Chapitre II:
La théorie de l'échantillonnage

MATH-F-207

Davy Paindaveine

Bloc 2, Bachelier en sciences mathématiques
Université libre de Bruxelles

2023–2024

Contenu du chapitre

Contenu du chapitre

Deux “mondes”

Soit le modèle statistique d’échantillonnage paramétrique

$$\left(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^n = \left\{ \mathrm{P}_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^k \right\}\right),$$

qui prévoit que l’observation $X^{(n)} = (X_1, \ldots, X_n)$ est faite de v.a. i.i.d. de distribution commune P_θ.

Les concepts introduits dans ce chapitre connaissent

1. une version population, “théorique”, relative à la distribution P_θ inconnue.
   Ces notions font typiquement intervenir le paramètre θ.
2. une version échantillon, “empirique”, associée aux observations $X_1, \ldots, X_n$.
   Ces notions ne peuvent pas faire intervenir le paramètre θ.

Statistiques

Définition 1

Une statistique $T(X^{(n)}) = T(X_1, \ldots, X_n)$ *est une fonction mesurable des observations* $X_1, \ldots, X_n$.

Exemples :

$$X_1, \quad \frac{1}{n}\sum_{i=1}^{n} X_i \ (\overset{\text{not}}{=} \bar{X}), \quad \text{et} \quad \arctan\left(X_{n-1} + X_n^3\right).$$

Une statistique peut également être à valeurs vectorielles ou ensemblistes. Ainsi,

$$\begin{pmatrix} X_1 \\ \bar{X} \end{pmatrix}, \quad X^{(n)} = (X_1, \ldots, X_n) \quad \text{et} \quad \left[\bar{X} - |X_1|, \bar{X} + |X_1|\right]$$

sont aussi des statistiques.

Statistiques d'ordre

Définition 2

Si, pour tout $\theta \in \Theta$, les X_i prennent des valeurs deux à deux différentes avec $P_\theta^{(n)}$-probabilité 1, la statistique d'ordre de $X^{(n)}$ est $(X_{(1)}, \ldots, X_{(n)})$, où $X_{(i)}$ est la $i$ème plus petite valeur parmi $X_1, \ldots, X_n$.

Terminologie: $X_{(i)}$ est la $i$ème statistique d'ordre.

En particulier, $X_{(1)} = \min(X_1, \ldots, X_n)$ et $X_{(n)} = \max(X_1, \ldots, X_n)$.

Question : Si $x^{(n)} = (1, -1, 0.5, -5, 7)$, alors que valent $x_{(1)}$, $x_{(5)}$ et $x_{(4)}$?

Attention: $X^{(n)} \neq X_{(n)}$

Fonction de répartition "population"

La distribution P_θ peut être décrite par sa fonction de répartition, définie par

$$\begin{aligned} F_\theta : \mathbb{R} &\to [0,1] \\ x &\mapsto F_\theta(x) = P_\theta[X_1 \leq x]. \end{aligned}$$

La fonction F_θ satisfait les propriétés caractéristiques suivantes:

(i) $\lim_{x\to-\infty} F_\theta(x) = 0$ et $\lim_{x\to\infty} F_\theta(x) = 1$

(ii) F_θ est non décroissante

(iii) F_θ est continue à droite

Pour tout $a < b$, on a $\mathrm{P}_\theta[a < X_1 \leq b] = F_\theta(b) - F_\theta(a)$.
Donc $\mathrm{P}_\theta[X_1 = b] = F_\theta(b) - F_\theta(b-0)$, où $F_\theta(b-0) \overset{\text{not}}{=} \lim_{x\overset{<}{\to}b} F_\theta(x)$.

Fonction de répartition empirique

La fonction de répartition empirique est l'équivalent empirique de la fonction de répartition population F_θ.

Elle est définie par

$$\begin{aligned} F_n : \mathbb{R} &\to [0,1] \\ x &\mapsto F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}[X_i \leq x]. \end{aligned}$$

Remarque : la fonction F_n est aléatoire. Pour chaque $x^{(n)}$ fixé, elle satisfait les propriétés caractéristiques (i)–(iii) d'une fonction de répartition.

