

Chapitre V : Intervalles de confiance

MATH-F-207

Davy Paindaveine

Bloc 2, Bachelier en sciences mathématiques
Université libre de Bruxelles

2023–2024

Contenu du chapitre

Introduction

Intervalles exacts

Intervalles asymptotiques

Lien avec les tests d'hypothèses

Zones de confiance

Contenu du chapitre

Introduction

Intervalles exacts

Intervalles asymptotiques

Lien avec les tests d'hypothèses

Zones de confiance

Introduction

Soit le modèle statistique paramétrique à paramètre scalaire

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

où les mesures de probabilité $P_{\theta}^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \dots, X_n)$.

L'estimation ponctuelle de θ fournit une valeur unique, $T(X^{(n)}) \in \Theta$, de θ .

On cherche maintenant à donner une collection $S(X^{(n)}) \subset \Theta$ de valeurs de θ contenant la vraie valeur de θ avec une certaine probabilité cible.

Définition

Ce sous-ensemble $S(X^{(n)})$ est une **zone de confiance**.

S'il s'agit d'un intervalle, on parle d'**intervalle de confiance (IC)**.

La probabilité cible est le **niveau de confiance**.

Définition 1

Soit un modèle statistique paramétrique à paramètre scalaire. Soit $\alpha \in]0, 1[$. Alors $[L(X^{(n)}), U(X^{(n)})]$ est un IC pour θ au niveau de confiance $1 - \alpha$ si et seulement si les statistiques $L(X^{(n)})$ et $U(X^{(n)})$ vérifient les deux conditions suivantes pour tout $\theta \in \Theta$:

- (i) $L(X^{(n)}) \leq U(X^{(n)})$ $P_\theta^{(n)}$ -p.s.*
- (ii) $P_\theta^{(n)}[\theta \in [L(X^{(n)}), U(X^{(n)})]] \geq 1 - \alpha$.*

Attention

On ne dira pas

“ $[L(x^{(n)}), U(x^{(n)})]$ contient la vraie valeur de θ avec probabilité $\geq 1 - \alpha$ ”

On dira plutôt

“ $[L(x^{(n)}), U(x^{(n)})]$ est la réalisation d'un intervalle aléatoire contenant la vraie valeur de θ avec probabilité $\geq 1 - \alpha$ ”

Contenu du chapitre

Introduction

Intervalles exacts

Intervalles asymptotiques

Lien avec les tests d'hypothèses

Zones de confiance

Exemple du verre de bière (avec variance connue)

Soit $X^{(n)} = (X_1, \dots, X_n)$, où les X_i sont i.i.d. $\mathcal{N}(\mu, \sigma_0^2)$, avec σ_0^2 un réel strictement positif fixé. Sous $P_\mu^{(n)}$, la variable aléatoire

$$g(X^{(n)}, \mu) = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma_0}$$

est de loi $\mathcal{N}(0, 1)$. Donc

$$\forall \mu \in \mathbb{R}, \quad P_\mu^{(n)} \left[\underbrace{\Phi^{-1}\left(\frac{\alpha}{2}\right)}_{=-z_{\alpha/2}} \leq \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma_0} \leq \underbrace{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}_{=z_{\alpha/2}} \right] = 1 - \alpha,$$

ce qui se réécrit

$$\forall \mu \in \mathbb{R}, \quad P_\mu^{(n)} \left[\bar{X} - z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}} \right] = 1 - \alpha.$$

Intervalle de confiance gaussien

Par conséquent,

$$[L(X^{(n)}), U(X^{(n)})] \stackrel{\text{def}}{=} \left[\bar{X} - z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}} \right]$$

est un IC pour μ au niveau de confiance $1 - \alpha$.

La longueur de l'IC

$$2z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}$$

est une fonction décroissante de n , mais croissante de σ_0 et de $1 - \alpha$.

