Chap 2 : Vraisemblance, Estimation ponctuelle et Intervalle de Confiance

Sommaire

1. Vraisemblance et Estimation
   - Vraisemblance, exhaustivité, liberté, complétude
   - Information de Fisher
   - L'estimateur du maximum de vraisemblance
   - Comportement asymptotique des estimateurs
2. Intervalle de Confiance
   - Construction pratique d'un Intervalle de Confiance
   - Intervalle de Confiance Asymptotique

Vraisemblance et Estimation Ponctuelle

Vraisemblance, exhaustivité, liberté, complétude

Vraisemblance d'un échantillon

- Soit $s_n = (x_1, \ldots, x_n)$ réalisation d'un échantillon (i.i.d) de loi de densité $f(x;\theta)$, dans un modèle statistique $\mathscr{M}$, la vraisemblance de l'échantillon s_n est la fonction $L(\theta, x_1, \ldots, x_n)$:

$$\begin{array}{rcl} L : \Theta & \longrightarrow & \mathbb{R}^+ \\ \theta & \mapsto & f(x_1, \ldots, x_n; \theta) = \prod_{i=1}^n f(x_i;\theta) \end{array}$$

- La log-vraisemblance est

$$\begin{array}{rcl} \mathscr{L}(\theta, x_1, \ldots, x_n) & = & \log L(\theta, x_1, \ldots, x_n) \\ & = & \sum_{i=1}^n \log(f(x_i;\theta)) \end{array}$$

- Exemple :

Exhaustivité

Définition

Soit $\mathscr{M} = \{f(x,\theta), x \in \mathscr{X}, \theta \in \Theta\}$un modèle paramétrique dominé et $S : \mathscr{X} \longrightarrow \mathscr{Y}$ une statistique. La statistique S est dite **exhaustive** si la loi de X sachant $S = s$ est indépendante de θ.

- Caractérisation via **le théorème de factorisation** : Une statistique $S(x_1,\ldots,x_n)$ est exhaustive du modèle paramétrique (dominé par μ) ssi la densité s'écrit

$$f(x_1,\ldots,x_n;\theta) = g_\theta\left(S(x_1,\ldots,x_n);\theta\right) h(x)$$

- Interprétation : $S(x_1,\ldots,x_n)$ contient toute l'information de l'échantillon sur θ, c'est un **résumé** qui préserve l'information.
- Propriété : Si S_1, S_2 sont des statistiques telles que $S_1 = h(S_2)$. Alors si S_1 est exhaustive pour θ, alors S_2 est exhaustive pour θ.
- **Définition** : Une statistique S est **exhaustive minimale** si tout autre statistique exhaustive T s'écrit $S = h(T)$.

Exemples

- L'échantillon lui-même
- Le cas multinomial
- Le cas gaussien
- Le cas uniforme

Liberté, Complétude

- Dans un modèle paramétrique $\mathscr{M}$, une statistique S est **libre** en θ si sa loi ne dépend pas de θ.
- Une statistique S est complète ssi

$$\forall \theta \in \Theta,\ E_{\theta}\left[h(S)\right] = 0 \implies h = 0,\, P_{\theta} - ps$$

- Théorème : Une statistique exhaustive complète est minimale.
- Théorème de Basu : Si S est une statistique exhaustive complète pour le modèle $\mathscr{M}$, alors toute statistique libre T est indépendante de S pour toute loi P_{θ} de $\mathscr{M}$.
  - Interprétation : Les statistiques exhaustives complètes sont des résumés qui compressent et sépare l'échantillon de manière optimale, en deux parties **indépendantes.**

Estimation et exhaustivité

- Théorème de **Rao-Blackwell** : Soit T une statistique exhaustive, et $\hat{\theta}$ un estimateur de θ tel que $\mathscr{R}(\hat{\theta},\theta) = E_\theta\left[\left(\hat{\theta}-\theta\right)^2\right] < \infty$, alors l'estimateur $\hat{\eta} = E_\theta\left[\hat{\theta}|T\right]$ est l'amélioré de Rao-Blackwell :

$$\mathscr{R}(\hat{\eta},\theta) < \mathscr{R}(\hat{\theta},\theta)$$

- Théorème de **Lehmann-Scheffé** : Soit S une statistique exhaustive complète pour un modèle paramétrique. Si $\hat{\theta}$ est un estimateur sans biais de θ, alors $\hat{\eta} = E_\theta\left[\hat{\theta}|T\right]$ est un estimateur sans biais uniformément de variance minimale.
- Conclusion et interprétation : de bons résumés optimaux de l'échantillon permettent d'augmenter la précision. .

