CentraleSupélec

Statistique et apprentissage

Chargés de cours (ordre alphabétique) :

Julien Bect, Gilles Faÿ, Ziad Kobeissi, Laurent Le Brusquet, Vincent Lescarret, Arshak Minasyan, Arthur Tenenhaus† & Xujia Zhu

† Coordinateur du cours

Cours 3/9

Lois asymptotiques
et intervalles de confiance

Objectifs du cours 3

- Compléter l'analyse asymptotique des estimateurs déjà présentée (consistance) par l'étude de leur vitesse de convergence.
- Montrer la/les démarche(s) utilisée(s) pour construire des intervalles de confiance.

Plan du cours

Rappel : cadre mathématique

Pour toute la section :

- On considère un modèle statistique

$$\left(\underline{\mathcal{X}}, \underline{\mathscr{A}}, \left\{\mathbb{P}_{\theta}^{\underline{X}},\, \theta \in \Theta\right\}\right),$$

  le plus souvent paramétrique ($\Theta \subset \mathbb{R}^p$).

- $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathrm{P}_\theta$, définis sur un même $(\Omega, \mathscr{F}, \mathbb{P}_\theta)$.

- On veut estimer un « paramètre d'intérêt » :
  - soit θ lui-même (on supposera dans ce cas $\Theta \subset \mathbb{R}^p$),
  - soit, plus généralement, $\eta = g(\theta) \in \mathbb{R}^q$.

Rappel : cadre mathématique

Pour toute la section :

- On considère un modèle statistique

$$\left(\underline{\mathcal{X}}, \underline{\mathscr{A}}, \left\{\mathbb{P}_\theta^{X},\ \theta \in \Theta\right\}\right),$$

  le plus souvent paramétrique ($\Theta \subset \mathbb{R}^p$).

- $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathrm{P}_\theta$, définis sur un même $(\Omega, \mathscr{F}, \mathbb{P}_\theta)$.

- On veut estimer un « paramètre d'intérêt » :
  - soit θ lui-même (on supposera dans ce cas $\Theta \subset \mathbb{R}^p$),
  - soit, plus généralement, $\eta = g(\theta) \in \mathbb{R}^q$.

Rappel : cadre mathématique

Pour toute la section :

- On considère un modèle statistique

$$\left(\underline{\mathcal{X}}, \underline{\mathscr{A}}, \left\{\mathbb{P}_\theta^{\underline{X}},\, \theta \in \Theta\right\}\right),$$

  le plus souvent paramétrique ($\Theta \subset \mathbb{R}^p$).

- $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathrm{P}_\theta$, définis sur un même $(\Omega, \mathscr{F}, \mathbb{P}_\theta)$.

- On veut estimer un « paramètre d'intérêt » :
  - soit θ lui-même (on supposera dans ce cas $\Theta \subset \mathbb{R}^p$),
  - soit, plus généralement, $\eta = g(\theta) \in \mathbb{R}^q$.

Plan du cours

Plan du cours

Rappel probabilités : Théorème Central Limite (TCL)

Théorème

Soient

- une suite $(X_n)_{n\in\mathbb{N}^*}$ de vecteurs aléatoires iid, à valeurs dans $\mathbb{R}^d$ et admettant un moment d'ordre 2.
- $\mu = \mathbb{E}(X_1)$ et $\Sigma = \mathrm{var}(X_1) \in \mathbb{R}^{d\times d}$.

Alors : $$\boxed{\sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,\Sigma),}$$

avec $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ la moyenne empirique.

$\Rightarrow$ On dit que la moyenne empirique $\bar{X}_n$

- est un estimateur asymptotiquement gaussien de $\mu = \mathbb{E}(X_1)$

  définition : normalité asymptotique

- qui converge à la vitesse $\frac{1}{\sqrt{n}}$.

  définition : vitesse de convergence

Rappel probabilités : Théorème Central Limite (TCL)

Théorème

Soient

- une suite $(X_n)_{n\in\mathbb{N}^*}$ de vecteurs aléatoires iid, à valeurs dans $\mathbb{R}^d$ et admettant un moment d'ordre 2.
- $\mu = \mathbb{E}(X_1)$ et $\Sigma = \mathrm{var}(X_1) \in \mathbb{R}^{d\times d}$.

Alors : $$\boxed{\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,\Sigma),}$$

avec $\bar{X}_n = \dfrac{1}{n}\sum_{i=1}^n X_i$ la moyenne empirique.

$\Rightarrow$ On dit que la moyenne empirique $\bar{X}_n$

- est un estimateur asymptotiquement gaussien de $\mu = \mathbb{E}(X_1)$

  définition : normalité asymptotique
- qui converge à la vitesse $\frac{1}{\sqrt{n}}$.

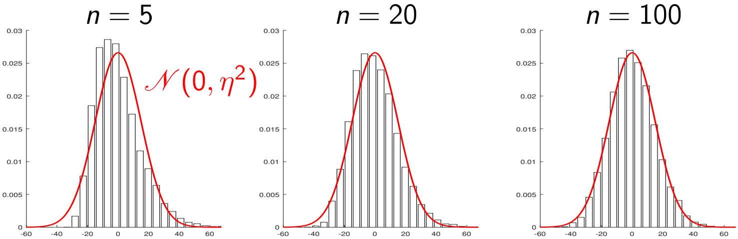
  définition : vitesse de convergence

Exemple : application « fiabilité composant »

Rappels :

- $X_i \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$, $\theta > 0$, et $\eta = \mathbb{E}_\theta(X_1) = \frac{1}{\theta}$.
- $\hat{\eta}_n = \bar{X}_n$ est obtenu par MV et méthode des moments.

➠ Application directe du TCL : $\sqrt{n}\left(\bar{X}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$.



Histogrammes de $\sqrt{n}\left(\bar{X}_n - \eta\right)$ obtenus à partir de 10000 réalisations de $\underline{X}_n$

Vitesse de convergence

Soit $\hat{\eta}_n = \hat{\eta}_n(X_1, \dots, X_n)$ un estimateur consistant de $\eta = g(\theta)$.

Définition

S'il existe une suite $(a_n)_{n\in\mathbb{N}^*}$ à valeurs réelles telle que :

- $\lim\limits_{n\to\infty} a_n = \infty$,
- $a_n\left(\hat{\eta}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} Z$,
- avec Z variable aléatoire non dégénérée*,

alors $\hat{\eta}_n$ converge vers η à la vitesse $\frac{1}{a_n}$.

* On dit que Z est dégénérée si :

- cas scalaire : $\exists c \in \mathbb{R},\ \ Z = c$ p.s. ;
- cas vectoriel : $\exists a \in \mathbb{R}^q \setminus \{0\}, \exists c \in \mathbb{R},\ \ \sum_{j=1}^q a_j Z^{(j)} = c$ p.s.

Remarque. Lorsque Z admet un moment d'ordre 2, on peut montrer que :

Z est non dégénérée ssi sa matrice de covariance est inversible. Démonstration

Vitesse de convergence

Soit $\hat{\eta}_n = \hat{\eta}_n(X_1, \ldots, X_n)$ un estimateur consistant de $\eta = g(\theta)$.

Définition

S'il existe une suite $(a_n)_{n\in\mathbb{N}^*}$ à valeurs réelles telle que :

- $\lim\limits_{n\to\infty} a_n = \infty$,
- $a_n\left(\hat{\eta}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} Z$,
- avec Z variable aléatoire non dégénérée*,

alors $\hat{\eta}_n$ converge vers η à la vitesse $\frac{1}{a_n}$.

* On dit que Z est dégénérée si :

- cas scalaire : $\exists c \in \mathbb{R},\ \ Z = c\ \ \text{p.s.}$;
- cas vectoriel : $\exists a \in \mathbb{R}^q \setminus \{0\},\ \exists c \in \mathbb{R},\ \ \sum_{j=1}^q a_j Z^{(j)} = c\ \ \text{p.s.}$

Remarque. Lorsque Z admet un moment d'ordre 2, on peut montrer que :

Z est non dégénérée ssi sa matrice de covariance est inversible. Démonstration

Normalité asymptotique

Soit $\hat{\eta}_n = \hat{\eta}_n(X_1, \dots, X_n)$ un estimateur consistant de $\eta = g(\theta)$.

Définition

S'il existe

- une suite $(a_n)_{n\in\mathbb{N}^*}$ à valeurs réelles telle que $\lim_{n\to\infty} a_n = \infty$,
- une matrice $\Sigma(\theta)$ symétrique définie positive,

telles que

$$a_n\left(\hat{\eta}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \Sigma(\theta)\right), \tag{1}$$

alors $\hat{\eta}_n$ est dit asymptotiquement normal.

Vocabulaire. $\Sigma(\theta)$ s'appelle la matrice de covariance asymptotique (variance asymptotique dans le cas scalaire).

Note : on peut m.q. (1) avec $a_n \to +\infty$ implique la consistance (faible).

Convergence en loi vers une constante

Normalité asymptotique

Soit $\hat{\eta}_n = \hat{\eta}_n(X_1, \ldots, X_n)$ un estimateur consistant de $\eta = g(\theta)$.

Définition

S'il existe

- une suite $(a_n)_{n \in \mathbb{N}^*}$ à valeurs réelles telle que $\lim_{n \to \infty} a_n = \infty$,
- une matrice $\Sigma(\theta)$ symétrique définie positive,

telles que

$$a_n \left(\hat{\eta}_n - \eta\right) \xrightarrow[n \to \infty]{\text{loi}} \mathcal{N}\left(0, \Sigma(\theta)\right), \tag{1}$$

alors $\hat{\eta}_n$ est dit asymptotiquement normal.

Vocabulaire. $\Sigma(\theta)$ s'appelle la matrice de covariance asymptotique (variance asymptotique dans le cas scalaire).

Note : on peut m.q. (1) avec $a_n \to +\infty$ implique la consistance (faible).

Convergence en loi vers une constante

Normalité asymptotique

Soit $\hat{\eta}_n = \hat{\eta}_n(X_1, \ldots, X_n)$ un estimateur consistant de $\eta = g(\theta)$.

Définition

S'il existe

- une suite $(a_n)_{n\in\mathbb{N}^*}$ à valeurs réelles telle que $\lim_{n\to\infty} a_n = \infty$,
- une matrice $\Sigma(\theta)$ symétrique définie positive,

telles que

$$a_n \left(\hat{\eta}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \Sigma(\theta)\right), \tag{1}$$

alors $\hat{\eta}_n$ est dit asymptotiquement normal.

Vocabulaire. $\Sigma(\theta)$ s'appelle la matrice de covariance asymptotique (variance asymptotique dans le cas scalaire).

Note : on peut m.q. (1) avec $a_n \to +\infty$ implique la consistance (faible).

Convergence en loi vers une constante

Plan du cours

Théorème de continuité

Théorème (Mann-Wald)

Soient

- $h : \mathbb{R}^d \to \mathbb{R}^q$ une fonction mesurable,
- Y une VA à valeurs dans $\mathbb{R}^d$,

telles que

$$\boxed{h \text{ est continue au point } Y\text{, presque sûrement.}}$$

Alors, pour toute suite $(Y_n)_{n\in\mathbb{N}^*}$ de VA à valeurs dans $\mathbb{R}^d$,

$$\begin{array}{llcl}
\text{(i)} & Y_n \xrightarrow{\text{ps}} Y & \Rightarrow & h(Y_n) \xrightarrow{\text{ps}} h(Y), \\
\text{(ii)} & Y_n \xrightarrow{\mathbb{P}} Y & \Rightarrow & h(Y_n) \xrightarrow{\mathbb{P}} h(Y), \\
\text{(iii)} & Y_n \xrightarrow{\text{loi}} Y & \Rightarrow & h(Y_n) \xrightarrow{\text{loi}} h(Y).
\end{array}$$

Démonstration : cf. CIP pour le cas où h est continu. Le cas général est admis.

