Intégration - Résumé

October 25, 2023

THEVENET Louis

Table des matières

TD1 : Maximum & méthode des moments

1. Estimation

Qualités d'un estimateur

1.1. Modèle statistique, estimateurs

Définition 1.1.1:

On note $\hat{\theta}(X_1, ..., X_n)$, $\hat{\theta}_n$ ou $\hat{\theta}$ l'estimateur lié aux n VA *iid* $X_1, ..., X_n$ elles-mêmes liées aux n observations $x_1, ..., x_n$

- Biais : $b_n(\theta) = E\left(\hat{\theta}_n\right) - \theta \in \mathbb{R}^p$
- Variance : $v_n(\theta) = E\left[\left(\hat{\theta}_n - E\left(\hat{\theta}_n\right)\right)^2\right]$
- Matrice de covariance : $E\left[\left(\hat{\theta}_n - E\left(\hat{\theta}_n\right)\right)\left(\hat{\theta}_n - E\left(\hat{\theta}_n\right)\right)^T\right]$
- Erreur quadratique moyenne (MSE) : $e_n(\theta) = E\left[\left(\hat{\theta}_n - \theta\right)^2\right] = v_n(\theta) + b_n^2(\theta)$
- un estimateur $\hat{\theta}_n$ est convergent si $\lim_{n \to +\infty} b_n(\theta) = \lim_{n \to +\infty} v_n(\theta) = 0$

1.2. Inégalité de Cramér Rao

Théorème 1.2.1:

$$\operatorname{Var}\left(\hat{\theta}_n\right) \geq \frac{\left[1 + b_n'(\theta)\right]^2}{\left(-E\left[\frac{\partial^2 \ln(L(X_1,...,X_n;\theta))}{\partial \theta^2}\right)\right]}) = \operatorname{BCR}(\theta)$$

- BCR : Borne de Cramér-Rao
- $L(X_1, ..., X_n; \theta)$: vraisemblance
- **Hypothèses** :
  1. log-vraisemblance deux fois dérivable
  2. suport de la loi indépendant de θ

1.3. Maximum de vraisemblance

Définition 1.3.1: Maximum de vraisemblance

$$\hat{\theta}_{\mathrm{MV}} = \arg\max_{\theta} L(X_1, ..., X_n; \theta)$$

Théorème 1.3.1: Recherche de $\hat{\theta}_{\mathrm{MV}}$

- Cherche les points fixes de la vraisemblances ou de la log-vraisemblances
- Tableau de variations pour vérifier ou alors étudier $\frac{\partial^2 \ln L\left(X_1,...,X_n;\hat{\theta}_{\mathrm{MV}}\right)}{\partial \theta^2} < 0$

Définition 1.3.2: Régularité

Comment construire un estimateur

1.4. Méthode des moments

1.5. Estimation Bayésienne

Définition 1.5.1: Estimation Bayésienne

On va estimer un paramètre inconnu $\theta \in \mathbb{R}^p$ à l'aide de l'estimation paramétrée par θ, et ‹ une loi à priori $p(\theta)$. Pour celà on minimise une fonction de coût $c\left(\theta, \hat{\theta}\right)$ qui représente l'erreur entre θ et $\hat{\theta}$. Deux estimateurs principaux :

- MMSE : moyenne de la loi à posteriori $\hat{\theta}_{\text{MSEE}} = E(\theta(X_1, ..., X_n))$
- MAP : estimateur du maximum à posteriori de θ est définie par $\hat{\theta}_{\text{MAP}} = \underset{\theta}{\arg\max}\, p(\theta \mid X_1, ..., X_n)$

Théorème 1.5.1: MMSE

L'estimateur MMSE minimise l'erreur quadratique moyenne (Root Mean Square)
On a

$$c\left(\theta, \hat{\theta}\right) = E\left[\left(\theta - \hat{\theta}\right)^T \left(\theta - \hat{\theta}\right)\right]$$

Théorème 1.5.2: MAP

à vérifier si ça minimise la moyenne ou la f° de coût tout court On a :

- L'estimateur MAP minimise la fonction de coût $E\left[c\left(\theta, \hat{\theta}\right)\right]$ avec

$$c\left(\theta, \hat{\theta}\right) = \begin{cases} 1 \text{ si } \left\|\theta - \hat{\theta}\right\| > \delta \\ 0 \text{ si } \left\|\theta - \hat{\theta}\right\| < \delta \end{cases}$$

1.6. Intervalles de confiance

2. Tests Statistiques

Définition 2.1:

- Risque de première espèce (fausse alarme) : $\alpha = \text{PFA} = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}]$
- Risque de seconde espèce (non détection) : $\beta = \text{PND} = P[\text{Rejeter } H_1 \mid H_1 \text{ vraie}]$
- Puissance du test (proba de détection) : $\pi = 1 - \beta$

Pour faire un test, on a H_0, H_1 etc bien posées

Statistique du test ? $T(x_1, ..., x_n)$

Règle du test : $\equiv$ zone critique

ex1 : si $T(x_1, ..., x_n) \begin{cases} > S_\alpha \text{: rejet } H_0 \\ < S_\alpha \text{: accepte } H_0 \end{cases}$