Le modèle exponentiel

Définition

Un modèle statistique $\{f(x;\theta)|\theta \in \Theta\}$ sur $\mathscr{X} = \mathbb{R}^d$ est une **famille exponentielle** si les densités $f(x;\theta)$ s'écrivent

$$f(x;\theta) = \exp\left(\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right) h(x)$$

avec h, B, η_i, T_i, $i = 1, \ldots, s$ à valeurs réels, et $h(x) \geq 0$.

- $T = (T_1 \ldots T_s)$ est la statistique privilégiée du modèle (les T_i sont linéairement libres = de rang plein)
- Il est plus commode d'utiliser directement les η_i comme paramètre : la densité s'écrit alors sous **forme canonique**

$$f(x;\eta) = \exp\left(\sum_{i=1}^{s} \eta_i T_i(x) - A(\eta)\right) h(x)$$

et $\eta = (\eta_1, \ldots, \eta_s)$ est appelé **paramètre naturel**.

Exemple : La loi Gamma

Propriétés

- $C(\eta) = \exp(-A(\eta))$ est la constante de normalisation.
- L'ensemble $H = \{\eta \in \mathbb{R}^s \,|\, \int_{\mathscr{X}} \exp\left(\sum_{i=1}^s \eta_i T_i(x)\right) < \infty\}$ est un convexe.
- La fonction $\eta \mapsto A(\eta)$ est indéfiniment dérivable et
  1. $E_\eta[T(X)] = \nabla_\eta A(\eta)$
  2. $\mathrm{cov}_\eta(T_i, T_j) = \frac{\partial^2}{\partial \eta_i \partial \eta_j} A(\eta)$
- $\theta \mapsto \eta(\theta)$ est le plus souvent un difféomorphisme.
- Pour une famille exponentielle, le modèle d'échantillonnage est encore une famille exponentielle avec la statistique privilégiée :

$$T(x_1, \ldots, x_n) = \sum_{i=1}^n T(x_i).$$

Exhaustivité dans les familles exponentielles

- Dans une famille exponentielle, si le modèle est de rang plein (égale à s), la **statistique privilégiée est exhaustive minimale**.

Réciproquement,

- **Théorème de Pitman-Koopman** : Si le modèle est homogène (le support ne dépend pas du paramètre) et est tel que pour une taille d'échantillon n assez grand, il existe une statistique exhaustive de taille fixe s, alors la famille est exponentielle.

Information de Fisher

Score et Information de Fisher

- On suppose que $\mathscr{M} = \{f(x;\theta), \theta \in \Theta\}$ est **régulier**. Le **score** est le vecteur aléatoire

$$S(X;\theta) = \nabla_\theta \log f(X;\theta)$$

Dans ce cas, $\forall \theta \in \Theta$,

$$E_\theta \left[\nabla_\theta \log f(X;\theta)\right] = 0$$

Attention au support !

- **La matrice d'information de Fisher** $I(\theta)$, ou information apportée par X sur θ :

$$\mathscr{I}(X;\theta) = \mathrm{cov}_\theta \left(\nabla_\theta \log f(X;\theta)\right)$$

- **Additivité de l'Information de Fisher** pour un échantillon i.i.d :

$$\mathscr{I}(X_1, \ldots, X_n;\theta) \triangleq \mathscr{I}_n(\theta) = n\mathscr{I}_1(\theta)$$

Calcul de l'Information de Fisher

- Théorème : Si $\frac{\partial^2}{\partial\theta_i\partial\theta_j} f(x;\theta)$ existe pour tout x et tout θ, et si on peut dériver 2 fois sous le signe intégral $\int f(x;\theta)\mu(dx)$ alors

$$\mathscr{I}(X;\theta) = -\left(E_\theta\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log f(X;\theta)\right]\right)_{1\leq i,j\leq p}$$