Fonction de répartition: illustration

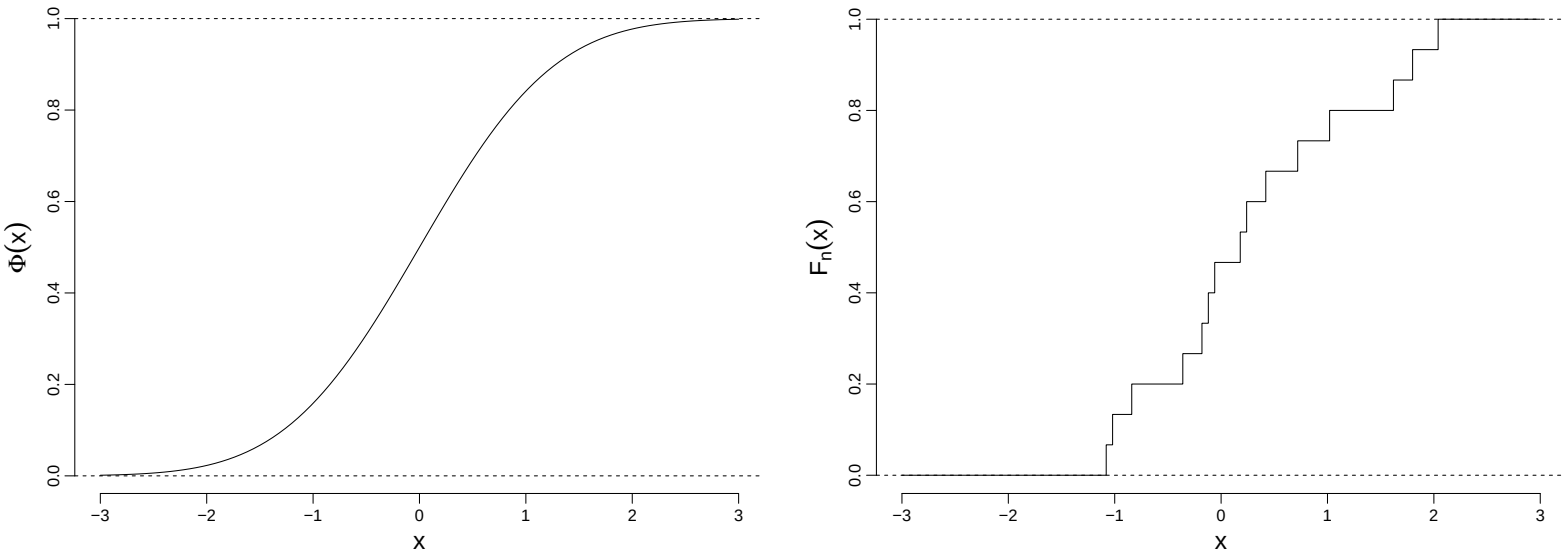

Figure 1: (Gauche:) fonction de répartition Φ de la loi $\mathcal{N}(0,1)$. (Droite:) fonction de répartition empirique F_n d'un échantillon de 15 v.a. i.i.d. $\mathcal{N}(0,1)$.

Fonction de répartition: illustration

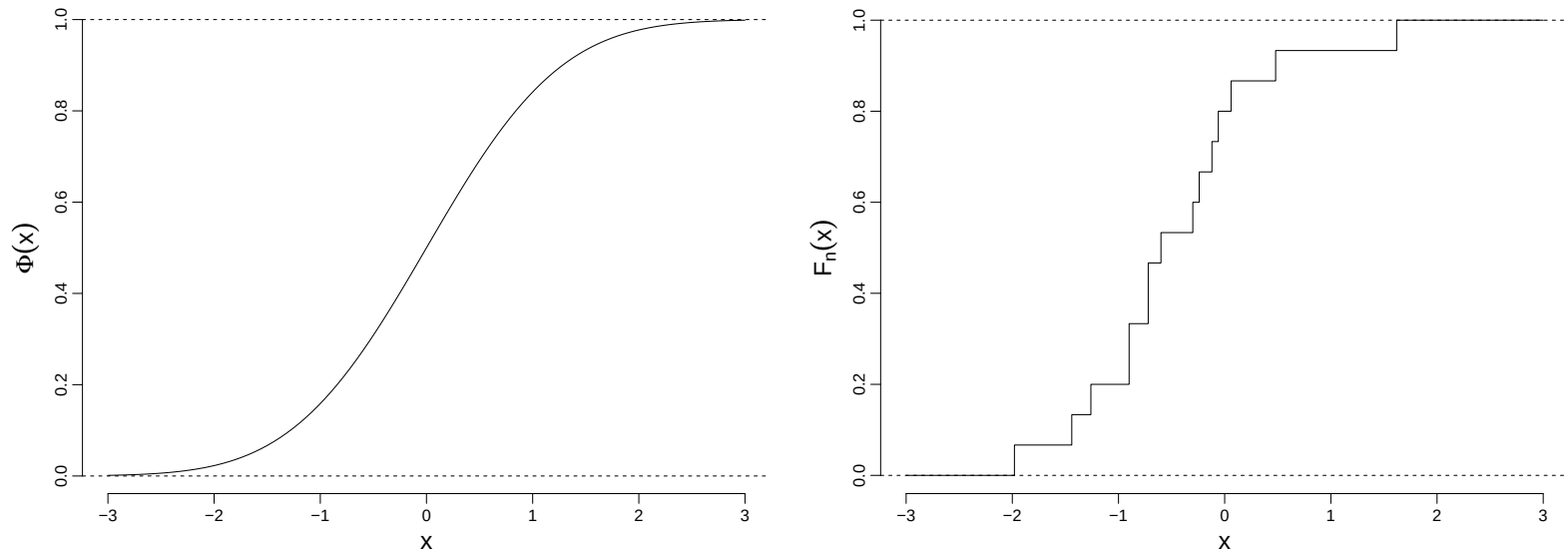

Figure 1: (Gauche:) fonction de répartition Φ de la loi $\mathcal{N}(0,1)$. (Droite:) fonction de répartition empirique F_n d'un échantillon de 15 v.a. i.i.d. $\mathcal{N}(0,1)$.