Intervalle asymétriques

Pour tout $u \in]0, \alpha[$ fixé, on a aussi

$$\forall \mu \in \mathbb{R}, \quad P_{\mu}^{(n)} \left[\Phi^{-1}(\alpha - u) \leq \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma_0} \leq \Phi^{-1}(1 - u) \right] = 1 - \alpha,$$

ce qui mène à l'IC

$$\begin{aligned} [L_u(X^{(n)}), U_u(X^{(n)})] &\stackrel{\text{def}}{=} \left[\bar{X} - \Phi^{-1}(1 - u) \frac{\sigma_0}{\sqrt{n}}, \bar{X} - \Phi^{-1}(\alpha - u) \frac{\sigma_0}{\sqrt{n}} \right] \\ &= \left[\bar{X} - z_u \frac{\sigma_0}{\sqrt{n}}, \bar{X} + z_{\alpha - u} \frac{\sigma_0}{\sqrt{n}} \right]. \end{aligned}$$

L'IC d'abord considéré, celui associé à $u = \alpha/2$, est celui de longueur minimale (en $u \in]0, \alpha[$). Il respecte aussi les propriétés de symétrie du problème.

Fonction pivotale

C'est le fait qu'on dispose d'une **fonction pivotale** qui a rendu possible la construction de l'IC.

Définition 2

Désignons par $\mathcal{X}^{(n)}$ l'ensemble des valeurs possibles de $x^{(n)}$. Alors **la fonction $g : \mathcal{X}^{(n)} \times \Theta \rightarrow \mathbb{R}$ est pivotale pour θ** si et seulement si les trois conditions suivantes sont satisfaites:

- (i) $x^{(n)} \mapsto g(x^{(n)}, \theta)$ est mesurable pour tout $\theta \in \Theta$;
- (ii) $\theta \mapsto g(x^{(n)}, \theta)$ est strictement monotone pour tout $x^{(n)} \in \mathcal{X}^{(n)}$;
- (iii) la distribution de $g(X^{(n)}, \theta)$ sous $P_\theta^{(n)}$ ne dépend pas de θ .

Méthode des fonctions pivotales

Si g est pivotale, alors il existe des réels a, b **indépendants de θ** tels que

$$\forall \theta \in \Theta, \quad P_{\theta}^{(n)} \left[a \leq g(X^{(n)}, \theta) \leq b \right] \geq 1 - \alpha.$$

Puisque $\theta \mapsto g(X^{(n)}, \theta)$ est strictement monotone, elle est inversible, ce qui permet d'écrire le résultat ci-dessus sous la forme

$$\forall \theta \in \Theta, \quad P_{\theta}^{(n)} \left[\theta \in g^{-1}(X^{(n)}, \cdot)([a, b]) \right] \geq 1 - \alpha,$$

où $g^{-1}(x^{(n)}, \cdot)([a, b])$ est l'image de $[a, b]$ par la réciproque de $\theta \mapsto g(x^{(n)}, \theta)$.

Comme on peut montrer que cette réciproque est continue, $g^{-1}(X^{(n)}, \cdot)([a, b])$ est un **intervalle de confiance** pour θ au niveau de confiance $1 - \alpha$.

Remarques

- ▶ On essayera, si possible, d'avoir " $= \alpha$ " (plutôt que " $\geq \alpha$ ").
- ▶ a et b ne sont pas uniques (// l'exemple du verre de bière)
- ▶ Souvent, a et b peuvent être fixés en identifiant l'IC de longueur minimale.

Exemple du bus

Soit $X^{(n)} = (X_1, \dots, X_n)$, où les X_i sont i.i.d. $\text{Unif}([0, \theta])$, avec $\theta \in \Theta = \mathbb{R}_0^+$.

La fonction

$$g(x^{(n)}, \theta) = \frac{x^{(n)}}{\theta}$$

est pivotale pour θ .

Les IC pour θ au niveau de confiance $1 - \alpha$ qui en résultent prennent la forme

$$[L_u(X^{(n)}), U_u(X^{(n)})] = \left[\frac{X_{(n)}}{(1-u)^{1/n}}, \frac{X_{(n)}}{(\alpha-u)^{1/n}} \right],$$

où $u \in [0, \alpha]$ est arbitraire (pour $u = \alpha$, l'IC est $[X_{(n)}, X_{(n)}/(1-\alpha)^{1/n}, +\infty[)$).

Exercice : montrer que $[X_{(n)}, X_{(n)}/\alpha^{1/n}]$ est l'IC de longueur minimale.

Paramètres de nuisance

Soit le modèle statistique paramétrique

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

où les mesures de probabilité $P_{\theta}^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \dots, X_n)$.

On souhaite construire un IC au niveau $1 - \alpha$ pour l'une des composantes θ_{ℓ} de $\theta = (\theta_1, \theta_2, \dots, \theta_k)$. Sans perte de généralité, on va se restreindre à $\ell = 1$.