Exemple « fiabilité composant » (suite)

Rappels :

- $X_i \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$, $\theta > 0$, et $\eta = \mathbb{E}_\theta(X_1) = \frac{1}{\theta}$.
- $\hat{\eta}_n = \bar{X}_n$ est obtenu par MV et méthode des moments.

Loi des grands nombres (forte et en m.q.) :

$$\hat{\eta}_n = \bar{X}_n \xrightarrow{\text{ps}, L^2} \eta.$$

Par le théorème de continuité :

$$\hat{\theta}_n = \frac{1}{\hat{\eta}_n} \xrightarrow{\text{ps}} \frac{1}{\eta} = \theta,$$

donc $\hat{\theta}_n$ est fortement consistant pour l'estimation de θ.

Remarque : on peut montrer que $\hat{\theta}_n$ est également consistant dans L^2.

Exemple « fiabilité composant » (suite)

Rappels :

- $X_i \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$, $\theta > 0$, et $\quad \eta = \mathbb{E}_\theta(X_1) = \frac{1}{\theta}$.
- $\hat{\eta}_n = \bar{X}_n$ est obtenu par MV et méthode des moments.

Loi des grands nombres (forte et en m.q.) :

$$\hat{\eta}_n = \bar{X}_n \xrightarrow{\text{ps}, L^2} \eta.$$

Par le théorème de continuité :

$$\hat{\theta}_n = \frac{1}{\hat{\eta}_n} \xrightarrow{\text{ps}} \frac{1}{\eta} = \theta,$$

donc $\hat{\theta}_n$ est fortement consistant pour l'estimation de θ.

Remarque : on peut montrer que $\hat{\theta}_n$ est également consistant dans L^2.

Théorème de Slutsky

Théorème

Soient

- $(X_n)_{n\in\mathbb{N}^*}$ une suite de vecteurs aléatoires qui converge en loi vers une VA X :
$$X_n \xrightarrow[n\to\infty]{\text{loi}} X,$$
- $(Y_n)_{n\in\mathbb{N}^*}$ une suite de vecteurs aléatoires qui converge en loi vers une constante c :
$$Y_n \xrightarrow[n\to\infty]{\text{loi}} c,$$

Alors
$$(X_n, Y_n) \xrightarrow[n\to\infty]{\text{loi}} (X, c).$$

Remarque : on peut m.q. $Y_n \xrightarrow[n\to\infty]{\text{loi}} c$ implique $Y_n \xrightarrow[n\to\infty]{\mathbb{P}} c$ (limite constante).

Théorème de Slutsky

Théorème

Soient

- $(X_n)_{n\in\mathbb{N}^*}$ une suite de vecteurs aléatoires qui converge en loi vers une VA X :
$$X_n \xrightarrow[n\to\infty]{\text{loi}} X,$$
- $(Y_n)_{n\in\mathbb{N}^*}$ une suite de vecteurs aléatoires qui converge en loi vers une constante c :
$$Y_n \xrightarrow[n\to\infty]{\text{loi}} c,$$

Alors
$$(X_n, Y_n) \xrightarrow[n\to\infty]{\text{loi}} (X, c).$$

Remarque : on peut m.q. $Y_n \xrightarrow[n\to\infty]{\text{loi}} c$ implique $Y_n \xrightarrow[n\to\infty]{\mathbb{P}} c$ (limite constante).

Exemple « fiabilité composant » (suite)

Rappel (TCL) $\sqrt{n}\left(\bar{X}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$.

Puisque $\bar{X}_n \xrightarrow[n\to\infty]{\text{ps}} \eta$ (constante), on a par le théorème de Slutsky :

$$\left(\sqrt{n}\left(\bar{X}_n - \eta\right), \bar{X}_n\right) \xrightarrow[n\to\infty]{\text{loi}} (Z, \eta) \quad \text{avec } Z \sim \mathcal{N}\left(0, \eta^2\right).$$

Donc, par le théorème de continuité,

$$\sqrt{n}\,\frac{\left(\bar{X}_n - \eta\right)}{\bar{X}_n} \xrightarrow[n\to\infty]{\text{loi}} \frac{Z}{\eta} \sim \mathcal{N}(0, 1),$$

puisque $(z, y) \mapsto \frac{z}{y}$ est continue en tout point où $y \neq 0$.

Remarque. Résultat utilisé pour construire un IC asymptotique exercice 4

Exemple « fiabilité composant » (suite)

Rappel (TCL) $\sqrt{n}\left(\bar{X}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$.

Puisque $\bar{X}_n \xrightarrow[n\to\infty]{\text{ps}} \eta$ (constante), on a par le théorème de Slutsky :

$$\left(\sqrt{n}\left(\bar{X}_n - \eta\right), \bar{X}_n\right) \xrightarrow[n\to\infty]{\text{loi}} (Z, \eta) \quad \text{avec } Z \sim \mathcal{N}\left(0, \eta^2\right).$$

Donc, par le théorème de continuité,

$$\sqrt{n}\,\frac{\left(\bar{X}_n - \eta\right)}{\bar{X}_n} \xrightarrow[n\to\infty]{\text{loi}} \frac{Z}{\eta} \sim \mathcal{N}\left(0, 1\right),$$

puisque $(z, y) \mapsto \frac{z}{y}$ est continue en tout point où $y \neq 0$.

Remarque. Résultat utilisé pour construire un IC asymptotique exercice 4

Exemple « fiabilité composant » (suite)

Rappel (TCL) $\sqrt{n}\left(\bar{X}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$.

Puisque $\bar{X}_n \xrightarrow[n\to\infty]{\text{ps}} \eta$ (constante), on a par le théorème de Slutsky :

$$\left(\sqrt{n}\left(\bar{X}_n - \eta\right), \bar{X}_n\right) \xrightarrow[n\to\infty]{\text{loi}} (Z, \eta) \quad \text{avec } Z \sim \mathcal{N}\left(0, \eta^2\right).$$

Donc, par le théorème de continuité,

$$\sqrt{n}\,\frac{\left(\bar{X}_n - \eta\right)}{\bar{X}_n} \xrightarrow[n\to\infty]{\text{loi}} \frac{Z}{\eta} \sim \mathcal{N}(0,1),$$

puisque $(z, y) \mapsto \frac{z}{y}$ est continue en tout point où $y \neq 0$.

Remarque. Résultat utilisé pour construire un IC asymptotique

exercice 4

Méthode de linéarisation (« delta méthode »)

« Delta théorème » (cas **scalaire**)

Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de VA à valeurs dans $\mathbb{R}$, t.q.

$$\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} Z,$$

avec Z une VA à valeurs dans $\mathbb{R}$ et $m \in \mathbb{R}$.

Alors, pour toute fonction $h : \mathbb{R} \to \mathbb{R}$ dérivable en m,

$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} h'(m)\,Z,$$

Intuition : $h(y) - h(m) \approx h'(m)\,(y - m)$.

Démonstration

Méthode de linéarisation (« delta méthode »)

« Delta théorème » (cas **scalaire**)

Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de VA à valeurs dans $\mathbb{R}$, t.q.

$$\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} Z,$$

avec Z une VA à valeurs dans $\mathbb{R}$ et $m \in \mathbb{R}$.

Alors, pour toute fonction $h : \mathbb{R} \to \mathbb{R}$ dérivable en m,

$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} h'(m)\,Z,$$

Intuition : $h(y) - h(m) \approx h'(m)\,(y-m)$.

Démonstration

Méthode de linéarisation (« delta méthode »)

« Delta théorème » (cas **scalaire**)

Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de VA à valeurs dans $\mathbb{R}$, t.q.

$$\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} Z,$$

avec Z une VA à valeurs dans $\mathbb{R}$ et $m \in \mathbb{R}$.

Alors, pour toute fonction $h : \mathbb{R} \to \mathbb{R}$ dérivable en m,

$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} h'(m)\,Z,$$

Intuition : $h(y) - h(m) \approx h'(m)\,(y - m)$.

Démonstration

Méthode de linéarisation (« delta méthode »)

« Delta théorème » (cas **vectoriel**)

Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de VA à valeurs dans $\mathbb{R}^d$, t.q.

$$\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} Z,$$

avec Z une VA à valeurs dans $\mathbb{R}^d$ et $m \in \mathbb{R}^d$.

Alors, pour toute fonction $h : \mathbb{R}^d \to \mathbb{R}^q$ différentiable en m,

$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} (Dh)(m)\,Z,$$

avec $(Dh)(m)$ matrice jacobienne de h au point m :

$$(Dh)(m) = \Big((\partial_j h_i)(m)\Big)_{1\le i\le q,\ 1\le j\le d}.$$

Méthode de linéarisation (« delta méthode »)

« Delta théorème » (cas **vectoriel**)

Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de VA à valeurs dans $\mathbb{R}^d$, t.q.

$$\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} Z,$$

avec Z une VA à valeurs dans $\mathbb{R}^d$ et $m \in \mathbb{R}^d$.

Alors, pour toute fonction $h : \mathbb{R}^d \to \mathbb{R}^q$ différentiable en m,

$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} (Dh)(m)\,Z,$$

avec $(Dh)(m)$ matrice jacobienne de h au point m :

$$(Dh)(m) = \Big((\partial_j h_i)(m)\Big)_{1\le i\le q,\ 1\le j\le d}.$$

Delta théorème dans le cas gaussien

Cas scalaire.
Si $\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,\, \sigma^2)$, alors
$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0,\; (h'(m))^2\sigma^2\right).$$

Cas vectoriel

Si $\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,\, \Sigma)$, alors
$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0,\; (Dh)(m)\,\Sigma\,(Dh)(m)^\top\right).$$

Delta théorème dans le cas gaussien

Cas scalaire.
Si $\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,\, \sigma^2)$, alors

$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0,\ (h'(m))^2\sigma^2\ \right).$$

Cas vectoriel

Si $\sqrt{n}\,(Y_n - m) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,\, \Sigma)$, alors

$$\sqrt{n}\,(h(Y_n) - h(m)) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0,\ (Dh)(m)\ \Sigma\ (Dh)(m)^\top\right).$$

Exemple : « fiabilité composant ».

On a déja vu que :

- $\hat{\theta}_n = 1/\bar{X}_n$ est un estimateur fortement consistant de θ,
- $\sqrt{n}\left(\bar{X}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$, où $\eta = \frac{1}{\theta}$.

Delta-méthode avec $h(\eta) = \frac{1}{\eta}$

$$\sqrt{n}\left(\frac{1}{\bar{X}_n} - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0,\, \eta^2\left(h'(\eta)\right)^2\right),$$

$$h'(\eta) = -\frac{1}{\eta^2} \quad \Longrightarrow \quad \sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \theta^2\right).$$

Conclusion : $\hat{\theta}_n$ est asymptotiquement gaussien,

et sa vitesse de convergence est $\frac{1}{\sqrt{n}}$.

Exemple : « fiabilité composant ».

On a déja vu que :

- $\hat{\theta}_n = 1/\bar{X}_n$ est un estimateur fortement consistant de θ,
- $\sqrt{n}\left(\bar{X}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$, où $\eta = \frac{1}{\theta}$.

Delta-méthode avec $h(\eta) = \frac{1}{\eta}$

$$\sqrt{n}\left(\frac{1}{\bar{X}_n} - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0,\ \eta^2\left(h'(\eta)\right)^2\right),$$

$$h'(\eta) = -\frac{1}{\eta^2} \quad \Longrightarrow \quad \sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \theta^2\right).$$

Conclusion : $\hat{\theta}_n$ est asymptotiquement gaussien,

et sa vitesse de convergence est $\frac{1}{\sqrt{n}}$.