Fonction de répartition: illustration

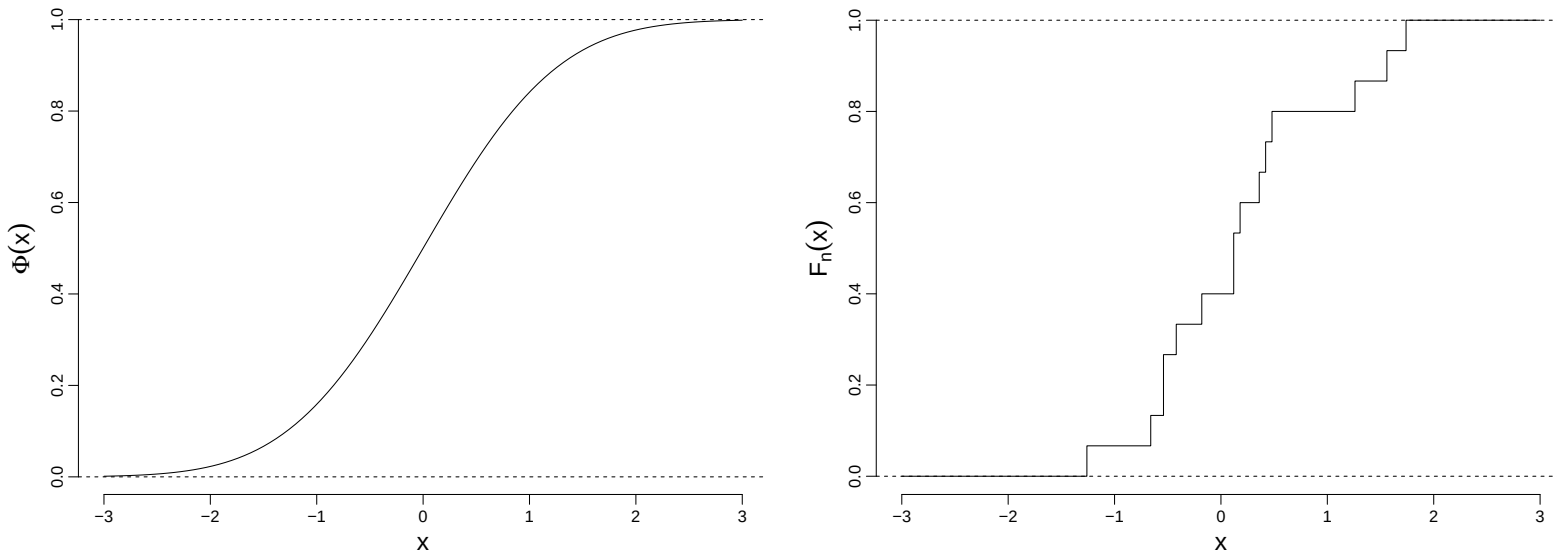

Figure 1: (Gauche:) fonction de répartition Φ de la loi $\mathcal{N}(0,1)$. (Droite:) fonction de répartition empirique F_n d'un échantillon de 15 v.a. i.i.d. $\mathcal{N}(0,1)$.

Convergence

La loi forte des grands nombres montre que, pour tout $\theta \in \Theta$ et tout $x \in \mathbb{R}$,

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}[X_i \leq x] \overset{\text{p.s.}}{\to} \mathrm{E}_\theta[\mathbb{I}[X_1 \leq x]] = \mathrm{P}_\theta[X_1 \leq x] = F_\theta(x)$$

sous $\mathrm{P}_\theta^{(n)}$.

Ce résultat de convergence p.s. ponctuelle peut être renforcé en le résultat de convergence p.s. uniforme suivant: pour tout $\theta \in \Theta$,

$$\sup_{x \in \mathbb{R}} |F_n(x) - F_\theta(x)| \overset{\text{p.s.}}{\to} 0 \quad \text{sous } \mathrm{P}_\theta^{(n)}.$$

Ce résultat est connu sous le nom de *théorème de Glivenko–Cantelli*.

Quantiles populations: version continue

Au niveau population, la fonction quantile est essentiellement l'inverse de la fonction de répartition F_θ.

Si $F_\theta : \mathbb{R} \to [0,1]$ est continue et strictement croissante, le quantile d'ordre $\alpha (\in (0,1))$ de la loi P_θ est défini comme

$$x_\alpha(\theta) \overset{\text{not}}{=} F_\theta^{-1}(\alpha).$$

Ainsi, $x_\alpha(\theta)$ est le nombre (unique) x tel que $X_1 (\sim \mathrm{P}_\theta)$ ait une probabilité α de prendre une valeur inférieure ou égale à $x_\alpha(\theta)$: $F_\theta(x_\alpha(\theta)) = \alpha$.

Quantiles populations: version générale

Si F_θ n'est pas continue, il se pourrait que l'équation $F_\theta(x) = \alpha$ n'ait pas de solution en x. Pour définir le quantile d'ordre α de façon générale, on a recours a un concept d'inverse généralisé.

Définition 3

Le quantile d'ordre $\alpha(\in (0,1))$ de X_1 de la loi P_θ est le nombre

$$x_\alpha(\theta) \stackrel{\text{def}}{=} \inf\Big\{x \in \mathbb{R} : F_\theta(x) \geq \alpha\Big\}.$$

La médiane, les 1er et 3ème quartiles, les déciles, les percentiles, ...

Les quantiles sont des mesures de position. La différence entre deux quantiles est une mesure de dispersion (intervalle interquartile : $x_{3/4}(\theta) - x_{1/4}(\theta)$).