Intervalle de confiance pour une composante

Définition 3

Soit le modèle statistique paramétrique ci-dessus. Soit $\alpha \in]0, 1[$.

Alors $[L(X^{(n)}), U(X^{(n)})]$ est un IC pour θ_1 au niveau de confiance $1 - \alpha$ si et seulement si les statistiques $L(X^{(n)})$ et $U(X^{(n)})$ vérifient les deux conditions suivantes pour tout $\theta \in \Theta$:

- (i) $L(X^{(n)}) \leq U(X^{(n)})$ $P_\theta^{(n)}$ -p.s.
- (ii) $P_\theta^{(n)} [\theta_1 \in [L(X^{(n)}), U(X^{(n)})]] \geq 1 - \alpha$.

Fonction pivotale pour une composante

Définition 4

Dans le cadre du modèle paramétrique ci-dessus, notons Θ_1 l'ensemble des valeurs que peut prendre le paramètre θ_1 lorsque θ parcourt Θ . Alors *la fonction $g : \mathbb{R}^n \times \Theta_1 \rightarrow \mathbb{R}$ est pivotale pour θ_1* si et seulement si les trois conditions suivantes sont satisfaites:

- (i) $x^{(n)} \mapsto g(x^{(n)}, \theta_1)$ est mesurable pour tout $\theta_1 \in \Theta_1$;
- (ii) $\theta_1 \mapsto g(x^{(n)}, \theta_1)$ est strictement monotone pour tout $x^{(n)} \in \mathcal{X}^{(n)}$;
- (iii) la distribution de $g(X^{(n)}, \theta_1)$ sous $P_\theta^{(n)}$ ne dépend pas de θ .

Exemple du verre de bière (avec variance inconnue)

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

Alors, en écrivant $s^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$,

$$g(x^{(n)}, \mu) = \frac{\sqrt{n-1}(\bar{x} - \mu)}{s}$$

est pivotale pour μ (si $n \geq 2$), et $g(X^{(n)}, \mu) \sim t_{n-1}$.

L'IC (symétrique) pour μ au niveau de confiance $1 - \alpha$ qui en résulte est

$$\left[\bar{X} - t_{n-1, 1-\alpha/2} \frac{s}{\sqrt{n-1}}, \bar{X} + t_{n-1, 1-\alpha/2} \frac{s}{\sqrt{n-1}} \right].$$

Fonction pivotale pour σ^2

Dans le même modèle, la fonction

$$g(x^{(n)}, \sigma^2) = \frac{ns^2}{\sigma^2}$$

est pivotale pour σ^2 (si $n \geq 2$), et $g(X^{(n)}, \sigma^2) \sim \chi_{n-1}^2$.

Un intervalle de confiance pour σ^2 au niveau de confiance $1 - \alpha$ est donc

$$\left[\frac{ns^2}{\chi_{n-1, 1-\alpha/2}^2}, \frac{ns^2}{\chi_{n-1, \alpha/2}^2} \right].$$

On vérifiera encore que la longueur de l'IC est une fonction croissante de $1 - \alpha$ et décroissante de n .

Contenu du chapitre

Introduction

Intervalles exacts

Intervalles asymptotiques

Lien avec les tests d'hypothèses

Zones de confiance

Introduction

Soit le modèle statistique paramétrique

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

où les mesures de probabilité $P_{\theta}^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \dots, X_n)$.

En pratique, **trouver des fonctions pivotales est difficile!** De nombreux résultats asymptotiques permettent cependant de construire des **IC asymptotiques**.

Intervalle asymptotiques

Définition 5

Soit le modèle statistique paramétrique ci-dessus. Soit $\alpha \in]0, 1[$. Alors $[L(X^{(n)}), U(X^{(n)})]$ est un IC asymptotique pour θ_1 au niveau de confiance $1 - \alpha$ si et seulement si les statistiques $L(X^{(n)})$ et $U(X^{(n)})$ vérifient les deux conditions suivantes pour tout $\theta \in \Theta$:

(i) pour tout n , $L(X^{(n)}) \leq U(X^{(n)})$ $P_\theta^{(n)}$ -p.s.

(ii) $P_\theta^{(n)} [\theta_1 \in [L(X^{(n)}), U(X^{(n)})]] \rightarrow 1 - \alpha$ quand $n \rightarrow \infty$.