Exemple : « fiabilité composant ».

On a déja vu que :

- $\hat{\theta}_n = 1/\bar{X}_n$ est un estimateur fortement consistant de θ,
- $\sqrt{n}\left(\bar{X}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$, où $\eta = \frac{1}{\theta}$.

Delta-méthode avec $h(\eta) = \frac{1}{\eta}$

$$\sqrt{n}\left(\frac{1}{\bar{X}_n} - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0,\ \eta^2\left(h'(\eta)\right)^2\right),$$

$$h'(\eta) = -\frac{1}{\eta^2} \quad \Longrightarrow \quad \sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \theta^2\right).$$

Conclusion : $\hat{\theta}_n$ est asymptotiquement gaussien,

et sa vitesse de convergence est $\frac{1}{\sqrt{n}}$.

Comparaison asymptotique d'estimateurs (scalaires) (1/2)

A l'aide des **variances asymptotiques.**
Illustration sur l'exemple « fiabilité composant » pour $\eta = \mathbb{E}_\theta(X_1)$.

1) Pour $\hat{\eta}^{(1)} = \bar{X}_n$, on a (TCL) : $\sqrt{n}\left(\hat{\eta}^{(1)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$.

2) Pour $\hat{\eta}^{(2)} = \sqrt{\frac{1}{2n}\sum_{i=1}^n X_i^2}$ (cf. cours 1) ?

- Comme $\mathbb{E}\left(X_1^2\right) = 2\eta^2$ et $\mathbb{E}\left(X_1^4\right) = 24\eta^4$, on a (TCL) :

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - 2\eta^2\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, 20\,\eta^4\right).$$

- D'où, en utilisant la delta méthode avec $h(z) = \sqrt{\frac{1}{2}z}$,

$$\sqrt{n}\left(\hat{\eta}^{(2)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \frac{5}{4}\eta^2\right).$$

Comparaison asymptotique d'estimateurs (scalaires) (1/2)

A l'aide des **variances asymptotiques.**
Illustration sur l'exemple « fiabilité composant » pour $\eta = \mathbb{E}_\theta(X_1)$.

1) Pour $\hat{\eta}^{(1)} = \bar{X}_n$, on a (TCL) : $\sqrt{n}\left(\hat{\eta}^{(1)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right)$.

2) Pour $\hat{\eta}^{(2)} = \sqrt{\frac{1}{2n}\sum_{i=1}^n X_i^2}$ (cf. cours 1) ?

- Comme $\mathbb{E}\left(X_1^2\right) = 2\eta^2$ et $\mathbb{E}\left(X_1^4\right) = 24\eta^4$, on a (TCL) :

$$\sqrt{n}\Big(\frac{1}{n}\sum_{i=1}^n X_i^2 - 2\eta^2\Big) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, 20\,\eta^4\right).$$

- D'où, en utilisant la delta méthode avec $h(z) = \sqrt{\frac{1}{2}z}$,

$$\sqrt{n}\left(\hat{\eta}^{(2)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \frac{5}{4}\eta^2\right).$$

Comparaison asymptotique d'estimateurs (scalaires) (2/2)

En résumé :

$$\sqrt{n}\left(\hat{\eta}^{(1)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right),$$
$$\sqrt{n}\left(\hat{\eta}^{(2)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \frac{5}{4}\eta^2\right).$$

On observe que

- les deux estimateurs sont asymptotiquement gaussiens,
- ont la même vitesse de convergence,
- mais la variance asymptotique de $\hat{\eta}^{(1)}$ est plus faible.

➠ On dit que $\hat{\eta}^{(1)}$ est asymptotiquement préférable à $\hat{\eta}^{(2)}$.

Comparaison asymptotique d'estimateurs (scalaires) (2/2)

En résumé :

$$\sqrt{n}\left(\hat{\eta}^{(1)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right),$$
$$\sqrt{n}\left(\hat{\eta}^{(2)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \frac{5}{4}\eta^2\right).$$

On observe que

- les deux estimateurs sont asymptotiquement gaussiens,
- ont la même vitesse de convergence,
- mais la variance asymptotique de $\hat{\eta}^{(1)}$ est plus faible.

➠ On dit que $\hat{\eta}^{(1)}$ est asymptotiquement préférable à $\hat{\eta}^{(2)}$.

Comparaison asymptotique d'estimateurs (scalaires) (2/2)

En résumé :

$$\sqrt{n}\left(\hat{\eta}^{(1)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right),$$
$$\sqrt{n}\left(\hat{\eta}^{(2)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \frac{5}{4}\eta^2\right).$$

On observe que

- les deux estimateurs sont asymptotiquement gaussiens,
- ont la même vitesse de convergence,
- mais la variance asymptotique de $\hat{\eta}^{(1)}$ est plus faible.

➠ On dit que $\hat{\eta}^{(1)}$ est asymptotiquement préférable à $\hat{\eta}^{(2)}$.

Comparaison asymptotique d'estimateurs (scalaires) (2/2)

En résumé :

$$\sqrt{n}\left(\hat{\eta}^{(1)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right),$$
$$\sqrt{n}\left(\hat{\eta}^{(2)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \frac{5}{4}\eta^2\right).$$

On observe que

- les deux estimateurs sont asymptotiquement gaussiens,
- ont la même vitesse de convergence,
- mais la variance asymptotique de $\hat{\eta}^{(1)}$ est plus faible.

➠ On dit que $\hat{\eta}^{(1)}$ est asymptotiquement préférable à $\hat{\eta}^{(2)}$.

Comparaison asymptotique d'estimateurs (scalaires) (2/2)

En résumé :

$$\sqrt{n}\left(\hat{\eta}^{(1)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta^2\right),$$
$$\sqrt{n}\left(\hat{\eta}^{(2)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \frac{5}{4}\eta^2\right).$$

On observe que

- les deux estimateurs sont asymptotiquement gaussiens,
- ont la même vitesse de convergence,
- mais la variance asymptotique de $\hat{\eta}^{(1)}$ est plus faible.

➠ On dit que $\hat{\eta}^{(1)}$ est asymptotiquement préférable à $\hat{\eta}^{(2)}$.

Plan du cours

Efficacité asymptotique

Rappel (borne de Cramér-Rao pour un paramètre scalaire) :

$\forall \hat{\theta}$ ESB régulier de θ, $\forall \theta \in \Theta$,

$$R_\theta\left(\hat{\theta}\right) = \mathrm{var}_\theta\left(\hat{\theta}\right) \geq \frac{1}{n} I_1^{-1}(\theta),$$

avec $I_1(\theta) = \mathrm{var}_\theta\left(S_\theta(X_1)\right)$.

➠ Si l'égalité est atteinte, alors $\hat{\theta}$ est dit efficace.

Efficacité asymptotique

Définition. Un estimateur est dit asymptotiquement efficace si

- il est asymptotiquement normal à la vitesse $\frac{1}{\sqrt{n}}$,
- avec pour variance asymptotique $I_1^{-1}(\theta)$.

Remarque : définition valable également dans le cas vectoriel, en remplaçant la variance par la matrice de covariance.

Efficacité asymptotique

Rappel (borne de Cramér-Rao pour un paramètre scalaire) :

$\forall \hat{\theta}$ ESB régulier de θ, $\forall \theta \in \Theta$,

$$R_\theta\left(\hat{\theta}\right) = \mathrm{var}_\theta\left(\hat{\theta}\right) \geq \frac{1}{n} I_1^{-1}(\theta),$$

avec $I_1(\theta) = \mathrm{var}_\theta\left(S_\theta(X_1)\right)$.

➠ Si l'égalité est atteinte, alors $\hat{\theta}$ est dit efficace.

Efficacité asymptotique

Définition. Un estimateur est dit asymptotiquement efficace si

- il est asymptotiquement normal à la vitesse $\frac{1}{\sqrt{n}}$,
- avec pour variance asymptotique $I_1^{-1}(\theta)$.

Remarque : définition valable également dans le cas vectoriel, en remplaçant la variance par la matrice de covariance.

Efficacité asymptotique de l'EMV

Contexte : $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathrm{P}_\theta$ et, $\forall\theta \in \Theta$, P_θ admet une densité f_θ.

Définition : modèle régulier

Le modèle statistique est dit régulier si

- les conditions C_0–C_2 sont vérifiées (déf. données au cours 2)
- les conditions C_3 & C_4 sont vérifiées

Conditions C_3 & C_4

- $\forall\theta \in \Theta$, l'information de Fisher $I_1(\theta)$ est définie positive.

Théorème

Si le modèle statistique est régulier et si l'EMV $\hat{\theta}_n$ est consistant, alors il est asymptotiquement efficace :

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, I_1^{-1}(\theta)\right).$$

Contre-exemple avec un modèle non-régulier

Efficacité asymptotique de l'EMV

Contexte : $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathrm{P}_\theta$ et, $\forall \theta \in \Theta$, P_θ admet une densité f_θ.

Définition : modèle régulier

Le modèle statistique est dit régulier si

- les conditions C_0–C_2 sont vérifiées (déf. données au cours 2)
- les conditions C_3 & C_4 sont vérifiées Conditions C_3 & C_4
- $\forall \theta \in \Theta$, l'information de Fisher $I_1(\theta)$ est définie positive.

Théorème

Si le modèle statistique est régulier et si l'EMV $\hat{\theta}_n$ est consistant, alors il est asymptotiquement efficace :

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, I_1^{-1}(\theta)\right).$$

Contre-exemple avec un modèle non-régulier

Information de Fisher dans les modèles réguliers

Rappel. L'information de Fisher apportée par $\underline{X}$ est la matrice

$$I(\theta) = \mathrm{var}_\theta(S_\theta) = \mathbb{E}_\theta\left(S_\theta\, S_\theta^\top\right).$$

Proposition : autre expression de l'information de Fisher

Dans un modèle régulier, on a l'égalité

$$I(\theta) = -\,\mathbb{E}_\theta\left(\nabla_\theta\left(S_\theta^\top\right)\right), \qquad (\star)$$

Autrement dit : $\forall \theta \in \Theta$, $\forall j \leq p$, $\forall k \leq p$,

$$\left(I(\theta)\right)_{j,k} = -\,\mathbb{E}_\theta\left(\frac{\partial}{\partial\theta_j} S_\theta^{(k)}\right) = -\,\mathbb{E}_\theta\left(\frac{\partial^2}{\partial\theta_j\partial\theta_k} \ln f_\theta(\underline{X})\right).$$

Remarque : en fait, si C_0–C_3 sont vérifiées, alors C_4 et $(\star)$ sont équivalents.

Information de Fisher dans les modèles réguliers

Rappel. L'information de Fisher apportée par $\underline{X}$ est la matrice

$$I(\theta) = \mathrm{var}_\theta(S_\theta) = \mathbb{E}_\theta\left(S_\theta\, S_\theta^\top\right).$$

Proposition : autre expression de l'information de Fisher

Dans un modèle régulier, on a l'égalité

$$I(\theta) = -\,\mathbb{E}_\theta\left(\nabla_\theta\left(S_\theta^\top\right)\right), \tag{$\star$}$$

Autrement dit : $\forall \theta \in \Theta$, $\forall j \leq p$, $\forall k \leq p$,

$$\left(I(\theta)\right)_{j,k} = -\,\mathbb{E}_\theta\left(\frac{\partial}{\partial \theta_j} S_\theta^{(k)}\right) = -\,\mathbb{E}_\theta\left(\frac{\partial^2}{\partial \theta_j \partial \theta_k} \ln f_\theta(\underline{X})\right).$$

Remarque : en fait, si C_0–C_3 sont vérifiées, alors C_4 et $(\star)$ sont équivalents.