- Pour une famille exponentielle, si on utilise la paramétrisation canonique :

$$\left\{\begin{array}{lcl} S(X;\eta) & = & T(X) - \nabla_\eta A(\eta) \\ \mathscr{I}(X;\eta) & = & \nabla^2_\eta A(\eta) = \mathrm{cov}_\eta(T(X)) \end{array}\right.$$

Inégalité d'information de Cramér-Rao

Théorème

Soit $\mathscr{M}$ *modèle statistique* ***régulier***, g *fonction dérivable et* $T_n = T(X_1, \ldots, X_n)$ *un estimateur sans biais de* $g(\theta)$.
Cas θ *scalaire : la borne de Cramer-Rao est définie par* $BCR(\theta) = \frac{g'(\theta)}{\mathscr{I}_n(\theta)}$ *et*

$$V_\theta\,(T_n) \geq BCR(\theta) = \frac{g'(\theta)^2}{\mathscr{I}_n(\theta)}$$

Cas θ *vectoriel :*

$$V_\theta\,(T_n) \geq J(g)(\theta)^\top \mathscr{I}_n(\theta)^{-1} J(g)(\theta)$$

avec $J(g)(\theta)$= *Jacobien de* g *évalué au point* θ.

Efficacité d'un estimateur

- La comparaison entre les matrices de covariances est au sens de l'ordre (partiel) matrices s.d.p : A, B s.d.p. $A \leq B \iff B - A$ s.d.p
- Inégalité d'information = $BCR(\theta)$ est la meilleure variance possible pour un estimateur sans biais.
- Un estimateur T_n sans biais de $g(\theta)$ est un estimateur **efficace** si $V_\theta(g(\theta)) = BCR(\theta)$:

Estimateur du maximum de vraisemblance

Estimateur du Maximum de Vraisemblance

- Soit $\mathscr{M}$ un modèle paramétrique et $x_1, \ldots, x_n$ un échantillon de vraisemblance $L(\theta; x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i; \theta)$. Lorsqu'il existe, l'estimateur du maximum de vraisemblance est définie par

$$\begin{aligned} \hat{\theta}_n^{EMV} &= \arg\min_{\theta \in \Theta} L(\theta; x_1, \ldots, x_n) \\ &= \arg\min_{\theta \in \Theta} \mathscr{L}(\theta; x_1, \ldots, x_n) \end{aligned}$$

Remarque

Les conditions suffisantes d'optimalité pour déterminer $\hat{\theta}_n^{EMV}$ sont :

$$\begin{cases} \dfrac{\partial \mathscr{L}}{\partial \theta}(x, \theta)|_{\theta = \hat{\theta}_n^{EMV}} = 0 \\ \dfrac{\partial^2 \mathscr{L}}{\partial \theta^2}(x, \hat{\theta}_n^{EMV}) \leq 0 \\ \mathscr{L}(x, \hat{\theta}_n^{EMV}) \geq \mathscr{L}(x, \theta), \forall \theta \in \Theta. \end{cases}$$

Exemples d'estimateurs d'EMV

- Le cas Bernoulli
- Le cas Poisson

Propriétés de l'EMV

- Sous les hypothèses du théorème de factorisation, $\hat{\theta}_n^{EMV}$ est fonction de toute statistique exhaustive.
- Pour l'EMV, il y a des problèmes :
  - Existence d'une solution au problème de maximisation
  - Existence de plusieurs solutions et de minima locaux
  - Numérique (pour l'optimisation)
- Pour un modèle exponentiel de plein rang (avec la paramétrisation naturelle) et $\mathscr{I}(\eta) > 0$ pour tout η, $\mathscr{L}(\theta; x_1, \ldots, x_n)$ est strictement concave et l'EMV est l'unique solution de

$$\frac{1}{n}\sum_{j=1}^{n} T(x_j) = \nabla_\eta A(\hat{\eta}^{EMV})$$

Comportement asymptotique des estimateurs

Convergence de l'EMV

- Si le modèle paramétrique $\mathscr{M}$ est régulier et si $\theta^* \in \overset{\circ}{\Theta}$ est le vrai paramètre alors
$$\hat{\theta}_n^{EMV} \longrightarrow \theta^*$$
en probabilité P_{θ^*}.

