

Mise à niveau

Comparaison d'estimateurs

A. Godichon-Baggioni

OBJECTIFS

On s'intéresse à l'estimation d'un paramètre $\theta \in \Theta$ d'une variable aléatoire X , avec Θ un intervalle ouvert de \mathbb{R} . Les objectifs sont donc :

- ▶ Comparer différents estimateurs de θ .
- ▶ Savoir si on peut parler d'estimateur optimal.

I. Généralités

COMPARAISON DES ERREURS QUADRATIQUES MOYENNES

On rappelle qu'une façon de quantifier la qualité d'un estimateur $\hat{\theta}_n$ de θ est de considérer son risque quadratique

$$\text{EQM}(\hat{\theta}_n, \theta) = \mathbb{E} \left[(\hat{\theta}_n - \theta)^2 \right].$$

On considère que $\hat{\theta}_n$ est un meilleur estimateur que $\tilde{\theta}_n$ si

$$\forall \theta \in \Theta, \quad \text{EQM}(\hat{\theta}_n, \theta) \leq \text{EQM}(\tilde{\theta}_n, \theta).$$

BIAIS D'UN ESTIMATEUR

On donne trop souvent trop d'importance au biais d'un estimateur! Débiaiser un estimateur ne donne pas forcément de meilleurs résultats!

Exemple : Soit $\theta > 0$ et $X \sim \mathcal{U}([0, \theta])$. Comparer les erreurs quadratiques moyennes des estimateurs suivants :

$$\hat{\theta}_n = X_{(n)}$$

$$\tilde{\theta}_n = \frac{n+1}{n} X_{(n)}$$

$$\theta_{\alpha,n} = \alpha X_{(n)}$$

avec $\alpha > 0$ choisi judicieusement.

APPROCHE ASYMPTOTIQUE

Le plus souvent, on dispose d'estimateurs $\hat{\theta}_n, \tilde{\theta}_n$ asymptotiquement normaux, i.e il existe σ_1^2 et σ_2^2 telles que

$$\sqrt{n} \left(\hat{\theta}_n - \theta \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \sigma_1^2 \right),$$
$$\sqrt{n} \left(\tilde{\theta}_n - \theta \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \sigma_2^2 \right).$$

Si $\sigma_1^2 \leq \sigma_2^2$, on choisira alors l'estimateur $\hat{\theta}_n$.

II. Information de Fisher

ABSOLUE CONTINUITÉ

Définition

On dit qu'une fonction f définie sur un intervalle ouvert I de \mathbb{R} est absolument continue sur I si il existe une fonction intégrable f' (appelée dérivée de f) telle que pour tout $[a, b] \subset I$, on ait

$$f(b) - f(a) = \int_a^b f'(x) dx.$$

Attention! La fonction f' n'est définie que presque partout (cf Théorème de Lebesgue).

EXPÉRIENCE ET MODÈLE STATISTIQUE

Définition

Une expérience statistique est la donnée d'un objet aléatoire \mathbf{X} dans un espace mesurable (E, \mathcal{E}) , et d'une famille de lois $(P_\theta)_{\theta \in \Theta}$ sur cet espace, supposé contenir la loi $P_{\mathbf{X}}$, et appelé modèle statistique pour la loi de \mathbf{X} .

Dans ce qui suit, Θ est un intervalle ouvert de \mathbb{R} , et on considère un modèle statistique de la forme

$(P_\theta)_{\theta \in \Theta} = (g_\theta, \mu)_{\theta \in \Theta}$ où μ est une mesure sur E .

MODÈLE RÉGULIER

Définition

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

- ▶ pour μ presque tout $x \in E$, l'application $\theta \mapsto g_\theta(x)$ est absolument continue sur Θ .
- ▶ pour tout $\theta_0 \in \Theta$, pour μ presque tout $x \in E$, l'application $\theta \mapsto g'_\theta(x)$ est continue en θ_0 .
- ▶ pour tout $\theta \in \Theta$, l'application

$$x \mapsto \frac{(g'_\theta(x))^2}{g_\theta(x)} \mathbf{1}_{g_\theta(x) > 0}$$

est intégrable sur E par rapport à la mesure μ et d'intégrale

$$I(\theta) = \int_E \frac{(g'_\theta(x))^2}{g_\theta(x)} \mathbf{1}_{g_\theta(x) > 0} \mu(dx)$$

continue sur Θ .

$I(\theta)$ est alors appelée Information de Fisher du modèle.

REMARQUE

Remarque : Si pour tout $x \in E$ la fonction $\theta \mapsto g_\theta(x)$ est de classe C^1 , alors les deux premiers points de la définition précédente sont vérifiés.