Exemple : « fiabilité composant » (suite)

Question : $\hat{\theta}_n = 1/\bar{X}_n$ est-il asymptotiquement efficace ?

On a déjà calculé le score : $S_\theta(X_1) = \frac{1}{\theta} - X_1$.

Calcul de l'information de Fisher (deux approches) :

Calcul de $\mathbb{E}_\theta\left(S_\theta(X_1)^2\right)$	Calcul de $-\mathbb{E}_\theta\left(\frac{\partial S_\theta}{\partial \theta}(X_1)\right)$
$I_1(\theta) = \text{var}_\theta(X_1) = \eta^2 = \frac{1}{\theta^2}$	$I_1(\theta) = -\mathbb{E}_\theta\left(-\frac{1}{\theta^2}\right) = \frac{1}{\theta^2}$

Conclusion : puisque $\sqrt{n}\left(\frac{1}{\bar{X}_n} - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \theta^2\right)$,

$$\hat{\theta}_n = \frac{1}{\bar{X}_n} \text{ est asymptotiquement efficace.}$$

➠ On retrouve le résultat du théorème (C_0–C_4 sont vérifiées).

Exemple : « fiabilité composant » (suite)

Question : $\hat{\theta}_n = 1/\bar{X}_n$ est-il asymptotiquement efficace ?

On a déjà calculé le score : $S_\theta(X_1) = \frac{1}{\theta} - X_1$.

Calcul de l'information de Fisher (deux approches) :

Calcul de $\mathbb{E}_\theta\left(S_\theta(X_1)^2\right)$	Calcul de $-\mathbb{E}_\theta\left(\frac{\partial S_\theta}{\partial\theta}(X_1)\right)$
$I_1(\theta) = \mathrm{var}_\theta(X_1) = \eta^2 = \frac{1}{\theta^2}$	$I_1(\theta) = -\mathbb{E}_\theta\left(-\frac{1}{\theta^2}\right) = \frac{1}{\theta^2}$

Conclusion : puisque $\sqrt{n}\left(\frac{1}{\bar{X}_n} - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \theta^2\right)$,

$$\hat{\theta}_n = \frac{1}{\bar{X}_n} \text{ est asymptotiquement efficace.}$$

➠ On retrouve le résultat du théorème (C_0–C_4 sont vérifiées).

Exemple : « fiabilité composant » (suite)

Question : $\hat{\theta}_n = 1/\bar{X}_n$ est-il asymptotiquement efficace ?

On a déjà calculé le score : $S_\theta(X_1) = \frac{1}{\theta} - X_1$.

Calcul de l'information de Fisher (deux approches) :

Calcul de $\mathbb{E}_\theta\left(S_\theta(X_1)^2\right)$	Calcul de $-\mathbb{E}_\theta\left(\frac{\partial S_\theta}{\partial \theta}(X_1)\right)$
$I_1(\theta) = \mathrm{var}_\theta(X_1) = \eta^2 = \frac{1}{\theta^2}$	$I_1(\theta) = -\mathbb{E}_\theta\left(-\frac{1}{\theta^2}\right) = \frac{1}{\theta^2}$

Conclusion : puisque $\sqrt{n}\left(\frac{1}{\bar{X}_n} - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \theta^2\right)$,

$$\hat{\theta}_n = \frac{1}{\bar{X}_n} \text{ est asymptotiquement efficace.}$$

➠ On retrouve le résultat du théorème (C_0–C_4 sont vérifiées).

Plan du cours

Plan du cours

Motivation

Problème

Un estimateur ponctuel commet nécessairement une erreur d'estimation.
Comment « rendre compte » de cette erreur ?

Deux approches :

- fournir en plus de la valeur estimée,
  - la loi de l'estimateur $\hat{\eta}$, exacte ou approchée,
  - ou au moins une « mesure de dispersion » (par ex. son écart-type) ;
- donner, plutôt qu'une estimation ponctuelle $\hat{\eta}$,

$$\boxed{\text{un intervalle de confiance pour } \eta.}$$

Motivation

Problème

Un estimateur ponctuel commet nécessairement une erreur d'estimation.
Comment « rendre compte » de cette erreur ?

Deux approches :

- fournir en plus de la valeur estimée,
  - la loi de l'estimateur $\hat{\eta}$, exacte ou approchée,
  - ou au moins une « mesure de dispersion » (par ex. son écart-type) ;
- donner, plutôt qu'une estimation ponctuelle $\hat{\eta}$,

  $\boxed{\text{un intervalle de confiance pour } \eta.}$

Motivation

Problème

Un estimateur ponctuel commet nécessairement une erreur d'estimation.
Comment « rendre compte » de cette erreur ?

Deux approches :

- fournir en plus de la valeur estimée,
  - la loi de l'estimateur $\hat{\eta}$, exacte ou approchée,
  - ou au moins une « mesure de dispersion » (par ex. son écart-type) ;
- donner, plutôt qu'une estimation ponctuelle $\hat{\eta}$,

  un **intervalle de confiance** pour η.

Taux de couverture

Rappel. $\eta = g(\theta)$.

Soient

- $\mathcal{P}(N)$ l'ensemble des parties de $N = g(\Theta)$,
- $C(\underline{X})$ une statistique à valeurs dans $\mathcal{P}(N)$.

Souhait. Que $C(\underline{X})$ contienne η avec une forte probabilité.

Définition

Pour $\theta \in \Theta$, le taux de couverture de $C(\underline{X})$ pour η est défini par :

$$\mathbb{P}_\theta\left(\eta \in C\left(\underline{X}\right)\right)$$

⚠ En général, le taux de couverture dépend de la loi sous-jacente, c'est-à-dire de θ.

Taux de couverture

Rappel. $\eta = g(\theta)$.

Soient

- $\mathcal{P}(N)$ l'ensemble des parties de $N = g(\Theta)$,
- $C(\underline{X})$ une statistique à valeurs dans $\mathcal{P}(N)$.

Souhait. Que $C\left(\underline{X}\right)$ contienne η avec une forte probabilité.

Définition

Pour $\theta \in \Theta$, le taux de couverture de $C\left(\underline{X}\right)$ pour η est défini par :

$$\mathbb{P}_\theta\left(\eta \in C\left(\underline{X}\right)\right)$$

⚠ En général, le taux de couverture dépend de la loi sous-jacente, c'est-à-dire de θ.

Taux de couverture

Rappel. $\eta = g(\theta)$.

Soient

- $\mathcal{P}(N)$ l'ensemble des parties de $N = g(\Theta)$,
- $C(\underline{X})$ une statistique à valeurs dans $\mathcal{P}(N)$.

Souhait. Que $C\left(\underline{X}\right)$ contienne η avec une forte probabilité.

Définition

Pour $\theta \in \Theta$, le taux de couverture de $C\left(\underline{X}\right)$ pour η est défini par :

$$\mathbb{P}_\theta\left(\eta \in C\left(\underline{X}\right)\right)$$

⚠ En général, le taux de couverture dépend de la loi sous-jacente, c'est-à-dire de θ.

Régions et intervalles de confiance

On cherche à contrôler le taux de couverture.

Soit $\alpha \in]0,1[$.

Définition : région de confiance de niveau $1-\alpha$

Une région de confiance de niveau (au moins) $1-\alpha$ pour η est une statistique $C(\underline{X})$ à valeurs dans $\mathcal{P}(N)$ telle que :

$$\forall \theta \in \Theta, \quad \mathbb{P}_\theta(\eta \in C(\underline{X})) \;\geq\; 1-\alpha.$$

On dit que $C(\underline{X})$ est de niveau *exactement* $1-\alpha$ si

$$\forall \theta \in \Theta, \quad \mathbb{P}_\theta(\eta \in C(\underline{X})) \;=\; 1-\alpha.$$

(On dit aussi : de « taille » $1-\alpha$.)

Cas scalaire. Si $C(\underline{X})$ est un intervalle de $\mathbb{R}$, on parle d'intervalle de confiance.

Régions et intervalles de confiance

On cherche à contrôler le taux de couverture.

Soit $\alpha \in]0, 1[$.

Définition : région de confiance de **niveau** $1 - \alpha$

Une région de confiance de niveau (au moins) $1 - \alpha$ pour η est une statistique $C(\underline{X})$ à valeurs dans $\mathcal{P}(N)$ telle que :

$$\forall \theta \in \Theta, \quad \mathbb{P}_\theta\left(\eta \in C\left(\underline{X}\right)\right) \geq 1 - \alpha.$$

On dit que $C(\underline{X})$ est de niveau *exactement* $1 - \alpha$ si

$$\forall \theta \in \Theta, \quad \mathbb{P}_\theta\left(\eta \in C\left(\underline{X}\right)\right) = 1 - \alpha.$$

(On dit aussi : de « taille » $1 - \alpha$.)

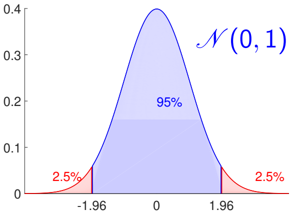
Cas scalaire. Si $C(\underline{X})$ est un intervalle de $\mathbb{R}$, on parle d'intervalle de confiance.

Exemple : n-échantillon $\mathcal{N}(\mu, \sigma_0^2)$, avec σ_0^2 connu

Comme $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma_0^2}{n}\right)$, $T = \sqrt{n}\,\frac{\bar{X}-\mu}{\sigma_0} \sim \mathcal{N}(0,1)$, donc pour $\alpha = 5\%$:

$$\mathbb{P}_\mu\left(\sqrt{n}\,\frac{\bar{X}-\mu}{\sigma_0} \in [-1.96,\, 1.96]\right) \approx 1-\alpha = 95\%,$$

où 1.96 est le quantile d'ordre 97.5% de la loi $\mathcal{N}(0,1)$. déf. : quantile



On "pivote" pour obtenir un IC de niveau exactement 95% :

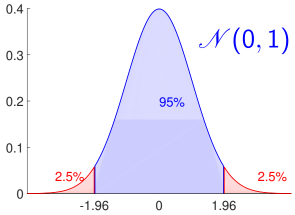
$$\sqrt{n}\,\frac{\bar{X}-\mu}{\sigma_0} \in [-1.96,\, 1.96] \quad \Leftrightarrow \quad \mu \in C(\underline{X}) = \left[\bar{X} - 1.96\,\frac{\sigma_0}{\sqrt{n}},\ \bar{X} + 1.96\,\frac{\sigma_0}{\sqrt{n}}\right].$$

Exemple : n-échantillon $\mathcal{N}(\mu, \sigma_0^2)$, avec σ_0^2 connu

Comme $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma_0^2}{n}\right)$, $T = \sqrt{n}\,\frac{\bar{X}-\mu}{\sigma_0} \sim \mathcal{N}(0,1)$, donc pour $\alpha = 5\%$:

$$\mathbb{P}_\mu\left(\sqrt{n}\,\frac{\bar{X}-\mu}{\sigma_0} \in [-1.96,\ 1.96]\right) \approx 1-\alpha = 95\%,$$

où 1.96 est le quantile d'ordre 97.5% de la loi $\mathcal{N}(0,1)$. déf. : quantile

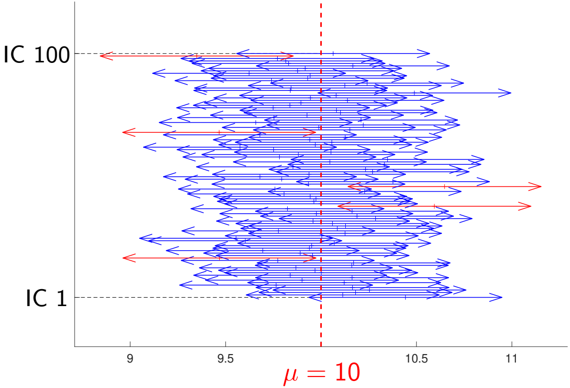


On "pivote" pour obtenir un IC de niveau exactement 95% :

$$\sqrt{n}\,\frac{\bar{X}-\mu}{\sigma_0} \in [-1.96,\ 1.96] \quad \Leftrightarrow \quad \mu \in C(\underline{X}) = \left[\bar{X} - 1.96\,\frac{\sigma_0}{\sqrt{n}},\ \bar{X} + 1.96\,\frac{\sigma_0}{\sqrt{n}}\right].$$

Interprétation : simulations

On simule 100 réalisations avec $\mu = 10$ et $\sigma_0 = 1$.