Remarque

La régularité suppose que l'information de Fisher $\mathscr{I}(\theta)$ existe et est définie positive. Une conséquence de $\mathscr{I}(\theta) > 0$ est que le modèle est **identifiable**, i.e $f(\cdot;\theta) = f(\cdot,\theta') \implies \theta = \theta'$, for all θ, θ'.

Normalité asymptotique de l'EMV

- Si le modèle paramétrique $\mathscr{M}$ est régulier et si $\theta^* \in \overset{\circ}{\Theta}$ est le vrai paramètre alors

$$\sqrt{n}\left(\hat{\theta}_n^{EMV} - \theta^*\right) \rightsquigarrow N(0, \mathscr{I}_1(\theta))$$

Remarque

- $\sqrt{n}$ est la vitesse de convergence de l'estimateur. De manière générale, la convergence des estimateurs paramétriques est "en $\frac{1}{\sqrt{n}}$".
- Un estimateur consistent $\hat{\theta}_n$ est dit **asymptotiquement normal** si il existe $r_n > 0$ tel que

$$r_n\left(\hat{\theta}_n - \theta^*\right) \rightsquigarrow N(0, \mathscr{J}(\theta)).$$

$\mathscr{J}(\theta)$ est la **variance asymptotique**.

Delta-Method

- Soit $\hat{\theta}_n$, un estimateur asymptotiquement normal $r_n\left(\hat{\theta}_n - \theta^*\right) \rightsquigarrow N(0, \mathscr{I}(\theta))$, et ψ une fonction différentiable autour de θ^*, alors $\psi(\hat{\theta}_n)$ est asymptotiquement normal et

$$r_n\left(\psi\left(\hat{\theta}_n\right) - \psi(\theta^*)\right) \rightsquigarrow N\left(0, \psi'(\theta)^\top \mathscr{I}(\theta)\psi'(\theta)\right)$$

- Lemme de Slutsky : Si $Y_n \rightsquigarrow Y$ et $A_n \stackrel{Proba}{\longrightarrow} A$ et $B_n \stackrel{Proba}{\longrightarrow} B$, alors

$$A_n + B_n Y_n \rightsquigarrow A + BY$$

- Ces résultats sont utilisées pour obtenir la normalité asymptotique de statistique complexe (mais régulière) de l'échantillon.

Intervalle de confiance

Intervalle de Confiance

On considËre un modËle d'Èchantillonnage paramÈtrique $X_1,...,X_n$ de taille n. Les variables alÈatoires $X_1,...,X_n$ sont indÈpendantes et de mÍme loi inconnue P, et elles sont à valeurs dans $\mathscr{X}$ ($\mathbb{R}$ ou $\mathbb{R}^d$).

On suppose que la loi P appartient à une famille de probabilitÈ $\mathscr{P} = \{P_\theta, \theta \in \Theta\}$.

DÈfinition 3.1 (RÈgion de Confiance)

Soit $x \mapsto \Lambda(x)$ une application de $\mathscr{X}^n$ à valeurs dans l'ensemble des borÈliens de $g(\Theta)$ telle que pour tout $\theta \in \Theta$, on a

$$P_\theta(g(\theta) \in \Lambda(X_1,...,X_n)) = 1 - \alpha.$$

On dit que l'ensemble alÈatoire Λ est une rÈgion de confiance de niveau $(1-\alpha)$.

Dans le cas où $g(\Theta) \subset \mathbb{R}$ et Λ est un intervalle, on parle d'un intervalle de confiance.

Intervalle de Confiance

DÈfinition 3.4 (Statistique asymptoptiquement libre)

On appelle statistique asymptotiquement libre pour θ, toute statistique dont sa loi asymptotique ne dÈpend pas de θ.

ple 3.1 Soit $X_1,...,X_n$ un Èchantillon de loi $\mathcal{N}(\mu,\sigma^2)$. Notons $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ et

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

- Si σ est connu, $\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma}$ est une statistique libre pour μ, de loi $\mathcal{N}(0,1)$.

- Si σ est inconnu $\frac{\sqrt{n}(\bar{X}_n-\mu)}{S_n}$ est une statistique libre pour μ, de loi T_{n-1}.