Quantiles empiriques

Considérons l'observation $X^{(n)} = (X_1, \ldots, X_n)$.

Définition 4

Le quantile empirique d'ordre $\alpha (\in (0,1))$ de $X^{(n)}$ est le nombre

$$x_\alpha^{(n)} \stackrel{\text{def}}{=} \inf \Big\{ x \in \mathbb{R} : F_n(x) \geq \alpha \Big\},$$

où F_n désigne la fonction de répartition empirique associée à $X^{(n)}$.

$\rightsquigarrow$ On peut considérer la médiane empirique, les 1er et 3ème quartiles empiriques, etc.

Questions : Si $x^{(n)} = (1, -1, 0.5, -5, 7)$, alors que valent $x_{1/2}^{(n)}$, $x_{1/4}^{(n)}$ et $x_{8/10}^{(n)}$?

Moments population

Définition 5

La distribution de X_1 (ou, par extension, X_1 elle-même) admet des moments finis d'ordre $r(>0)$ sous P_θ si et seulement si

$$\mathrm{E}_\theta[|X_1|^r] = \int_{-\infty}^{\infty} |x|^r \, dF_\theta(x)$$

existe et est finie.

Si une v.a. admet des moments finis d'ordre $r(>0)$, alors elle admet aussi des moments finis d'ordre s pour tout $s \in (0, r)$ (Preuve : en live).

Intégrales de Lebesgue–Stieltjes

Au slide précédent, l'intégrale

$$\int_{-\infty}^{\infty} |x|^r \, dF_\theta(x)$$

est une intégrale de Lebesgue–Stieltjes:

- Si P_θ est continue, de densité f_θ,

$$\int_{-\infty}^{\infty} g(x) \, dF_\theta(x) = \int_{-\infty}^{\infty} g(x) f_\theta(x) \, dx.$$

- Si P_θ est discrète, avec les valeurs possibles $x_i(\theta)$ et probabilités correspondantes $p_i(\theta)$, $i \in \mathcal{I}$ (au plus dénombrable),

$$\int_{-\infty}^{\infty} g(x) \, dF_\theta(x) = \sum_{i \in \mathcal{I}} g(x_i(\theta)) p_i(\theta).$$

Moments non centrés

Définition 6

Dans un modèle statistique d'échantillonnage paramétrique, supposons que, pour tout $\theta \in \Theta$, X_1 *admette des moments finis d'ordre* $r (\in \mathbb{N}_0)$ *sous* P_θ. *Alors la quantité*

$$\mu'_r(\theta) = \mathrm{E}_\theta[X_1^r]$$

est appelée moment population non centré d'ordre r *(sous* P_θ*).*

Le moment non centré d'ordre 1, $\mu'_1(\theta) = \mathrm{E}_\theta[X_1]$, est l'espérance de X_1, et est souvent noté μ ou $\mu(\theta)$. C'est une mesure de position de X_1.

Moments centrés

Définition 7

Dans un modèle statistique d'échantillonnage paramétrique, supposons que, pour tout $\theta \in \Theta$, X_1 *admette des moments finis d'ordre* $r (\in \mathbb{N}_0)$ *sous* P_θ. *Alors la quantité*

$$\mu_r(\theta) = \mathrm{E}_\theta[(X_1 - \mu(\theta))^r]$$

est appelée moment population centré d'ordre r *(sous* P_θ*).*

Le moment centré d'ordre 2 est la variance de X_1, $\mathrm{Var}_\theta[X_1]$, laquelle sera souvent notée σ^2 ou $\sigma^2(\theta)$. C'est une mesure de dispersion de X_1.

Rappel: $\mathrm{Var}_\theta[X_1] = \mathrm{E}_\theta[X_1^2] - (\mathrm{E}_\theta[X_1])^2$, c'est-à-dire $\mu_2(\theta) = \mu_2'(\theta) - (\mu_1'(\theta))^2$.

Coefficient d'asymétrie

Définition 8

Si les moments d'ordre 3 sont finis, alors la quantité

$$\gamma_1 = \frac{\mu_3(\theta)}{(\mu_2(\theta))^{3/2}}$$

est appelée le coefficient d'asymétrie de Fisher.