En rouge : les réalisations où l'IC ne contient pas $\mu = 10$.

➠ La proportion des cas où l'IC qui ne contient pas μ est (environ) α.

Plan du cours

Loi libre et fonction pivotale

La démarche peut être formalisée avec la notion de fonction pivotale.

Définitions

Une fonction

$$T : \underline{\mathcal{X}} \times N \;\to\; \mathbb{R}$$

est dite pivotale si la loi de la variable aléatoire $T = T(\underline{X}, \eta)$ ne dépend pas de θ. On dit que la loi de $T(\underline{X}, \eta)$ est libre.

Retour sur l'**exemple** : $X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma_0^2)$ avec σ_0 connu.

Alors $T = \sqrt{n}\,\frac{\bar{X}_n - \mu}{\sigma_0}$ est pivotale puisque

$$\sqrt{n}\,\frac{\bar{X}_n - \mu}{\sigma_0} \;\sim\; \mathcal{N}(0, 1).$$

Remarque : on peut aussi choisir $T = \sqrt{n}\,(\bar{X}_n - \mu) \sim \mathcal{N}(0, \sigma_0^2)$.

Loi libre et fonction pivotale

La démarche peut être formalisée avec la notion de fonction pivotale.

Définitions

Une fonction

$$T : \underline{\mathcal{X}} \times N \to \mathbb{R}$$

est dite pivotale si la loi de la variable aléatoire $T = T(\underline{X}, \eta)$ ne dépend pas de θ. On dit que la loi de $T(\underline{X}, \eta)$ est libre.

Retour sur l'**exemple** : $X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma_0^2)$ avec σ_0 connu.

Alors $T = \sqrt{n}\, \frac{\bar{X}_n - \mu}{\sigma_0}$ est pivotale puisque

$$\sqrt{n}\, \frac{\bar{X}_n - \mu}{\sigma_0} \sim \mathcal{N}(0, 1).$$

Remarque : on peut aussi choisir $T = \sqrt{n}\, (\bar{X}_n - \mu) \sim \mathcal{N}(0, \sigma_0^2)$.

Rappel de proba : quantiles

Définition : quantile d'ordre r

Soit $F(x)$ la fonction de répartition d'une loi sur $\mathbb{R}$.

Pour $0 < r < 1$, le quantile d'ordre r de la loi est défini par :

$$q_r = \inf\left\{x \in \mathbb{R},\ F(x) \geq r\right\} = \min\left\{x \in \mathbb{R},\ F(x) \geq r\right\}.$$

Propriétés :

- Si F est continue, alors $F(q_r) = r$.
- Si de plus F est strictement croissante, alors $q_r = F^{-1}(r)$.

Rappel de proba : quantiles

Définition : quantile d'ordre r

Soit $F(x)$ la fonction de répartition d'une loi sur $\mathbb{R}$.

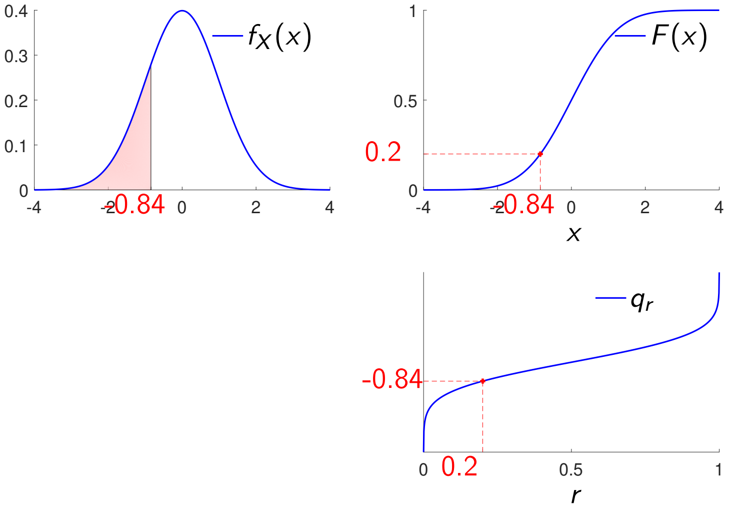
Pour $0 < r < 1$, le quantile d'ordre r de la loi est défini par :

$$q_r = \inf \{x \in \mathbb{R},\ F(x) \geq r\} = \min \{x \in \mathbb{R},\ F(x) \geq r\}.$$

Propriétés :

- Si F est continue, alors $F(q_r) = r$.
- Si de plus F est strictement croissante, alors $q_r = F^{-1}(r)$.

Fonction quantile de la loi $\mathcal{N}(0,1)$



Utilisation des fonctions pivotales

Soient $T\left(\underline{X},\eta\right)$ une fonction pivotale et $\alpha \in \left]0,1\right[$.

Proposition

Supposons la fonction de répartition F de $T\left(\underline{X},\eta\right)$ continue et strictement croissante, et notons $q_r = F^{-1}(r)$ le quantile d'ordre r.

Alors, pour tout $\gamma \in [0,\alpha]$:

$$\begin{aligned} C^{\gamma}\left(\underline{X}\right) &= \left\{\eta \in N \text{ tel que } q_{\gamma} \leq T\left(\underline{X},\eta\right) \leq q_{\gamma+1-\alpha}\right\} \\ &= T^{-1}\left(\underline{X},\, [q_{\gamma}, q_{\gamma+1-\alpha}]\right) \end{aligned}$$

est un intervalle de confiance pour η de niveau exactement $1-\alpha$.

Démonstration. $$\begin{aligned} \mathbb{P}_{\theta}\left(g(\theta) \in C^{\gamma}\left(\underline{X}\right)\right) &= \mathbb{P}_{\theta}\left(q_{\gamma} \leq T\left(\underline{X},\eta\right) \leq q_{\gamma+1-\alpha}\right) \\ &= F\left(q_{\gamma+1-\alpha}\right) - F\left(q_{\gamma}\right) = 1-\alpha \end{aligned}$$

Utilisation des fonctions pivotales

Soient $T\left(\underline{X}, \eta\right)$ une fonction pivotale et $\alpha \in]0, 1[$.

Proposition

Supposons la fonction de répartition F de $T\left(\underline{X}, \eta\right)$ continue et strictement croissante, et notons $q_r = F^{-1}(r)$ le quantile d'ordre r.

Alors, pour tout $\gamma \in [0, \alpha]$:

$$\begin{aligned} C^{\gamma}\left(\underline{X}\right) &= \left\{\eta \in N \text{ tel que } q_{\gamma} \leq T\left(\underline{X}, \eta\right) \leq q_{\gamma+1-\alpha}\right\} \\ &= T^{-1}\left(\underline{X}, \left[q_{\gamma}, q_{\gamma+1-\alpha}\right]\right) \end{aligned}$$

est un intervalle de confiance pour η de niveau exactement $1 - \alpha$.

Démonstration.

$$\begin{aligned} \mathbb{P}_{\theta}\left(g(\theta) \in C^{\gamma}\left(\underline{X}\right)\right) &= \mathbb{P}_{\theta}\left(q_{\gamma} \leq T\left(\underline{X}, \eta\right) \leq q_{\gamma+1-\alpha}\right) \\ &= F\left(q_{\gamma+1-\alpha}\right) - F\left(q_{\gamma}\right) = 1 - \alpha \end{aligned}$$

Exemple : n-échantillon $\mathcal{N}(\mu, \sigma_0^2)$, avec σ_0^2 connu

Considérons à nouveau la fonction pivotale

$$T(\underline{X}, \mu) = \sqrt{n}\, \frac{(\bar{X} - \mu)}{\sigma_0} \;\sim\; \mathcal{N}(0, 1).$$

Pour tout $\gamma \leq \alpha$, on obtient un IC de niveau (exactement) $1 - \alpha$:

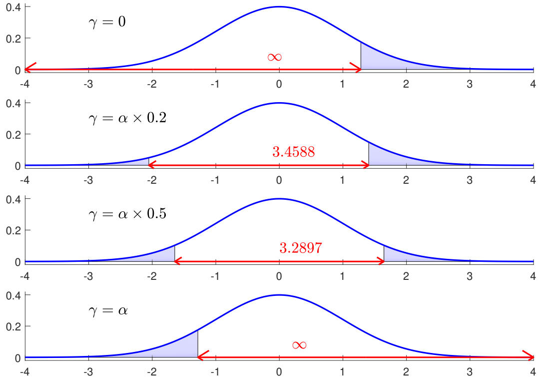
$$C^\gamma = \left[\bar{X} - \frac{\sigma_0}{\sqrt{n}}\, q_{1-\alpha+\gamma}, \quad \bar{X} - \frac{\sigma_0}{\sqrt{n}}\, q_\gamma \right],$$

avec q_r le quantile d'ordre r de la loi $\mathcal{N}(0, 1)$.

Par exemple, avec $\gamma = \frac{\alpha}{2}$ et $\alpha = 0.05$:

$$\begin{aligned} -q_{1-\alpha+\gamma} &= -q_{0.975} \approx -1.96 \\ -q_\gamma &= -q_{0.025} \approx +1.96 \end{aligned}$$

Choix du paramètre γ



Densité de la loi $\mathcal{N}(0,1)$ et quantiles associés pour $\alpha = 0.1$ et plusieurs valeurs de γ (valeurs en rouge : $q_{\gamma+1-\alpha} - q_\gamma$).

Critère usuel : valeur t.q. l'IC soit de longueur minimale (ici $\gamma = \dfrac{\alpha}{2}$).

Exemple : « fiabilité composant » (suite)

On peut montrer que :

$$T\left(\underline{X}, \eta\right) = \frac{\bar{X}}{\eta} \sim \Gamma\left(n, n\right).$$

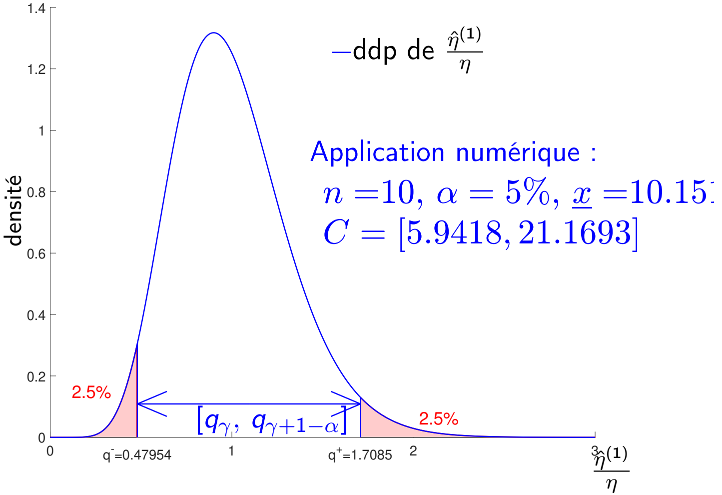
D'où un IC de niveau (exactement) $1 - \alpha$ pour η :

$$C^{\gamma} = \left[\frac{\bar{X}}{q_{\gamma+1-\alpha}}, \frac{\bar{X}}{q_{\gamma}}\right],$$

avec q_r le quantile d'ordre r de la loi $\Gamma\left(n, n\right)$.

