

PERFORMANCES

Sommaire

1.1	Qualité d'un estimateur	2
-----	-----------------------------------	---

Dans ce chapitre nous allons étudier comment comparer des estimateurs. Un estimateur T_n de θ sera un bon estimateur s'il s'approche suffisamment du paramètre dans un sens que nous allons préciser.

1.1 Qualité d'un estimateur

1.1.1 Risque quadratique, biais et variance

Définition : *risque quadratique.*

Le risque quadratique ou erreur quadratique moyenne d'un estimateur T_n est donné par l'expression suivante :

$$EQM(T_n) = \mathbb{E}[(T_n - \theta)^2]$$

Le risque quadratique permet de comparer deux estimateurs, c'est une mesure assez brute de la qualité d'un estimateur qui peut être raffiné avec les notions de biais et de variance.

Définition : *bias d'un estimateur*

On appelle biais de T_n pour θ la valeur $b_\theta(T_n) := \mathbb{E}(T_n) - \theta$

Par conséquent on dira d'un estimateur T_n de θ qu'il est sans biais si et seulement si $\mathbb{E}(T_n) = \theta$, dans le cas contraire, on dira que T_n est un estimateur biaisé. En d'autres termes un estimateur est sans biais si son espérance est égale au paramètre recherché.

Exercice : Montrer que $EQM(T_n) = Var(T_n) + b_\theta(T_n)^2$

Le risque quadratique d'un estimateur est égal à sa variance plus le carré de son biais. Lorsqu'on évalue plusieurs estimateurs le meilleur sera sans biais et de variance minimale. Cependant, le biais et la variance ne sont pas les seules propriétés caractérisant la qualité d'un estimateur. En effet, augmenter la taille de l'échantillon est aussi à prendre en considération quand on compare des estimateurs. D'ailleurs le succès du big data revient à utiliser d'énormes échantillons, ce qui conduit à bénéficier des performances asymptotiques des

estimateurs.

Exercice : On considère une variable aléatoire X telle que $\mathbb{E}[X] = \mu$ et $\text{Var}(X) = \sigma^2$.

- Montrer que les estimateurs de la moyenne et de la variance d'une loi Gaussienne $\mathcal{N}(\mu, \sigma^2)$ que vous avez trouvé à l'aide du maximum de vraisemblance sont sans biais.
- Montrer que l'estimateur de la variance empirique $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ est biaisé. Pour cela commencez par développer et simplifier cette somme.
- Déduisez en un estimateur de la variance non biaisé et retrouvez l'estimateur de la variance empirique corrigé présenté plus haut \hat{S}_n^2 .