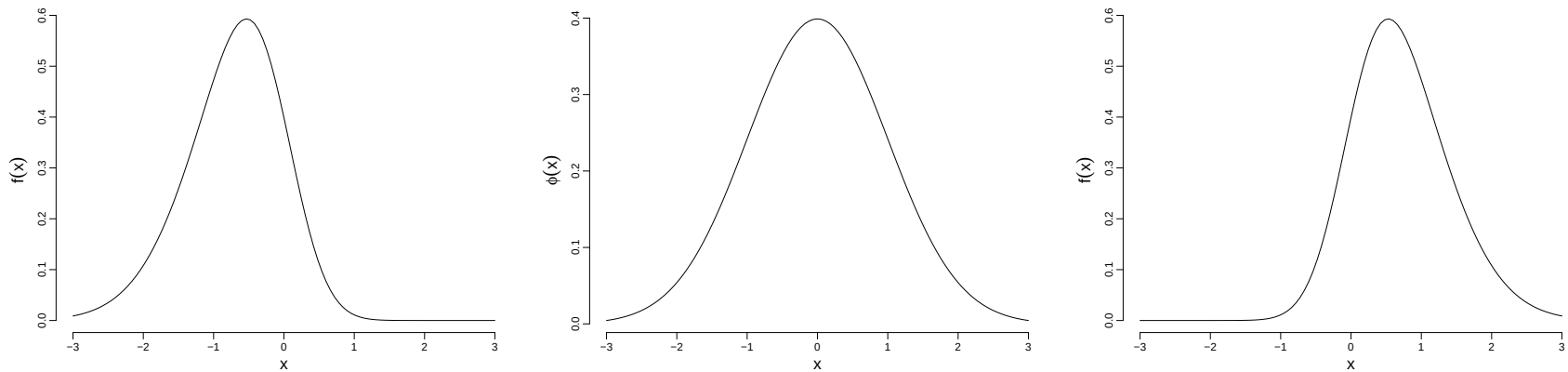

Figure 2: (Gauche:) densité asymétrique à gauche ($\gamma_1 < 0$). (Centre:) densité symétrique ($\gamma_1 = 0$). (Droite:) densité asymétrique à droite ($\gamma_1 > 0$).

Coefficient d'aplatissement

Définition 9

Si les moments d'ordre 4 sont finis, alors la quantité

$$\gamma_2 = \frac{\mu_4(\theta)}{(\mu_2(\theta))^2} - 3$$

est appelée le coefficient d'aplatissement de Fisher.

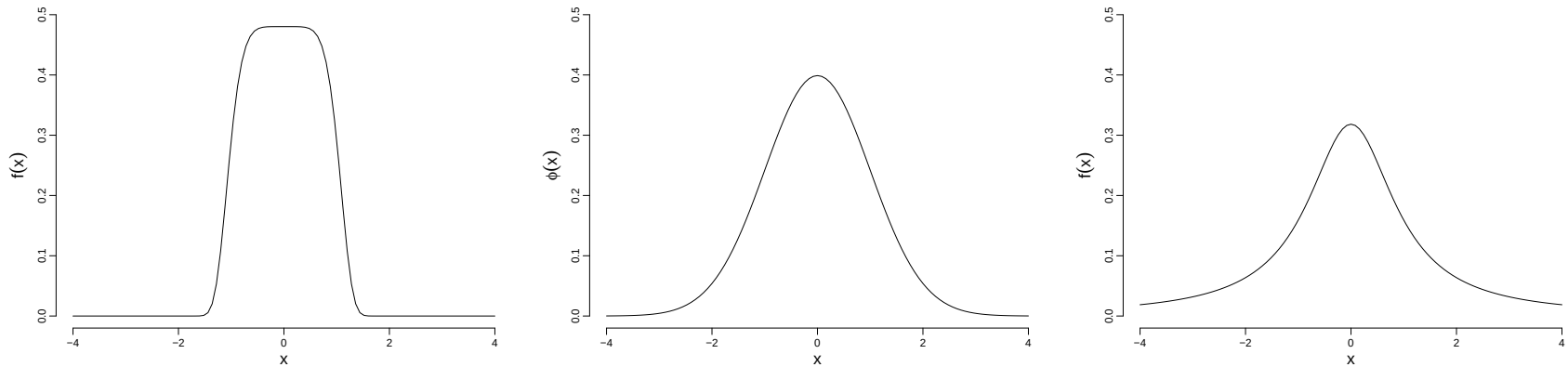

Figure 3: (De gauche à droite:) densité *platykurtique* ($\gamma_2 < 0$), normale (*mésokurtique*; $\gamma_2 = 0$) et *leptokurtique* ($\gamma_2 > 0$).

Interprétations

- Les lois symétriques livrent $\gamma_1 = 0$.
- Les lois normales livrent $\gamma_2 = 0$. Des distributions "à queues lourdes" ($\gamma_2 > 0$) prévoient plus d'observations extrêmes que pour les lois normales (à variances égales). Ce sont ces queues lourdes qui provoquent l'inexistence de certains moments.
- C'est un souci d'invariance qui explique la présence du dénominateur dans les expressions de γ_1 et γ_2 : quels que soient $a \in \mathbb{R}_0^+$ et $b \in \mathbb{R}$, γ_1 et γ_2 ne changent pas s'ils sont calculés sur $aX_1 + b$ plutôt que sur X_1.

Moments empiriques

Au niveau échantillon, les moments empiriques non centrés et centrés d'ordre r sont respectivement les statistiques

$$m'_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r \quad \text{et} \quad m_r = \frac{1}{n}\sum_{i=1}^{n} (X_i - m'_1)^r.$$

- Ils ne requièrent pas de condition de moments finis...
- ...Mais de telles conditions permettent de garantir que ces moments empiriques sont proches des moments population associés.
- En effet: si X_1 admet des moments finis d'ordre r sous P_θ, alors la loi forte des grands nombres assure que, sous P_θ,

$$m'_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r \overset{\text{p.s.}}{\to} \mathrm{E}_\theta[X_1^r] = \mu'_r(\theta) \quad \text{quand } n \to \infty.$$

Moments empiriques (2)

Les équivalents empiriques de la moyenne et de la variance population sont la moyenne empirique

$$\bar{X} \overset{\text{not}}{=} m'_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

et la variance empirique

$$s^2 \overset{\text{not}}{=} m_2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2.$$

Les coefficients d'asymétrie et d'aplatissement de Fisher empiriques sont

$$\frac{m_3}{m_2^{3/2}} \quad \text{et} \quad \frac{m_4}{m_2^2} - 3,$$

respectivement.