Choix de γ : on peut prendre $\gamma = \frac{\alpha}{2}$ par simplicité, ou chercher numériquement γ tel que la longueur $1/q_{\gamma} - 1/q_{1+\gamma-\alpha}-$ soit minimale.

Exemple : « fiabilité composant » (suite)



Densité de la loi pivotale $\Gamma(n, n)$
et quantiles associés pour $\alpha = 0.05$ et $\gamma = \frac{\alpha}{2}$.

Plan du cours

Motivation et objectif

Problème

Il est parfois (souvent) difficile de trouver une fonction pivotale.

Solution : recourir à nouveau à une approche asymptotique.

- Obtention d'intervalles « approximatifs ».
- Calculs facilités grâce aux outils déjà introduits (TCL, Slutsky, delta méthode…).

⚠ Toute analyse menée dans le cadre asymptotique est approximative lorsque n est fini.

➠ Les résultats obtenus peuvent être mauvais pour n petit…

Motivation et objectif

Problème

Il est parfois (souvent) difficile de trouver une fonction pivotale.

Solution : recourir à nouveau à une approche asymptotique.

- Obtention d'intervalles « approximatifs ».
- Calculs facilités grâce aux outils déjà introduits (TCL, Slutsky, delta méthode...).

⚠ Toute analyse menée dans le cadre asymptotique est approximative lorsque n est fini.

➠ Les résultats obtenus peuvent être mauvais pour n petit...

Motivation et objectif

Problème

Il est parfois (souvent) difficile de trouver une fonction pivotale.

Solution : recourir à nouveau à une approche asymptotique.

- Obtention d'intervalles « approximatifs ».
- Calculs facilités grâce aux outils déjà introduits (TCL, Slutsky, delta méthode...).

⚠ Toute analyse menée dans le cadre asymptotique est

approximative lorsque n est fini.

➠ Les résultats obtenus peuvent être mauvais pour n petit...

Région (intervalle) de confiance asymptotique

On note $\underline{X}_n = (X_1, \ldots, X_n)$. Rappel : $\eta = g(\theta)$ et $N = g(\Theta)$.

Définition : région de confiance asymptotique

Une région de confiance asymptotique de niveau (au moins) $1 - \alpha$ est une statistique $C_n(\underline{X}_n)$ à valeurs dans $\mathcal{P}(N)$, telle que

$$\forall \theta \in \Theta, \quad \lim_{n \to \infty} \mathbb{P}_\theta\left(g(\theta) \in C_n\left(\underline{X}_n\right)\right) \geq 1 - \alpha.$$

(variante : « exactement » si égalité pour tout θ.)

Rappel : pour une RC « exacte » de niveau (au moins) $1 - \alpha$,

$$\forall \theta \in \Theta, \quad \mathbb{P}_\theta\left(g(\theta) \in C_n\left(\underline{X}_n\right)\right) \geq 1 - \alpha$$

(ici « exacte » signifie « non asymptotique »).

Construction d'intervalles de confiance asymptotique

On a recours aux **fonctions pivotales asymptotiques**.

Leur utilisation est illustrée pour :

- le paramètre d'une loi de Rayleigh — exercice 3

  Il s'agit d'un exercice mêlant **définitions** et questions.

- l'application « fiabilité composant » — exercice 4

  On montre que

$$C_n = \left[\left(1 - \frac{1}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}\right) \bar{X}_n,\ \left(1 + \frac{1}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}\right) \bar{X}_n \right]$$

  est un IC asymptotique de niveau $1-\alpha$ pour η où q_r est le quantile d'ordre r de la loi $\mathcal{N}(0,1)$.

⚠ *La construction d'IC asymptotique est au programme du cours (et de l'examen).*

Construction d'intervalles de confiance asymptotique

On a recours aux **fonctions pivotales asymptotiques**.

Leur utilisation est illustrée pour :

- le paramètre d'une loi de Rayleigh — exercice 3

  Il s'agit d'un exercice mêlant **définitions** et questions.

- l'application « fiabilité composant » — exercice 4

  On montre que

  $$C_n = \left[\left(1 - \frac{1}{\sqrt{n}} q_{1-\frac{\alpha}{2}}\right) \bar{X}_n, \left(1 + \frac{1}{\sqrt{n}} q_{1-\frac{\alpha}{2}}\right) \bar{X}_n \right]$$

  est un IC asymptotique de niveau $1-\alpha$ pour η où q_r est le quantile d'ordre r de la loi $\mathcal{N}(0,1)$.

⚠ *La construction d'IC asymptotique est au programme du cours (et de l'examen).*

Construction d'intervalles de confiance asymptotique

On a recours aux **fonctions pivotales asymptotiques**.

Leur utilisation est illustrée pour :

- le paramètre d'une loi de Rayleigh exercice 3

  Il s'agit d'un exercice mêlant **définitions** et questions.

- l'application « fiabilité composant » exercice 4

  On montre que

$$C_n = \left[\left(1 - \frac{1}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}\right) \bar{X}_n,\ \left(1 + \frac{1}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}\right) \bar{X}_n \right]$$

  est un IC asymptotique de niveau $1-\alpha$ pour η où q_r est le quantile d'ordre r de la loi $\mathcal{N}(0,1)$.

⚠ *La construction d'IC asymptotique est au programme du cours (et de l'examen).*

Conclusion du cours et transition vers la séquence suivante

Nous avons vu et développerons en TD 3 :

- des outils pour établir la convergence en loi et la vitesse de convergence d'une suite d'estimateurs,
- l'utilisation de la loi (asymptotique) d'une suite d'estimateurs pour construire des intervalles ou régions de confiance.

Nous traiterons dans la séquence 4 :

- de la prise de décision par un test d'hypothèses statistique,
- de la construction d'un tel test,
- des risques associés à une telle décision.

Conclusion du cours et transition vers la séquence suivante

Nous avons vu et développerons en TD 3 :

- des outils pour établir la convergence en loi et la vitesse de convergence d'une suite d'estimateurs,
- l'utilisation de la loi (asymptotique) d'une suite d'estimateurs pour construire des intervalles ou régions de confiance.

Nous traiterons dans la séquence 4 :

- de la prise de décision par un test d'hypothèses statistique,
- de la construction d'un tel test,
- des risques associés à une telle décision.

Plan du cours

Plan du cours

Exercice 1 (Estimation de la probabilité d'un évènement)

corrigé

Soit $(X_n)_{n\geq 1}$ une suite de VA iid à valeurs dans $(\mathcal{X}, \mathscr{A})$.

Pour un $A \in \mathscr{A}$ donné, on estime $\eta = \mathbb{P}(X_1 \in A)$ à l'aide de la **fonction indicatrice** :

$$\hat{\eta}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{X_i \in A}.$$

Question

Etudier le comportement asymptotique de $\hat{\eta}_n$.

Exercice 2 (Loi asymptotique)

corrigé

Soient $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$, avec $\theta > 0$.

Soit η la probabilité de dépasser un seuil $x_0 > 0$ donné :

$$\eta = \mathbb{P}_\theta(X \geq x_0) = \exp(-\theta x_0) .$$

Questions

1. Étudier le comportement asymptotique de la moyenne empirique $\bar{X}_n$.
2. Proposer un estimateur $\hat{\eta}_n^{(1)}$ fonction de $\bar{X}_n$, par substitution.
3. Étudier le comportement asymptotique de $\hat{\eta}_n^{(1)}$.
4. Soit $\hat{\eta}_n^{(2)} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{X_i \geq x_0}$. L'un des deux estimateurs est-il asymptotiquement préférable à l'autre ?

Exercice 3 (Loi de Rayleigh : IC asymptotique)

corrigé

Il s'agit d'un exercice long qui aborde la notion **d'intervalle de confiance asymptotique**.

Soient $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathscr{R}\left(\sigma^2\right)$, avec $\sigma^2 > 0$.

Loi de Rayleigh

Les questions ❶-❸ détaillent l'obtention d'IC asymptotique à l'aide de **fonctions pivotales asymptotiques.**

Les questions ❹-❺ illustrent la notion de **taux de couverture** appliquée au contexte des IC asymptotiques.

corrigé

Définition

Une (suite de) fonction(s)

$$T_n : \mathcal{X}^n \times N \rightarrow \mathbb{R}$$

est une fonction pivotale asymptotique si la loi limite de $T_n(\underline{X}_n, \eta)$ ne dépend pas de θ :

$$T_n(\underline{X}_n, \eta) \xrightarrow[n\to\infty]{\text{loi}} T_\infty.$$

où T_∞ est une VA dont la loi est libre.

Définition donnée avec les notations du cours

➠ Dans le cadre de l'exercice, $\eta = \theta = \sigma^2$.

Exercice 3 (Loi de Rayleigh : IC asymptotique)

➠ corrigé

Utilisation des fonctions pivotales asymptotiques :

- ➠ Identique à celle des fonctions pivotales dans le cas exact !
- ➠ Les intervalles obtenus sont des intervalles de confiance asymptotique.

Questions

1. Déterminer la loi asymptotique de $\bar{X}_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$
2. En utilisant la loi asymptotique de $\bar{X}_n$, proposer une fonction pivotale asymptotique,
3. en déduire un intervalle de confiance asymptotique de niveau $1-\alpha$.

Exercice 3 (Loi de Rayleigh : IC asymptotique)

➠ corrigé

Rappel : le **taux de couverture** d'un IC est son niveau **réel**.

Le calcul du taux de couverture de $C_n\left(\underline{X}_n\right)$ nécessite de manipuler la **Fonction de répartition (FR) de** T_n.

Ici, T_n dépend de $\bar{X}_n$ dont la loi n'est pas une loi standard.
➠ sa FR est cependant accessible numériquement.

Questions

4. Montrer que $\displaystyle \frac{1}{\sigma}\sum_{i=1}^{n} X_i \sim \mathcal{SR}(n,1)$
   où $F^{(n)}$ est la FR de la loi $\mathcal{SR}(n,1)$. ➠ Somme de lois de Rayleigh
5. Exprimer le taux de couverture de $C_n\left(\underline{X}_n\right)$ à l'aide de $F^{(n)}$.

Exercice 4 (IC asymptotique Fiabilité composant)

corrigé

Contexte de l'application « Fiabilité composant »

➠ $(X_1, \ldots, X_n) \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$ et $\eta = \dfrac{1}{\theta}$

Questions

1. Montrer que

$$T_n = \sqrt{n}\,\frac{(\bar{X}_n - \eta)}{\bar{X}_n} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1).$$

est une fonction pivotale asymptotique (voir exercice 3 pour une définition de ce terme).

2. Utiliser cette fonction pivotale pour construire un IC asymptotique de niveau $1-\alpha$.
3. Exprimer la couverture de l'IC asymptotique obtenu.

retour au slide 38

Plan du cours

Corrigé de l'exercice 1

retour au sujet

En appliquant le TCL à $Y_i = \mathbb{1}_{X_i \in A} \overset{\text{iid}}{\sim} \mathrm{Ber}(\eta)$:

$$\sqrt{n}\left(\hat{\eta}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta(1-\eta)\right).$$

Concl. : si $0 < \eta < 1$, alors $\hat{\eta}_n$ est asymptotiquement gaussien, avec

- vitesse de convergence : $\frac{1}{\sqrt{n}}$,
- variance asymptotique : $\eta(1-\eta)$.