- Si μ est connu $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$ est une statistique libre pour σ, de loi χ^2_n.

- Si μ est inconnu $(n-1)\frac{S_n^2}{\sigma^2}$ est une statistique libre pour σ, de loi χ^2_{n-1}.

Construction pratique d'un IC

ple 3.2 $X_1, ..., X_n$ un Èchantillon de loi $P(\frac{1}{\lambda})$. $T_n = \frac{\sqrt{n}(\bar{X}_n - \lambda)}{\lambda}$ est une statistique asymptotiquement libre pour λ. En effet, par le thÈorËme central-limite $T_n \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$.

Construction pratique d'un IC

Soit T_n un estimateur du paramËtre θ. On cherche $\varphi_1(T_n)$ et $\varphi_2(T_n)$ de sorte que leurs lois ne dÈpendent pas de θ et

$$P(\varphi_1(T_n) \leq \theta \leq \varphi_2(T_n)) = 1 - \alpha$$

marque Le choix d'une statistique libre revient à Èviter d'avoir les bornes de l'IC dÈpendantes de θ

IC - EspÈrance d'une variable normale : σ connu

connu Soit $X_1,...,X_n$ un Èchantillon de loi $\mathcal{N}(\mu,\sigma^2)$.

- Prenons $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ comme estimateur de μ.

- On utilise le fait que $\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma}$ est une statistique libre pour μ, de loi $\mathcal{N}(0,1)$.

- L'intervalle de probabilitÈ de $\bar{X}_n$ à $1-\alpha$ est :

$$\begin{array}{rcl} -u_{\frac{\alpha}{2}} & \leq \frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma} \leq & u_{\frac{\alpha}{2}} \\ \mu - u_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}} & \leq \bar{X}_n \leq & \mu + u_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}} \end{array}$$

où $u_{\frac{\alpha}{2}}$ est telle que $\phi(u_{\frac{\alpha}{2}}) = P(N(0,1) \leq u_{\frac{\alpha}{2}}) = 1-\frac{\alpha}{2}$.
d'où l'intervalle de confiance :

$$\bar{x}_n - u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \quad \leq \mu \leq \quad \bar{x}_n + u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

- Application numÈrique : si $(1-\alpha) = 0,95$, on a $u_{\frac{\alpha}{2}} = 1,96$. L'intervalle de confiance de niveau 95% pour μ est

IC - EspÈrance d'une variable normale : σ inconnu

On utilise le fait que $T_{n-1} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n}$ suit une loi de Student à $(n-1)$ degrÈs de libertÈ.

L'intervalle de probabilitÈ pour T_{n-1} :

$$-t_{(n-1)\,,\,\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \leq t_{(n-1)\,,\,\frac{\alpha}{2}}$$

d'où l'intervalle de confiance :

$$\bar{x}_n - t_{(n-1)\,,\,\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x}_n + t_{(n-1)\,,\,\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

marque Si la dÈfinition $S_n^{2*} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ alors l'intervalle de confiance devient :

$$\bar{x}_n - t_{(n-1)\,,\,\frac{\alpha}{2}} \frac{s^*}{\sqrt{n-1}} \leq \mu \leq \bar{x}_n + t_{(n-1)\,,\,\frac{\alpha}{2}} \frac{s^*}{\sqrt{n-1}}$$

ple 3.3 On prend, par exemple $\alpha = 0{,}05$, et $n = 25$, on a $t_{(n-1)\,,\,\frac{\alpha}{2}} = 2{,}064$ et l'intervalle de confiance de niveau de confiance 95% pour μ est :

$$\left[\bar{x}_n - \frac{2{,}064}{\sqrt{25}} s_{25}\ ,\ \bar{x}_n + \frac{2{,}064}{\sqrt{25}} s_{25}\right]$$

IC - Variance d'une variable normale

connu soit $\hat{U}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$. on utilise le fait que $\frac{n\hat{U}_n^2}{\sigma^2}$ suit une loi de χ_n^2 comme somme de n carrÈs de $\mathcal{N}(0,1)$ indÈpendantes.