Contenu du chapitre

Loi échantillonnée

Soit le modèle statistique paramétrique

$$\left(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{(n)} = \left\{ \mathrm{P}_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^k \right\}\right),$$

où les mesures de probabilité $\mathrm{P}_\theta^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \ldots, X_n)$. Soit $T(X^{(n)})$ une statistique à valeurs dans $\mathbb{R}^m$.

Définition 10

La distribution échantillonnée de $T(X^{(n)})$ sous $\mathrm{P}_\theta^{(n)}$ est la mesure de probabilité définie par

$$\mathrm{P}_\theta^{T(n)}[B] = \mathrm{P}_\theta^{(n)}[T(X^{(n)}) \in B] = \mathrm{P}_\theta^{(n)}\left[\{x^{(n)} \in \mathbb{R}^n : T(x^{(n)}) \in B\}\right]$$

pour tout B dans la sigma-algèbre de Borel $\mathcal{B}^m$ sur $\mathbb{R}^m$.

Exemple du bus

Soit $X^{(n)} = (X_1, \ldots, X_n)$, où les X_i sont i.i.d. $\mathrm{Unif}([0, \theta])$, avec $\theta \in \mathbb{R}_0^+$.
Soit la statistique $T(X^{(n)}) = X_{(n)} = \max(X_1, \ldots, X_n)$.

Alors la loi échantillonnée de $T(X^{(n)})$ a pour fonction de répartition

$$\begin{aligned} F_\theta^{X_{(n)}}(x) &= \mathrm{P}_\theta^{(n)}[X_{(n)} \leq x] = \mathrm{P}_\theta^{(n)}[X_1 \leq x, \ldots, X_n \leq x] \\ &= \prod_{i=1}^{n} \mathrm{P}_\theta[X_i \leq x] = \begin{cases} 0 & \text{si } x < 0 \\ (\frac{x}{\theta})^n & \text{si } 0 \leq x \leq \theta \\ 1 & \text{si } x > \theta, \end{cases} \end{aligned}$$

donc pour densité

$$f_\theta^{X_{(n)}}(x) = \frac{n x^{n-1}}{\theta^n} \mathbb{I}[0 \leq x \leq \theta].$$

Exemple électoral

Soit $X^{(n)} = (X_1, \ldots, X_n)$, où les X_i sont i.i.d. $\mathrm{Bern}(p)$, avec $p \in [0, 1]$.
Soit la statistique $T(X^{(n)}) = \sum_{i=1}^{n} X_i$

Alors la loi échantillonnée de $T(X^{(n)})$ est la loi binomiale de paramètres n et p ($\mathrm{Bin}(n, p)$).

Rappel : $X \sim \mathrm{Bin}(n, p) \Leftrightarrow P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$, pour $k = 0, \ldots, n$.

Exemple du verre de bière

Soit $X^{(n)} = (X_1, \ldots, X_n)$, où les X_i sont i.i.d. $\mathcal{N}(\mu, \sigma^2)$, avec $(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_0^+$.

Rappel : Soient $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ indépendantes. Alors, $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Donc

- la loi échantillonnée de $T(X^{(n)}) = \sum_{i=1}^n X_i$ est la loi $\mathcal{N}(n\mu, n\sigma^2)$.
- La loi échantillonnée de $T(X^{(n)}) = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ est la loi $\mathcal{N}(\mu, \frac{\sigma^2}{n})$.

Une remarque importante

Pouvoir calculer la distribution échantillonnée d'une statistique $T(X^{(n)})$ reste exceptionnel. Dans les autres cas, on peut souvent

(i) déterminer les (ou certains) moments de la distribution échantillonnée.

(ii) déterminer la distribution échantillonnée *asymptotique*, c'est-à-dire la limite en loi de la distribution échantillonnée *exacte*[1].

[1] Dans la suite, le qualificatif "exact" sera relatif à une valeur fixée de n.

Exemple pour (i)

Soit le modèle statistique non paramétrique pour lequel l'observation $X^{(n)} = (X_1, \ldots, X_n)$ regroupe des v.a. i.i.d. dont la loi commune est de fonction de répartition F et admet des moments finis d'ordre 2.

Soit la statistique $T(X^{(n)}) = \bar{X}$.

En posant $\mu = \mathrm{E}_F[X_1]$ et $\sigma^2 = \mathrm{Var}_F[X_1]$,

$$\mathrm{E}_F[\bar{X}] = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}_F[X_i] = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$

et

$$\mathrm{Var}_F[\bar{X}] = \frac{1}{n^2}\mathrm{Var}_F\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}_F[X_i] = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}.$$

On ne peut pas calculer la distribution échantillonnée (exacte) de $T(X^{(n)})$, mais on a pu calculer deux de ses moments.

Exemple pour (ii)

Pour le même modèle, le TCL livre directement que, sous $\mathrm{P}_F^{(n)}$,

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2) \quad \text{quand } n \to \infty.$$

Ceci décrit complètement la distribution échantillonnée asymptotique de $\bar{X}$ sous $\mathrm{P}_F^{(n)}$.