Corrigé de l'exercice 1

retour au sujet

En appliquant le TCL à $Y_i = \mathbb{1}_{X_i \in A} \overset{\text{iid}}{\sim} \mathrm{Ber}(\eta)$:

$$\sqrt{n}\left(\hat{\eta}_n - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta(1-\eta)\right).$$

Concl. : si $0 < \eta < 1$, alors $\hat{\eta}_n$ est asymptotiquement gaussien, avec

- vitesse de convergence : $\frac{1}{\sqrt{n}}$,
- variance asymptotique : $\eta(1-\eta)$.

En appliquant le TCL à $Y_i = \mathbb{1}_{X_i \in A} \overset{\text{iid}}{\sim} \mathrm{Ber}(\eta)$:

$$\sqrt{n}\,(\hat{\eta}_n - \eta) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\,(0, \eta(1-\eta))\,.$$

Concl. : si $0 < \eta < 1$, alors $\hat{\eta}_n$ est asymptotiquement gaussien, avec

- vitesse de convergence : $\frac{1}{\sqrt{n}}$,
- variance asymptotique : $\eta(1-\eta)$.

Corrigé de l'exercice 2

retour au sujet

❶ Appliquant le TCL :

$$\sqrt{n}\left(\bar{X}_n - \frac{1}{\theta}\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \frac{1}{\theta^2}\right)$$

❷ $\eta = \exp\left(-\dfrac{x_0}{\frac{1}{\theta}}\right) = h\left(\frac{1}{\theta}\right)$

avec $h : u \mapsto \exp\left(-\dfrac{x_0}{u}\right)$ continue sur $\mathbb{R}_+^*$.

Utilisant la méthode de substitution à $\bar{X}_n$ estimateur de $\frac{1}{\theta}$:

$$\hat{\eta}_n^{(1)} = h\left(\bar{X}_n\right) = \exp\left(-\frac{x_0}{\bar{X}_n}\right)$$

retour au sujet

❸ h est dérivable sur $\mathbb{R}_+^*$ avec $h'(u) = \dfrac{x_0}{u^2}\exp\left(-\dfrac{x_0}{u}\right)$.
Appliquant le Delta théorème dans le contexte gaussien :

$$\sqrt{n}\left(h\left(\bar{X}_n\right) - h\left(\frac{1}{\theta}\right)\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, h'\left(\frac{1}{\theta}\right)^2 \frac{1}{\theta^2}\right)$$

Soit :

$$\sqrt{n}\left(\hat{\eta}_n^{(1)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \left(x_0\theta\exp\left(-\theta x_0\right)\right)^2\right)$$

La variance asymptotique de $\hat{\eta}_n^{(1)}$ est $\sigma_1^2(\theta) = \left(x_0\theta\exp\left(-\theta x_0\right)\right)^2$.

Corrigé de l'exercice 2

retour au sujet

❹ $\hat{\eta}_n^{(2)} = \frac{1}{n}\sum_{i=1}^{n} Z_i$ avec $Z_i = \mathbb{1}_{X_i \geq x_0} \implies \begin{cases} Z_1, \dots, Z_n \text{ IID} \\ Z_1 \sim \mathrm{Ber}(\eta) \end{cases}$

Utilisant le résultat de l'exercice 1 :

Exercice 1

$$\sqrt{n}\left(\hat{\eta}_n^{(2)} - \eta\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \eta(1-\eta)\right)$$

avec $\eta = \exp\left(-\theta x_0\right)$, on obtient la variance asymptotique :

$$\sigma_2^2(\theta) = \exp\left(-\theta x_0\right)\left(1 - \exp\left(-\theta x_0\right)\right).$$

Corrigé de l'exercice 2

retour au sujet

Soit $\Delta(\theta) = \sigma_2^2(\theta) - \sigma_1^2(\theta)$.

$$\begin{aligned}\Delta(\theta) &= \exp\left(-\theta x_0\right)\left(1 - \exp\left(-\theta x_0\right) - x_0^2\theta^2 \exp\left(-\theta x_0\right)\right)\\ &= \exp\left(-\theta x_0\right)\varphi(\theta x_0)\end{aligned}$$

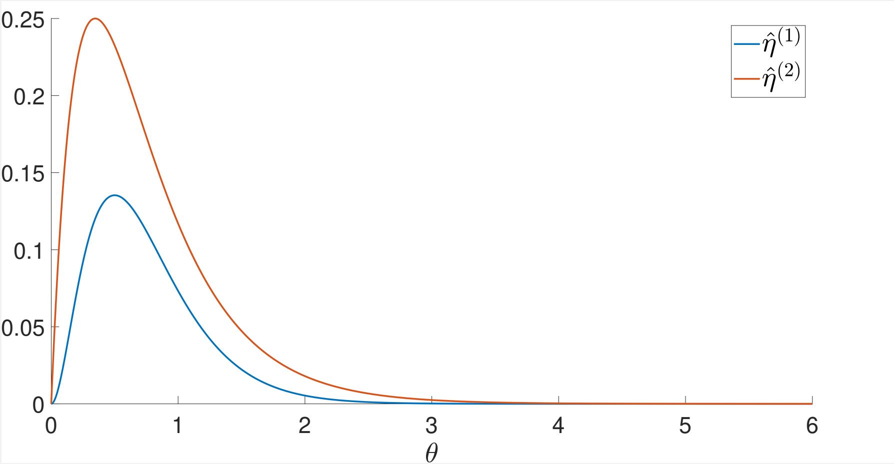
avec $\varphi(u) = 1 - \exp(-u)(1+u^2)$.

Un tableau de variation de φ montre que $\varphi > 0$ sur $\mathbb{R}_+$.

$\hat{\eta}_n^{(1)}$ est donc asymptotiquement préférable à $\hat{\eta}_n^{(2)}$.

Corrigé de l'exercice 2

retour au sujet



Tracés des 2 variances asymptotiques pour $x_0 = 2.0$.

Corrigé de l'exercice 3

retour au sujet

❶ Par application directe du TCL :

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \sigma\sqrt{\frac{\pi}{2}}\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\left(0, \sigma^2\left(2-\frac{\pi}{2}\right)\right).$$

$$\sqrt{n}\left(\frac{\frac{\bar{X}_n}{\sigma} - \sqrt{\frac{\pi}{2}}}{\sqrt{2-\frac{\pi}{2}}}\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1).$$

❷ Ainsi :

$T_n = \sqrt{n}\left(\frac{\frac{\bar{X}_n}{\sigma} - \sqrt{\frac{\pi}{2}}}{\sqrt{2-\frac{\pi}{2}}}\right)$ est une fonction pivotale asymptotique.

Corrigé de l'exercice 3

retour au sujet

❸ Comme $T_n \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$, il vient avec un niveau $1-\alpha$:

$$-q_{1-\frac{\alpha}{2}} \leq \sqrt{n}\left(\frac{\frac{\bar{X}_n}{\sigma} - \sqrt{\frac{\pi}{2}}}{\sqrt{2-\frac{\pi}{2}}}\right) \leq q_{1-\frac{\alpha}{2}}$$

où q_r est le quantile d'ordre r de la loi $\mathcal{N}(0,1)$.

D'où l'intervalle de confiance asymptotique de niveau $1-\alpha$:

$$\sqrt{\frac{2}{\pi}}\bar{X}_n \frac{1}{1+\frac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{\frac{4}{\pi}-1}} \leq \sigma \leq \sqrt{\frac{2}{\pi}}\bar{X}_n \frac{1}{1-\frac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{\frac{4}{\pi}-1}}$$

retour au sujet

On peut simplifier l'IC asymptotique en effectuant un développement limité :

IC asymptotique de niveau $1-\alpha$ pour σ

$$C_n = \sqrt{\frac{2}{\pi}}\bar{X}_n \left[1 - \frac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{\frac{4}{\pi}-1},\, 1 + \frac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{\frac{4}{\pi}-1}\right]$$

retour au sujet

❹ $X_i \sim \mathcal{R}(\sigma^2)$. σ étant un paramètre d'échelle : $\dfrac{X_i}{\sigma} \sim \mathcal{R}(1)$.

Les X_i étant IID, il vient $\dfrac{1}{\sigma}\displaystyle\sum_{i=1}^{n} X_i \sim \mathcal{SR}(n, 1)$.

❺ Taux de couverture de $C_n\left(\underline{X}_n\right)$

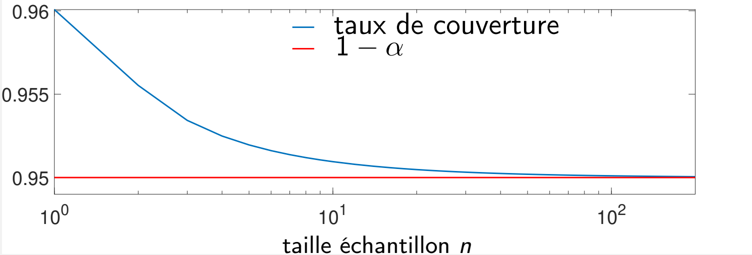
$$\begin{aligned}
\mathbb{P}_\theta\left(\sigma \in C_n\left(\underline{X}_n\right)\right) &= \mathbb{P}_\theta\left(a_n \leq \frac{1}{\sigma}\sum_{i=1}^{n} X_i \leq b_n\right) \\
&= F^{(n)}(b_n) - F^{(n)}(a_n)
\end{aligned}$$

$$\text{avec} \left\{\begin{aligned}
a_n &= n\sqrt{\frac{\pi}{2}} - \sqrt{n}\sqrt{2 - \frac{\pi}{2}}\, q_{1-\frac{\alpha}{2}} \\
b_n &= n\sqrt{\frac{\pi}{2}} + \sqrt{n}\sqrt{2 - \frac{\pi}{2}}\, q_{1-\frac{\alpha}{2}}
\end{aligned}\right.$$

Remarque. Ici le taux de couverture ne dépend pas de θ.
C'est un cas particulier car σ est un paramètre d'échelle.

Corrigé de l'exercice 3

retour au sujet



Taux de couverture de l'IC asymptotique $C_n(\underline{X}_n)$ avec $\alpha = 5\%$.

Remarque. On observe qu'on a bien un intervalle de confiance de niveau asymptotique (exactement) $1-\alpha$:

$$\forall \theta, \quad \lim_{n\to\infty} \mathbb{P}_\theta\left(\sigma \in C_n\left(\underline{X}_n\right)\right) = 1-\alpha.$$

❶ On a montré dans le cours (TCL, Slutski, Mann-Wald) que

$$T_n\left(\underline{X}_n, \eta\right) = \sqrt{n}\,\frac{\left(\bar{X}_n - \eta\right)}{\bar{X}_n} \xrightarrow[n\to\infty]{\text{loi}} \sim \mathcal{N}\left(0,1\right),$$

➠ T_n est donc une fonction pivotale asymptotique.

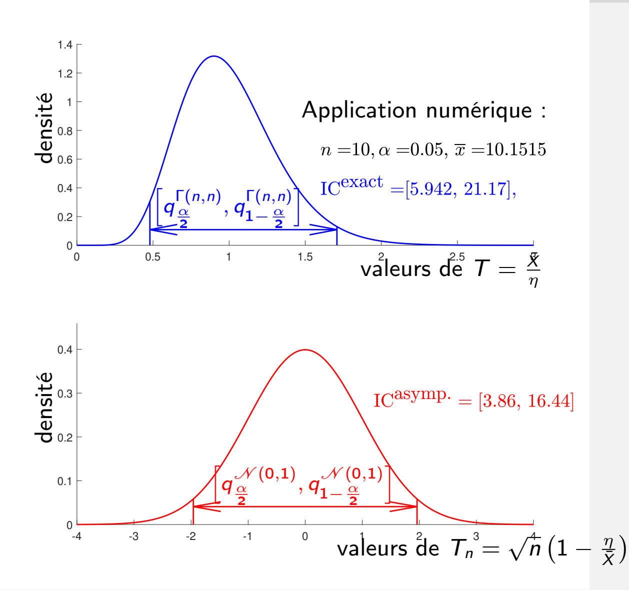
❷ IC asymptotique de niveau $1-\alpha$ pour η :

$$C_n = \left[\left(1 - \frac{1}{\sqrt{n}}\,q_{1-\frac{\alpha}{2}}\right)\bar{X}_n,\ \left(1 + \frac{1}{\sqrt{n}}\,q_{1-\frac{\alpha}{2}}\right)\bar{X}_n\right]$$

où q_r est le quantile d'ordre r de la loi $\mathcal{N}(0,1)$.