Soit k_1 et k_2, les bornes de l'intervalle de probabilitÈ d'un χ_n^2 :

$$P(k_1 \leq \frac{n\hat{U}_n^2}{\sigma^2} \leq k_2) = 1-\alpha$$

d'où l'intervalle de confiance :

$$\frac{n\hat{u}_n^2}{k_2} \leq \sigma^2 \leq \frac{n\hat{u}_n^2}{k_1}$$

nconnu On utilise $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2$, et on sait que $\frac{(n-1)S_n^2}{\sigma^2}$ suit une loi χ_{n-1}^2,

soit b_1 et b_2 les bornes de l'intervalle de probabilitÈ :

$$P(b_1 \leq \frac{(n-1)S_n^2}{\sigma^2} \leq b_2) = 1-\alpha$$

d'où l'intervalle de confiance :

$$\frac{(n-1)s_n^2}{b_2} \leq \sigma^2 \leq \frac{(n-1)\hat{s}_n^2}{b_1}$$

Intervalle de Confiance : Exemple

ple 3.4 Soit $X_1, ..., X_n$ un Èchantillon de loi uniforme $\mathscr{U}([0,\theta])$.

La vraisemblance associÈ à l'Èchantillon est dÈfinie par

$$L_n(x,\theta) = \Pi_{i=1}^{n} \frac{1}{\theta} \mathbf{1}_{[0,\theta]}(x_i)$$

DÈterminons l'EMV : $L_n(x,\theta)$ est maximum en θ ssi $\Pi_{i=1}^{n} \mathbf{1}_{[0,\theta[}(x_i) = 1$.

Autrement dit $\forall i = 1, ..., n \;\; \theta \geq x_i$, donc, $\sup\limits_{1 \leq i \leq n} x_i \leq \theta$. D'où l'EMV :

$$\hat{\theta}_n = \sup_{1 \leq i \leq n} X_i$$

Cherchons une statistique libre pour θ (à partir de $\hat{\theta}_n$). D'abord calculons la loi de $\hat{\theta}_n$, $\forall x \in [0,\theta]$,

$$\begin{aligned} F_{\hat{\theta}_n}(x) &= P(\hat{\theta}_n \leq x) \\ &= (\frac{x}{\theta})^n \end{aligned}$$

VÈrifions que $T_n = \frac{\hat{\theta}_n}{\theta}$ est une statistique libre pour θ. En effet, $\forall x \in [0,1]$, on a

Intervalle de Confiance : Exemple

IC MÈthode 1 : DÈterminons un $\varphi_1(\hat{\theta}_n)$ et $\varphi_2(\hat{\theta}_n)$ telle que

$$P\left(\varphi_1(\hat{\theta}_n) \leq \theta \leq \varphi_2(\hat{\theta}_n)\right) = 1-\alpha \tag{1}$$

A partir de (1), il faut construire autour de θ, l'estimateur T_n

$$\begin{aligned} P\left(\frac{1}{\varphi_2(\hat{\theta}_n)} \leq \frac{1}{\theta} \leq \frac{1}{\varphi_1(\hat{\theta}_n)}\right) &= 1-\alpha \\ P\left(\frac{\hat{\theta}_n}{\varphi_2(\hat{\theta}_n)} \leq T_n \leq \frac{\hat{\theta}_n}{\varphi_1(\hat{\theta}_n)}\right) &= 1-\alpha \end{aligned}$$

Comme on connait la loi de la statistique libre T_n, on cherche donc un $a := \frac{\hat{\theta}_n}{\varphi_2(\hat{\theta}_n)}$ et $b := \frac{\hat{\theta}_n}{\varphi_1(\hat{\theta}_n)}$ tels que

$$\begin{aligned} P(a \leq T_n \leq b) &= 1-\alpha \\ F_{T_n}(b) - F_{T_n}(a) &= 1-\alpha \\ b^n - a^n &= 1-\alpha \end{aligned}$$

Il y a une infinitÈ de a et b possibles. On choisit a et b tels que l'intervalle $[a, b]$ soit le plus court possible. Soit $b = 1$ et donc $a = \alpha^{\frac{1}{n}}$.