Une exemple paramétrique général

Supposons qu'on dispose d'un estimateur $T(X^{(n)}) = (T_1(X^{(n)}), \dots, T_k(X^{(n)}))$ de $\theta = (\theta_1, \dots, \theta_k)$ faiblement convergent et tel que, pour tout $\theta \in \Theta$,

$$\sqrt{n}(T_1(X^{(n)}) - \theta_1) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(\theta)) \quad \text{sous } P_\theta^{(n)}$$

quand $n \rightarrow \infty$, où $\sigma^2(\cdot)$ est une fonction continue. Alors la fonction

$$g(x^{(n)}, \theta_1) = \frac{\sqrt{n}(T_1(x^{(n)}) - \theta_1)}{\sigma(T(x^{(n)}))}$$

est “**asymptotiquement pivotale**”, et $g(X^{(n)}, \theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$ quand $n \rightarrow \infty$.

L'IC asymptotique pour θ_1 au niveau asymptotique $1 - \alpha$ qui en résulte est

$$\left[T_1(X^{(n)}) - z_{\alpha/2} \frac{\sigma(T(X^{(n)}))}{\sqrt{n}}, T_1(X^{(n)}) + z_{\alpha/2} \frac{\sigma(T(X^{(n)}))}{\sqrt{n}} \right].$$

Un exemple "semi-paramétrique"

Soit un modèle statistique d'échantillonnage pour lequel l'observation $X^{(n)} = (X_1, \dots, X_n)$ regroupe des variables aléatoires dont la loi commune admet des moments finis d'ordre 2. Soient $\mu = E_F[X_1]$ et $\sigma^2 = \text{Var}_F[X_1]$.

Alors la fonction

$$g(x^{(n)}, \mu) = \frac{\sqrt{n}(\bar{x}^{(n)} - \mu)}{s}$$

est **asymptotiquement pivotale**, et $g(X^{(n)}, \mu) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$ quand $n \rightarrow \infty$.

L'IC asymptotique pour μ au niveau asymptotique $1 - \alpha$ qui en résulte est

$$\left[\bar{X} - z_{\alpha/2} \frac{s}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \frac{s}{\sqrt{n}} \right].$$

Contenu du chapitre

Introduction

Intervalles exacts

Intervalles asymptotiques

Lien avec les tests d'hypothèses

Zones de confiance

Construction d'un test

Soit un modèle statistique paramétrique à paramètre scalaire

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

où les mesures de probabilité $P_{\theta}^{(n)}$ représentent les distributions possibles pour $X^{(n)} = (X_1, \dots, X_n)$.

Soit $[L(X^{(n)}), U(X^{(n)})]$ un IC pour θ au niveau de confiance $1 - \alpha$.

Pour le problème de test

$$\begin{cases} \mathcal{H}_0 : \theta = \theta_0 \\ \mathcal{H}_1 : \theta \neq \theta_0, \end{cases}$$

on peut alors considérer le test défini par

$$\phi(x^{(n)}) = \begin{cases} 1 & \text{si } \theta_0 \notin [L(x^{(n)}), U(x^{(n)})] \\ 0 & \text{si } \theta_0 \in [L(x^{(n)}), U(x^{(n)})]. \end{cases}$$

Construction d'un test

Le risque de première espèce de ϕ satisfait

$$\begin{aligned} E_{\theta_0}^{(n)}[\phi] &= P_{\theta_0}^{(n)}[\theta_0 \notin [L(X^{(n)}), U(X^{(n)})]] \\ &= 1 - P_{\theta_0}^{(n)}[\theta_0 \in [L(X^{(n)}), U(X^{(n)})]] \leq 1 - (1 - \alpha) = \alpha, \end{aligned}$$

de sorte que ϕ est de niveau α .

Il est donc possible, à l'aide d'un IC, de construire un test valide pour le problème de test bilatéral.

Construction d'un IC

Soit ϕ_{θ_0} un test (pur) de niveau α pour le problème de test bilatéral

$$\begin{cases} \mathcal{H}_0 : \theta = \theta_0 \\ \mathcal{H}_1 : \theta \neq \theta_0. \end{cases}$$

Posons

$$A(x^{(n)}) \stackrel{\text{def}}{=} \left\{ \theta_0 \in \Theta : \phi_{\theta_0}(x^{(n)}) = 0 \right\}.$$

Puisque

$$\mathbb{P}_{\theta}^{(n)}[\theta \in A(X^{(n)})] = \mathbb{P}_{\theta}^{(n)}[\phi_{\theta}(X^{(n)}) = 0] = 1 - \mathbb{P}_{\theta}^{(n)}[\phi_{\theta}(X^{(n)}) = 1] \geq 1 - \alpha,$$

l'ensemble aléatoire $A(X^{(n)})$ est une zone de confiance pour θ au niveau $1 - \alpha$.