EXEMPLES

Exemple 1 : la loi exponentielle Soit $X \sim \mathcal{E}(\theta)$ avec $\theta \in \Theta = \mathbb{R}_+^*$, alors

$$g_\theta(x) = \theta e^{-\theta x} \quad \text{et} \quad \mu(dx) = \mathbf{1}_{x \geq 0} dx.$$

Le modèle est régulier et

$$I(\theta) = \frac{1}{\theta^2}$$

EXEMPLES

Exemples 2 : la loi de Bernoulli Soit $X \sim \mathcal{B}(\theta)$ avec $\theta \in \Theta = (0, 1)$. Alors $\mu(dx) = \delta_0 + \delta_1$ et

$$g_\theta(0) = 1 - \theta \quad \text{et} \quad g_\theta(1) = \theta.$$

Le modèle est régulier et on a

$$I(\theta) = \frac{1}{\theta(1 - \theta)}.$$

EXEMPLES

Exemple 3 : la loi uniforme Soit $X \sim \mathcal{U}([0, \theta])$ avec $\theta \in \Theta = \mathbb{R}_+^*$, on a alors

$$g_\theta(x) = \frac{1}{\theta} \mathbf{1}_{[0, \theta]}(x) \quad \text{et} \quad \mu(dx) = \mathbf{1}_{x \geq 0} dx.$$

Le modèle n'est pas régulier.

SCORE ET INFORMATION DE FISHER

Définition

Notons $l_\theta(\mathbf{X}) = \log(g_\theta(\mathbf{X}))$ la log-vraisemblance de \mathbf{X} . Si les deux premiers points de la définition précédente sont satisfaits et si l'application

$$\theta \longmapsto \mathbb{E}_\theta \left[(l'_\theta(\mathbf{X}))^2 \right]$$

est continue sur Θ , alors le modèle est régulier et d'information de Fisher

$$I(\theta) = \mathbb{E}_\theta \left[(l'_\theta(\mathbf{X}))^2 \right] = \mathbb{E}_\theta \left[\left(\frac{\partial}{\partial \theta} \log(g_\theta(\mathbf{X})) \right)^2 \right].$$

La variable aléatoire

$$l'_\mathbf{X}(\theta) = \frac{\partial}{\partial \theta} \log(g_\theta(\mathbf{X})) = \frac{g'_\theta(\mathbf{X})}{g_\theta(\mathbf{X})}$$

est appelée le score.

SCORE ET INFORMATION DE FISHER

Proposition

Si le modèle $(g_\theta)_{\theta \in \Theta}$ est régulier, alors

$$I(\theta) = \mathbb{V}_\theta [l'_X(\theta)].$$

EXEMPLES

Exemple 1 : la loi exponentielle Soit $\theta > 0$ et $X \sim \mathcal{E}(\theta)$. On retrouve bien

$$\mathbb{V} [l'_X(\theta)] = \frac{1}{\theta^2} = I(\theta)$$

Exemple 2 : la loi de Bernoulli Soit $\theta \in (0, 1)$ et $X \sim \mathcal{B}(\theta)$. On retrouve bien

$$\mathbb{V} [l'_\theta(X)] = \frac{1}{\theta(1-\theta)} = I(\theta)$$

INFORMATION DE FISHER D'UN ÉCHANTILLON

Proposition

Soit $\mathbf{X} = (X_1, \dots, X_n)$ où les X_i variables aléatoires i.i.d et X_1 admet f_θ comme densité par rapport à une mesure μ . Si le modèle $(f_\theta)_{\theta \in \Theta}$ est régulier et d'information de Fisher $I(\theta)$, alors le modèle produit de densité

$$g_\theta(x_1, \dots, x_n) = \prod_{i=1}^n f_\theta(x_i)$$

par rapport à la mesure produit $\mu^{\otimes n}$ est encore régulier et d'information de Fisher

$$I_n(\theta) = \mathbb{V} [l'_\theta(X_1, \dots, X_n)] = nI(\theta).$$

EXEMPLES

Exemple 1 : la loi exponentielle Soit $\theta > 0$ et $\mathbf{X} = (X_1, \dots, X_n)$ où les X_i sont i.i.d avec $X_1 \sim \mathcal{E}(\theta)$. On retrouve bien

$$I_n(\theta) = \mathbb{V} [l'_\theta (X_1, \dots, X_n)] = \frac{n}{\theta^2} = nI(\theta)$$

Exemple 2 : la loi de Bernoulli Soit $\theta \in (0, 1)$ et $\mathbf{X} = (X_1, \dots, X_n)$ où les X_i sont i.i.d avec $X_1 \sim \mathcal{B}(\theta)$. On retrouve bien

$$I_n(\theta) = \mathbb{V} [l'_\mathbf{X}(\theta)] = \frac{n}{\theta(1-\theta)} = nI(\theta).$$

III. Inégalité de l'Information et borne de Cramer-Rao

OBJECTIF

L'objectif est de chercher à minorer l'erreur quadratique moyenne des estimateurs de θ . Dans ce qui suit on considère $\mathbf{X} = (X_1, \dots, X_n)$ où les X_i sont i.i.d et un estimateur $\hat{\theta}_n = \theta(\mathbf{X})$ de θ . On notera (f_θ, μ) le modèle associé à X_1 .