Corrigé de l'exercice 4

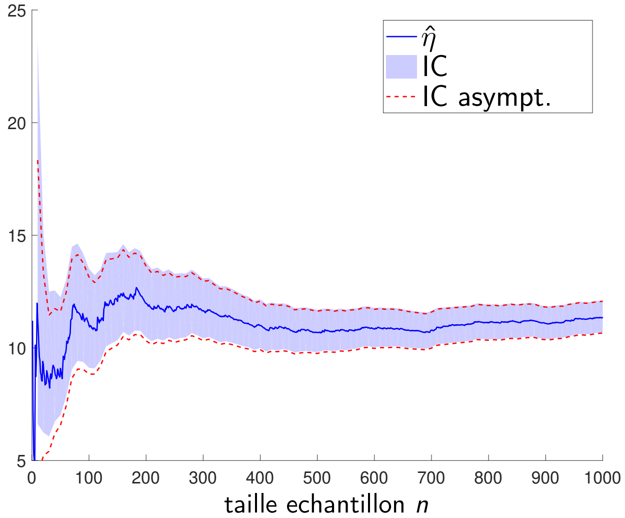
retour au sujet



⚠ Ne pas confondre les intervalles sur les fonctions pivotales $\left[q_{\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}\right]$ et les intervalles de confiance sur η.

Corrigé de l'exercice 4

retour au sujet



Comparaison des IC exact et asymptotique en fonction de n

retour au sujet

❸ Taux de couverture de $C_n\left(\underline{X}_n\right)$

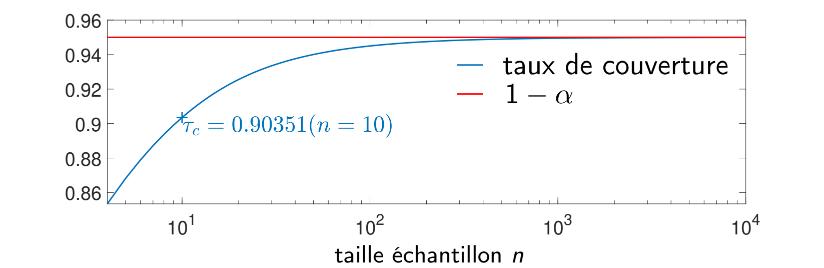
$$\begin{aligned}\mathbb{P}_\theta\left(\eta \in C_n\left(\underline{X}_n\right)\right) &= \mathbb{P}_\theta\left(\eta \in \left[\left(1-\frac{1}{\sqrt{n}}\,q_{1-\frac{\alpha}{2}}\right),\left(1+\frac{1}{\sqrt{n}}\,q_{1-\frac{\alpha}{2}}\right)\right]\bar{X}_n\right)\\ &= \mathbb{P}_\theta\left(\frac{1}{1+\frac{1}{\sqrt{n}}\,q_{1-\frac{\alpha}{2}}} \leq \frac{\bar{X}_n}{\eta} \leq \frac{1}{1-\frac{1}{\sqrt{n}}\,q_{1-\frac{\alpha}{2}}}\right)\end{aligned}$$

Comme (rappel) $\dfrac{\bar{X}_n}{\eta} \sim \Gamma\left(n,n\right)$, il vient :

$$\mathbb{P}_\theta\left(\eta \in C_n\left(\underline{X}_n\right)\right) = F^{\Gamma(n,n)}\left(\frac{1}{1-\frac{1}{\sqrt{n}}\,q_{1-\frac{\alpha}{2}}}\right) - F^{\Gamma(n,n)}\left(\frac{1}{1+\frac{1}{\sqrt{n}}\,q_{1-\frac{\alpha}{2}}}\right)$$

avec $F^{\Gamma(n,n)}$ fonction de répartition de la loi $\Gamma(n,n)$.

Corrigé de l'exercice 4



Taux de couverture de l'IC asympt. de niveau 95%

Remarques.

- On retrouve que $\forall \theta, \lim_{n\to\infty} \mathbb{P}_\theta\left(\eta \in C_n\left(\underline{X}_n\right)\right) \geq 1-\alpha$.
- En général le taux de couverture dépend de θ. Ici, ce n'est pas le cas car η est un paramètre d'échelle.

Plan du cours

Démonstration

Z admet un moment d'ordre 2. On peut donc définir :

- sa moyenne $\mu = \mathbb{E}(Z)$,
- sa matrice de covariance $\Sigma_Z = \mathbb{E}((Z-\mu)(Z-\mu)^\top)$.

Commençons par remarquer que s'il existe $a \in \mathbb{R}^q \setminus \{0\}$ et $c \in \mathbb{R}$ tel que $a^\top Z = c;$ p.s., alors $c = a^\top \mu$.

Un résultat intermédiaire

Soit V variable aléatoire scalaire **positive**. On a :

$$\mathbb{E}(V) = 0 \iff V = 0 \text{ p.s.} \qquad (*)$$

Démonstration (suite)

Soient $a \in \mathbb{R}^q \setminus \{0\}$ et $c \in \mathbb{R}$.

$$\begin{aligned} a^\top Z = c \text{ p.s.} &\iff a^\top (Z-\mu) = 0 \text{ p.s.} \\ &\iff a^\top (Z-\mu)(Z-\mu)^\top a = 0 \text{ p.s.} \\ &\iff \mathbb{E}(a^\top (Z-\mu)(Z-\mu)^\top a) = 0 \text{ (utilisant } (*)) \\ &\iff a^\top \Sigma_Z a = 0 \end{aligned}$$

La matrice Σ_Z étant définie positive, $a^\top \Sigma_Z a = 0$ (avec $a \neq 0$) signifie que $a \in \mathrm{Ker}(\Sigma_Z)$.
Ainsi,

$$\begin{aligned} Z \text{ dégénérée} &\iff \exists a \neq 0 \text{ t.q. } a^\top Z = c \text{ p.s.} \\ &\iff \exists a \neq 0 \in \mathrm{Ker}(\Sigma_Z) \\ &\iff \Sigma_Z \text{ non inversible} \end{aligned}$$

$\square$

retour au slide 7

Lien entre convergence en loi et en probabilité

On sait déja que la convergence en probabilité entraine la convergence en loi. Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de VA à valeurs dans $\mathbb{R}^d$.

Proposition

Si $Y_n \xrightarrow{\text{loi}} c$, avec $c \in \mathbb{R}^d$ une constante, alors $Y_n \xrightarrow{\mathbb{P}} c$.

Corollaire

S'il existe $c \in \mathbb{R}^d$,

- une VA Z à valeurs dans $\mathbb{R}^d$,
- une suite $(a_n)_{n\in\mathbb{N}^*}$ à valeurs réelles et telle que $\lim\limits_{n\to\infty} a_n = \infty$,

tels que

$$a_n (Y_n - c) \xrightarrow[n\to\infty]{\text{loi}} Z$$

alors

$$Y_n \xrightarrow[n\to\infty]{\mathbb{P}} c.$$

Démo. (exercice) : combiner la prop. ci-dessus et le thm de Slutsky (voir plus loin). □

retour au slide 8

Démonstration « delta théorème » (cas scalaire)

Soit la fonction ψ définie par :

$$\psi(y) = \begin{cases} \dfrac{h(y) - h(m)}{y - m} & \text{si } y \neq m, \\ h'(m) & \text{si } y = m; \end{cases}$$

ψ est continue en m car h est dérivable en m. Comme $Y_n \xrightarrow[n\to\infty]{\text{loi}} m$,

$$\psi(Y_n) \xrightarrow[n\to\infty]{\text{loi}} \psi(m) = h'(m),$$

et donc (Slutsky)

$$\left(\sqrt{n}(Y_n - m), \psi(Y_n)\right) \xrightarrow[n\to\infty]{\text{loi}} \left(Z, h'(m)\right).$$

Finalement, on a

$$\sqrt{n}\,(h(Y_n) - h(m)) = \sqrt{n}\,(Y_n - m)\,\psi(Y_n) \xrightarrow[n\to\infty]{\text{loi}} h'(m)\,Z. \quad \square$$

retour au slide 13

Modèles réguliers : conditions de régularité C_3 et C_4

Rappel : C_0, C_1 et C_2 ont été définies au cours précédent.

Condition de régularité C_3

$\theta \mapsto f_\theta(\underline{x})$ est deux fois continûment différentiable ν-presque pour tout $\underline{x}$.

Condition de régularité C_4

En tout point $\theta \in \Theta$, on a

$$\int_{\mathcal{S}} \nabla_\theta \nabla_\theta^\top f_\theta(\underline{x})\, \nu(\mathrm{d}\underline{x}) = \nabla_\theta \int_{\mathcal{S}} \nabla_\theta^\top f_\theta(\underline{x})\, \nu(\mathrm{d}\underline{x}).$$

Autrement dit : $\forall \theta \in \Theta$, $\forall k \leq p$, $\forall j \leq p$,

$$\int_{\mathcal{S}} \frac{\partial^2 f_\theta(\underline{x})}{\partial \theta_k \partial \theta_j}\, \nu(\mathrm{d}\underline{x}) = \frac{\partial}{\partial \theta_k} \int_{\mathcal{S}} \frac{\partial f_\theta(\underline{x})}{\partial \theta_j}\, \nu(\mathrm{d}\underline{x}).$$

retour au slide 20

Exemple d'EMV non asymptotiquement gaussien

Soient $X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathcal{U}_{[0,\theta]}$, avec $\theta > 0$ inconnu.

⚠ Ce modèle n'est pas régulier (pourquoi ?).

On montre que (cf. TD1, exercice 1.2)

- $\hat{\theta}_n = \max_{i \leq n} X_i$ est l'EMV de θ, et
- $n\left(\hat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{\text{loi}} -Z$ avec $Z \sim \mathcal{E}\left(\lambda = \dfrac{1}{\theta}\right)$.

Dans ce cas particulier,

➠ l'EMV n'est pas asymptotiquement gaussien ;

➠ la vitesse de convergence est $\frac{1}{n}$: plus rapide que $\frac{1}{\sqrt{n}}$.

➠ retour au slide 20

La loi Rayleigh $\mathscr{R}\left(\sigma^2\right)$

Soit $X \sim \mathscr{R}\left(\sigma^2\right)$ avec $\sigma > 0$. Sa densité est :

$$f(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \mathbb{1}_{\mathbb{R}^+}(x).$$

Moments

- moyenne : $\mathbb{E}_\sigma(X) = \sigma\sqrt{\frac{\pi}{2}}$
- variance : $\mathrm{var}_\sigma(X) = \sigma^2\left(2 - \frac{\pi}{2}\right)$

Propriété

si $X \sim \mathscr{R}(\sigma^2)$ alors $Y = X^2 \sim \mathcal{E}\left(\frac{1}{2\sigma^2}\right)$.

Loi d'une somme de lois de Rayleigh

On définit (pour l'exercice) la loi suivante :

$$\text{Si } (X_1, \ldots, X_n) \overset{\text{iid}}{\sim} \mathcal{R}(\sigma^2), \text{ alors } Z = \sum^n X_i \sim \mathcal{SR}(n, \sigma^2).$$