Intervalle de Confiance : Exemple

Soit $\varphi_1(\hat{\theta}_n) = \hat{\theta}_n$ et $\varphi_1(\hat{\theta}_n) = \alpha^{-\frac{1}{n}}\hat{\theta}_n$ Nous avons donc obtenu un intervalle de confiance :

$$\left[\hat{\theta}_n \,;\, \alpha^{-\frac{1}{n}}\hat{\theta}_n\right]$$

IC MÈthode 2 : On peut partir directement de T_n, et l'on cherche un a et b tel que

$$\begin{aligned} P(a \leq T_n \leq b) &= 1-\alpha \\ F_{T_n}(b) - F_{T_n}(a) &= 1-\alpha \\ b^n - a^n &= 1-\alpha \end{aligned}$$

On choisit les a et b tels que $[a, b]$ soit l'intervalle le plus court possible.

Puis on remonte le raisonnement pour arriver à un intervalle de confiance (voir exemple sur les espÈrances/variances normales).

Intervalle de Confiance Asymptotique

Èthode ModËle rÈgulier, c'est-à-dire l'existence de l'information de Fisher.

- On construit l'intervalle asymptotique en utilisant l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de θ.
- On sait que $\hat{\theta}_n$ est asymtotiquement normal, i.e. :

$$\sqrt{n}(\hat{\theta}_n - \theta) \underset{n\to\infty}{\xrightarrow{\text{loi}}} \mathcal{N}(0, \frac{1}{I(\theta)}).$$

- Autrement dit :

$$\sqrt{n\, I(\theta)}(\hat{\theta}_n - \theta) \underset{n\to\infty}{\xrightarrow{\text{loi}}} \mathcal{N}(0,1).$$

$$P\left(-u_{\frac{\alpha}{2}} \leq \sqrt{n\, I(\hat{\theta}_n)}(\hat{\theta}_n - \theta) \leq u_{\frac{\alpha}{2}}\right) = 1-\alpha$$

d'où l'intervalle de confiance asymptÙtique :

$$\left[\hat{\theta}_n - \frac{u_{\frac{\alpha}{2}}}{\sqrt{n\, I(\hat{\theta}_n)}}, \hat{\theta}_n + \frac{u_{\frac{\alpha}{2}}}{\sqrt{n\, I(\hat{\theta}_n)}}\right].$$

Intervalle de Confiance Asymptotique : Exemple

marque On a mentionnÈ l'EMV car il est asymptÙtiquement normal, donc il converge en loi vers une loi connue (ici la loi normale).

ple 3.5 Soit $X_1, ..., X_n$ un Èchantillon de loi uniforme $U([0, \theta])$.

L'Estimateur de Maximum de Vraisemblance : $\hat{\theta}_n = \sup\limits_{1 \leq i \leq n} X_i$.

Soit $T_n := n\left(1 - \frac{\hat{\theta}_n}{\theta}\right)$, T_n est une statistique asymptÙtiquement libre pour θ, en effet ;

$$\begin{array}{rcl} F_{T_n}(x) & = & P(T_n \leq x) \\ & = & P\left(\hat{\theta}_n \geq \theta(1 - \frac{x}{n})\right)^n \\ & = & 1 - (1 - \frac{x}{n})^n \end{array}$$

et $\lim\limits_{n \to \infty} F_{T_n}(x) = 1 - e^{-x}$. Donc, $T_n \to T$, où T suit une loi exponentielle $\mathscr{E}xp(1)$.

Intervalle de Confiance Asymptotique : Exemple

On cherche un intervalle de la forme $[\hat{\theta}_n\,;\,b]$ tel que pour un α donnÈ on ait

$$P(\hat{\theta}_n \leq \theta \leq b) = 1 - \alpha.$$

$$\begin{aligned} P(\hat{\theta}_n \leq \theta \leq b) &= P\left(0 \leq T_n \leq n\left(1 - \frac{\hat{\theta}_n}{b}\right)\right) \\ &\approx_{n\to\infty} F_T\left(n(1 - \frac{\hat{\theta}_n}{b})\right) = 1 - \alpha \end{aligned}$$

d'où $b = \frac{\hat{\theta}_n}{1 - \frac{1}{n} F_T^{-1}(1-\alpha)} = \frac{\hat{\theta}_n}{1 - \frac{1}{n}\log\alpha}$ On a donc

$$IC = \left[\hat{\theta}_n\,;\,\frac{\hat{\theta}_n}{1 - \frac{1}{n}\log\alpha}\right]$$