INÉGALITÉ DE L'INFORMATION

Proposition

Soit $(f_\theta, \mu)_{\theta \in \Theta}$ un modèle régulier et $\hat{\theta}_n$ un estimateur de θ dont l'erreur quadratique moyenne est localement bornée, et de biais

$b(\theta) = \mathbb{E}_\theta [\hat{\theta}_n] - \theta$. Alors, si $I(\theta) > 0$,

$$EQM(\hat{\theta}_n, \theta) = \mathbb{E}_\theta \left[(\hat{\theta}_n - \theta)^2 \right] \geq b(\theta)^2 + \frac{(1 + b'(\theta))^2}{I_n(\theta)}.$$

REMARQUE

Remarque : L'inégalité de l'Information de nous aide en aucun cas pour comparer des estimateurs. La borne est propre à chaque estimateur et l'atteindre ne signifie aucunement que l'on a un bon estimateur.

Exemple : On prend $\theta \in \mathbb{R}$ et $\hat{\theta}_n = 0$. On a bien

$$\text{EQM}(\hat{\theta}_n, \theta) = b(\theta)^2 + \frac{(1 + b'(\theta))^2}{I_n(\theta)}$$

mais l'estimateur "n'est pas bon".

BORNE DE CRAMER-RAO

Corollaire

Soit $(f_\theta, \mu)_{\theta \in \Theta}$ un modèle régulier et $\hat{\theta}_n$ un estimateur sans biais de θ de variance localement bornée, alors

$$EQM(\hat{\theta}_n, \theta) = \mathbb{V}_\theta[\hat{\theta}_n] \geq \frac{1}{I_n(\theta)}$$

Un estimateur atteignant cette borne est dit efficace.

EXEMPLE

Exemple : la loi de Bernoulli Soient X_1, \dots, X_n des variables aléatoires indépendantes suivant une loi de Bernoulli de paramètre $\theta \in (0, 1)$. L'estimateur $\hat{\theta}_n = \bar{X}_n$ est efficace.

IV. Efficacité asymptotique

EFFICACITÉ ASYMPTOTIQUE

Définition

Soit $(f_\theta)_{\theta \in \Theta}$ un modèle régulier d'information $I(\theta)$ ne s'annulant pas sur Θ . Soit $\hat{\theta}_n$ un estimateur de θ . On dit que $\hat{\theta}_n$ est asymptotiquement efficace si il existe $\sigma^2 \leq I(\theta)^{-1}$ tel que

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

EXEMPLES

Exemple 1 : la loi de Bernoulli Soient X_1, \dots, X_n des variables aléatoires indépendantes suivant une loi de Bernoulli de paramètre $\theta \in (0, 1)$. L'estimateur $\hat{\theta}_n$ est asymptotiquement efficace.

Exemple 2 : la loi exponentielle Soient X_1, \dots, X_n des variables aléatoires indépendantes suivant une loi exponentielle de paramètre $\theta > 0$. L'estimateur $\hat{\theta}_n = \bar{X}_n^{-1}$ est asymptotiquement efficace.

EMV ET EFFICACITÉ ASYMPTOTIQUE

Théorème

Soit $(f_\theta)_{\theta \in \Theta}$ un modèle régulier d'information de Fisher $I(\theta)$. Soit $\theta_0 \in \Theta$ vérifiant $I(\theta_0) > 0$ et X_1, \dots, X_n des variables aléatoires i.i.d de densité f_{θ_0} . Si il existe un estimateur du maximum de vraisemblance $\hat{\theta}_n$ qui converge vers θ_0 , et si il existe $h > 0$ tel que

$$\mathbb{E}_{\theta_0} \left[\sup_{\theta_0 - h \leq \theta \leq \theta_0 + h} l'_X(\theta)^2 \right] < +\infty,$$

alors

$$\sqrt{n} \left(\hat{\theta}_n - \theta_0 \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{1}{I(\theta_0)} \right).$$

REMARQUE

Remarque : Si le modèle est régulier, il n'est pas possible de trouver un estimateur qui soit meilleur qu'un estimateur asymptotiquement efficace pour tout $\theta \in \Theta$. Cependant, il est possible d'en trouver qui soient meilleur pour tout $\theta \in \Theta_0$ où Θ_0 est de mesure de Lebesgue nulle.

Exemple : estimateur de Hodge : Soit $X \sim \mathcal{N}(\theta, 1)$ et on considère l'estimateur

$$\hat{\theta}_n = \bar{X}_n \mathbf{1}_{|\bar{X}_n| > n^{-1/4}}.$$

Que se passe-t-il si $\theta = 0$?