En écrivant "$\approx$" pour "est approximativement de loi", on a donc pour n grand que $\sqrt{n}(\bar{X} - \mu) \approx \mathcal{N}(0, \sigma^2)$ sous $\mathrm{P}_F^{(n)}$, ou encore que

$$\bar{X} \approx \mathcal{N}(\mu, \tfrac{\sigma^2}{n})$$

sous $\mathrm{P}_F^{(n)}$.

Contenu du chapitre

Introduction

Le lemme de Fisher est un résultat classique, qui précise la distribution échantillonnée (exacte) de la statistique

$$T(X^{(n)}) = \begin{pmatrix} \bar{X} \\ s^2 \end{pmatrix}$$

dans le modèle d'échantillonnage paramétrique gaussien.

Nous savons déjà que $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$. Il reste donc à déterminer

- la distribution échantillonnée de s^2
- la structure de dépendance de $\bar{X}$ et s^2.

Lemme de Fisher

Théorème 11

Soient $X_1, \ldots, X_n$ $(n \geq 2)$ *i.i.d. de loi commune* $\mathcal{N}(\mu, \sigma^2)$. *Alors,*

(i) $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$,

(ii) $ns^2/\sigma^2 \sim ?$

(iii) $\bar{X}$ *et* s^2 *sont indépendantes*.

Variable aléatoire chi-carré

Soit $k \in \mathbb{N}_0$.

Définition 12

La v.a. Q est de loi chi-carré à k degrés de liberté (notation: $Q \sim \chi^2_k$) $\Leftrightarrow$ Q a la même distribution que $\sum_{i=1}^{k} Z_i^2$, où les Z_i sont i.i.d. de loi commune $\mathcal{N}(0,1)$.

On peut montrer que si $Q \sim \chi^2_k$, alors Q admet la densité

$$f^Q(x) = \frac{1}{2^{k/2}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} \exp(-x/2)\, \mathbb{I}[x > 0],$$

où $\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t)\, dt$ est la fonction Gamma d'Euler.

Exercice : Vérifier que $\mathrm{E}[Q] = k$ et $\mathrm{Var}[Q] = 2k$.

Si $Q_1 \sim \chi^2_{k_1}$ et $Q_2 \sim \chi^2_{k_2}$ sont indépendantes, alors $Q_1 + Q_2 \sim \chi^2_{k_1+k_2}$ (pourquoi?)

Lemme de Fisher

Théorème 13

Soient $X_1, \ldots, X_n$ $(n \geq 2)$ *i.i.d. de loi commune* $\mathcal{N}(\mu, \sigma^2)$. *Alors,*

(i) $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$,

(ii) $ns^2/\sigma^2 \sim \chi^2_{n-1}$

(iii) $\bar{X}$ *et* s^2 *sont indépendantes*.

Preuve du lemme de Fisher (1)

Preuve: posons

$$Z_i = \frac{X_i - \mu}{\sigma}, \quad i = 1, \ldots, n.$$

Puisque $X_i = \sigma Z_i + \mu$ pour tout i, on a (exercice)

$$\bar{X} = \sigma \bar{Z} + \mu, \quad \text{où } \bar{Z} := \frac{1}{n} \sum_{i=1}^{n} Z_i$$

et

$$\frac{ns^2}{\sigma^2} = ns_z^2, \quad \text{où } s_z^2 := \frac{1}{n} \sum_{i=1}^{n} (Z_i - \bar{Z})^2.$$

Donc il suffit de prouver que

- $ns_z^2 \sim \chi^2_{n-1}$
- $\bar{Z}$ et s_z^2 sont mutuellement indépendantes.

Preuve du lemme de Fisher (2)

Notons que $Z^{(n)} := (Z_1, \ldots, Z_n)$ a des composantes i.i.d. $\mathcal{N}(0,1)$.
Donc la densité de $Z^{(n)}$ en $z^{(n)} = (z_1, \ldots, z_n)$ est

$$f^{Z^{(n)}}(z^{(n)}) = \prod_{i=1}^{n} f^{Z_i}(z_i) = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z_i^2} \right) = \left(\frac{1}{\sqrt{2\pi}} \right)^n e^{-\frac{1}{2} \|z^{(n)}\|^2}.$$

Posons $Y^{(n)} := OZ^{(n)}$, où O est une matrice $n \times n$ orthogonale O.
Alors, pour tout borélien B de $\mathbb{R}^n$,

$$\int_B f^{Y^{(n)}}(y^{(n)})\, dy^{(n)} = P[Y^{(n)} \in B] = P[OZ^{(n)} \in B] = P[Z^{(n)} \in O^{-1}B]$$

$$= \int_{O^{-1}B} f^{Z^{(n)}}(z^{(n)})\, dz^{(n)} = \int_B f^{Z^{(n)}}(z^{(n)})\, dz^{(n)}.$$

Donc $f^{Y^{(n)}} = f^{Z^{(n)}}$, ce qui montre que $Y^{(n)}$ et $Z^{(n)}$ ont la même distribution.
On conclut que $Y^{(n)} = (Y_1, \ldots, Y_n)$ a des composantes i.i.d. $\mathcal{N}(0,1)$.