Il n'y a pas de garantie qu'il s'agisse d'un IC.

Remarques

- ▶ Ceci permet une construction systématique d'IC (via les tests de rapport de vraisemblance).
- ▶ Les IC ainsi construits héritent de propriétés d'optimalité des tests associés.

Contenu du chapitre

Introduction

Intervalles exacts

Intervalles asymptotiques

Lien avec les tests d'hypothèses

Zones de confiance

Zones de confiance

Soit le modèle statistique paramétrique

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

où les mesures de probabilité $P_{\theta}^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \dots, X_n)$.

On souhaite construire une zone de confiance au niveau $1 - \alpha$ pour θ .

Définition 6

Soit un modèle statistique paramétrique et soit $\alpha \in]0, 1[$. Alors la statistique $R^{(n)}(X^{(n)})$, à valeurs dans l'ensemble des parties de Θ , est une zone de confiance pour θ au niveau de confiance $1 - \alpha$ si et seulement si pour tout $\theta \in \Theta$,

$$P_{\theta}^{(n)}[\theta \in R(X^{(n)})] \geq 1 - \alpha.$$

Remarques

- ▶ Ces zones ne doivent pas nécessairement être connexes.
En ce sens, la définition est plus générale que celle des IC.
- ▶ Comme ci-dessus, on peut construire ces zones de façon systématique.
Mais les zones obtenues ne seront que rarement des hyper-rectangles.
- ▶ Que faire si l'on désire des hyper-rectangles?

Hyper-rectangles

Supposons que, pour chaque $\ell = 1, \dots, k$, $[L_\ell(X^{(n)}), U_\ell(X^{(n)})]$ soit un IC au niveau $1 - \alpha$ fixé pour la ℓ ème composante θ_ℓ de paramètre $\theta = (\theta_1, \dots, \theta_k)$.

Il est naturel de considérer la zone de confiance hyper-rectangulaire

$$\begin{aligned} R(X^{(n)}) &= [L_1(X^{(n)}), U_1(X^{(n)})] \times \dots \times [L_k(X^{(n)}), U_k(X^{(n)})] \\ &= \left\{ \theta \in \Theta : \theta_\ell \in [L_\ell(X^{(n)}), U_\ell(X^{(n)})] \quad \forall \ell = 1, \dots, k \right\}. \end{aligned}$$

Mais il est crucial de remarquer que, typiquement,

$$P_\theta^{(n)}[\theta \in R(X^{(n)})] < P_\theta^{(n)}[\theta_\ell \in [L_\ell(X^{(n)}), U_\ell(X^{(n)})]]$$

pour tout ℓ , de sorte que le niveau de confiance associé à $R(X^{(n)})$ n'est pas clair.

Hyper-rectangles

Une solution possible est de considérer des IC $[L_\ell(X^{(n)}), U_\ell(X^{(n)})]$ a un niveau de confiance de $1 - \frac{\alpha}{k}$. En effet, on a alors

$$\begin{aligned} P_\theta^{(n)}[\theta \notin R(X^{(n)})] &= P_\theta^{(n)}\left[\bigcup_{\ell=1}^k [\theta_\ell \notin [L_\ell(X^{(n)}), U_\ell(X^{(n)})]]\right] \\ &\leq \sum_{\ell=1}^k \underbrace{P_\theta^{(n)}[\theta_\ell \notin [L_\ell(X^{(n)}), U_\ell(X^{(n)})]]}_{\leq \alpha/k} \\ &\leq \alpha, \end{aligned}$$

ce qui livre $P_\theta^{(n)}[\theta \in R(X^{(n)})] \geq 1 - \alpha$: l'hyper-rectangle $R(X^{(n)})$ est une zone de confiance au niveau de confiance $1 - \alpha$.

Conclusion

Dans ce chapitre, nous avons vu

- ▶ ce que sont des intervalles/zones/régions de confiance
- ▶ comment construire des intervalles à l'aide de fonctions pivotales, en présence ou non de paramètres de nuisance
- ▶ comment définir et construire des intervalles asymptotiques
- ▶ Le lien entre IC et tests d'hypothèses