.d. $\mathcal{N}(\mu_2, \sigma_2^2)$. De plus, on suppose que les deux échantillons sont indépendants. Hypothèse supplémentaire: $\sigma_1^2 = \sigma_2^2 = \sigma^2$. On veut tester $H_0: \mu_1 = \mu_2$ contre $H_1: \mu_1 \neq \mu_2$

EXEMPLE 12.1 (EFFICACITÉ TRAITEMENT) — $\mu_1 = \mu_2$ contre $\mu_2 < \mu_1$ qui diminue le taux de cholestérol. \diamond

12.1 STATISTIQUE DE TEST

$\mu_1 - \mu_2 = 0$? idée:

- \hookrightarrow On estime $\hat{\mu}_1 - \hat{\mu}_2$ par $\bar{X} - \bar{Y}$
- \hookrightarrow Loi de $\bar{X} - \bar{Y}$

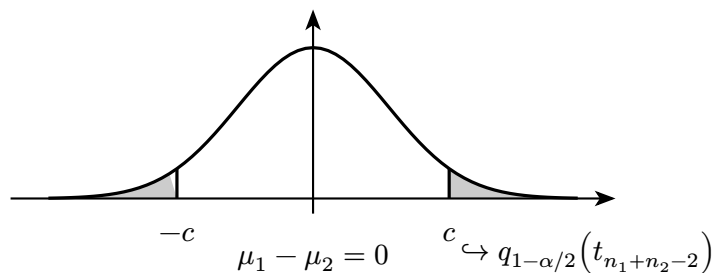
(X_n) et (Y_n) indépendantes donc CL de gaussiennes indépendantes = $\mathcal{N}(\mu_1 - \mu_2, \sigma^2(\frac{1}{n_1} + \frac{1}{n_2}))$

- $E[\bar{X} - \bar{Y}] = \mu_1 - \mu_2$ par linéarité de l'espérance + (X_j) et (Y_j) i.d.
- $\text{Var}(\bar{X} - \bar{Y}) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2(\frac{1}{n_1} + \frac{1}{n_2})$

Si σ^2 connue:

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\frac{\sigma}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}} \underset{H_0}{\sim} \mathcal{N}(0, 1)$$

loi de T sous H_0



PROPOSITION 12.2 — Sous les hypothèses de notre modèle

- 2 échantillons gaussiens indépendants
- $\sigma_1^2 = \sigma_2^2$

alors $S_n^2 = \frac{1}{n_1+n_2} \left(\sum_{i=1}^{n_1} (X_i - \bar{X})^2 + \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2 \right)$ est un estimateur sans biais de σ^2 et

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_n \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

a pour loi exacte la loi Student($n_1 + n_2 - 2$)

Preuve. admise □

12.2 RÉGION DE REJET

$$\mathcal{R} = \{|T| > c\}$$

12.3 RÈGLE DE DECISION

2 façon équivalents:

1.

calcul du seuil: α fixé (condition de niveau) $P_{\mu_1=\mu_2}(|T| > c_\alpha) \stackrel{T \text{ de loi continue}}{\equiv} \alpha \Rightarrow c_\alpha = \text{qt}_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$

2. Calcul de la p-valeur

$$\text{p-valeur} = P_{H_0}(|T| \geq |T^{\text{obs}}|) \stackrel{\substack{\text{loi de } T \\ \text{symétrique}}}{=} = 2P_{H_0}(T > |T^{\text{obs}}|) = 2(1 - F(|T^{\text{obs}}|))$$

où F de loi Student($n_1 + n_2 - 2$)

EXEMPLE 12.3 (APPLICATION NUMÉRIQUE) — $n_1 = 12$, $\bar{x} = 1.50$, $S_X^{\text{obs}} = 0.95 = \sqrt{\frac{1}{n_1-1} \sum (x_i - \bar{x})^2}$, $n_2 = 8$, $\bar{y} = 2.35$, $S_Y^{\text{obs}} = 1.35$

$$T^{\text{obs}} = \frac{\bar{x} - \bar{y}}{S_{(X,Y)}^{\text{obs}} \sqrt{\frac{1}{12} + \frac{1}{8}}}$$

$$S_{(X,Y)}^2 = \frac{1}{n_1 + n_2 - 2} \left(\underbrace{\sum (x_i - \bar{x})^2}_{(n_1-1)(S_X^{\text{obs}})^2} + \underbrace{\sum (y_i - \bar{y})^2}_{(n_2-1)(S_Y^{\text{obs}})^2} \right) \rightsquigarrow T^{\text{obs}} = 1.801$$

$$\alpha = 5\%, \quad c_\alpha = \text{qt}_{0.975} \left(\overbrace{12 + 8 - 2}^{ddl=18} \right) = 2.101, \quad \text{p-valeur} = 2 \left(1 - \underbrace{F_{\text{Student}(12+8-2)}(1.801)}_{\approx 0.95} \right) \approx$$

$0.10 > \alpha$ donc on ne rejette pas H_0

1. $|T^{\text{obs}}| < c_\alpha = 2.101$ on ne rejette pas H_0 , les 2 échantillons n'ont pas des moyennes différentes. ◇