Preuve du lemme de Fisher (3)

Fixons O de la forme

$$O = \begin{pmatrix} \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ & \Gamma & \end{pmatrix}.$$

Alors

$$Y_1 = (Y^{(n)})_1 = (OZ^{(n)})_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i = \sqrt{n}\bar{Z}$$

et

$$\begin{aligned} ns_z^2 &= \sum_{i=1}^n (Z_i - \bar{Z})^2 = \sum_{i=1}^n (Z_i^2 - 2Z_i\bar{Z} + \bar{Z}^2) \\ &= \left(\sum_{i=1}^n Z_i^2\right) - 2\left(\sum_{i=1}^n Z_i\right)\bar{Z} + n\bar{Z}^2 = \left(\sum_{i=1}^n Z_i^2\right) - n\bar{Z}^2 \\ &= \|Z^{(n)}\|^2 - (\sqrt{n}\bar{Z})^2 = \|Y^{(n)}\|^2 - Y_1^2 = Y_2^2 + \ldots + Y_n^2 \sim \chi^2_{n-1}. \end{aligned}$$

Et $\bar{Z} = \frac{1}{\sqrt{n}} Y_1$ et $s_z^2 = \frac{1}{n}(Y_2^2 + \ldots + Y_n^2)$ sont mutuellement indépendantes. $\square$

Un corollaire du lemme de Fisher

Corollaire 14

Soient $X_1, \ldots, X_n$ $(n \geq 2)$ *i.i.d. de loi commune* $\mathcal{N}(\mu, \sigma^2)$. *Alors*

$$\frac{\sqrt{n-1}(\bar{X} - \mu)}{s} \sim t_{n-1},$$

où s *désigne la racine carrée de la variance empirique* s^2.

Variable aléatoire de Student

Soit $k \in \mathbb{N}_0$.

Définition 15

La v.a. V est de loi de Student (ou de loi t) à k degrés de liberté (notation: $V \sim t_k$) $\Leftrightarrow$ V a la même distribution que

$$\frac{Z}{\sqrt{Q/k}},$$

où $Z \sim \mathcal{N}(0,1)$ et $Q \sim \chi^2_k$ sont indépendantes.

On peut montrer que si $V \sim t_k$, alors V admet la densité

$$f^V(x) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\,\Gamma(\frac{k}{2})}\left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}} \qquad \text{pour tout } x \in \mathbb{R}.$$

$\rightsquigarrow$ Si $V \sim t_1$ (*loi de Cauchy*), $f^V(x) = 1/\{\pi(1+x^2)\}$ pour tout $x \in \mathbb{R}$.

Variable aléatoire de Student (2)

Les queues de la loi t_k sont donc d'autant plus lourdes que k est petit, ce qui implique l'inexistence de certains moments.

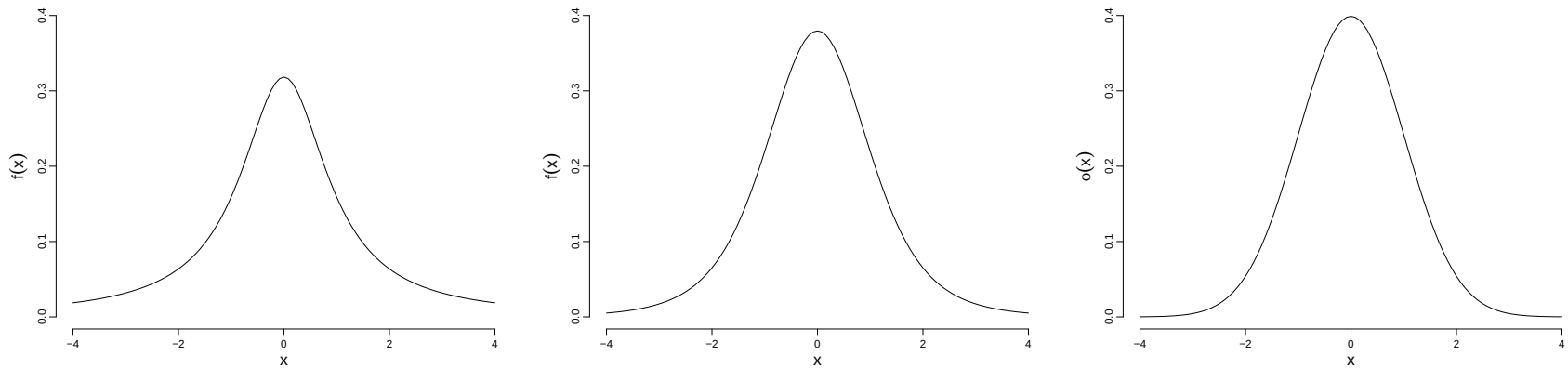

Figure 4: (De gauche à droite:) fonctions de densité de v.a. t_1, t_5 et $\mathcal{N}(0,1)$

Exercice : $V \sim t_k$ a des moments finis d'ordre $r \Leftrightarrow r < k$ ($\rightsquigarrow$ si $V \sim t_1$, alors V n'admet pas de moment fini d'ordre 1, et on ne peut donc pas parler de $\mathrm{E}[V]$!)

Exercice : Si $V_k \sim t_k$, alors $V_k \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ quand $k \to \infty$.

Résumé du chapitre

- Expressions population $><$ expressions empiriques
- Notions de fonction de répartition, moments et quantiles
- Lois échantillonnées (exactes et asymptotiques)
- Lemme de Fisher