

Biais et variance : exemple

Exemple Soient $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, $\hat{\mu} = \bar{X}_n$, et $\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$. On admet que $\text{var}_{\mu, \sigma^2}(\sum_{i=1}^n (X_i - \bar{X}_n)^2) = 2\sigma^4(n-1)$. Trouver le biais et la variance de $\hat{\mu}$. Trouver les valeurs de a qui minimisent le biais, la variance, et la EQM, pour $\hat{\sigma}^a := a\hat{\sigma}_n^2$.

Solution (début) : On a $\bar{X}_n \sim \mathcal{N}(\mu, \sigma^2/n)$ donc $\mathbb{E}_{\mu, \sigma^2}[\bar{X}] = \mu$ et $\text{var}_{\mu, \sigma^2}(\bar{X}_n) = \frac{\sigma^2}{n} = \text{EQM}_{\mu, \sigma^2}(\bar{X}_n)$ ($\hat{\mu}$ non-biaisé).

Pour $\hat{\sigma}_n^2$ on utilise la formule alternative $\hat{\sigma}_n^2 = \sum_{i=1}^n X_i^2/n - \bar{X}_n^2$. On a $\mathbb{E}_{\mu, \sigma^2}[X_1^2] = (\mathbb{E}_{\mu, \sigma^2}[X_1])^2 + \text{var}_{\mu, \sigma^2}[X_1] = \mu^2 + \sigma^2$ et de manière similaire $\mathbb{E}_{\mu, \sigma^2}[\bar{X}_n^2] = (\mathbb{E}_{\mu, \sigma^2}[\bar{X}_n])^2 + \text{var}_{\mu, \sigma^2}[\bar{X}_n] = \mu^2 + \sigma^2/n$. Donc

$$\mathbb{E}_{\mu, \sigma^2}[\hat{\sigma}^2] = \mathbb{E}_{\mu, \sigma^2}[X_1^2] - \mathbb{E}_{\mu, \sigma^2}[\bar{X}_n^2] = \mu^2 + \sigma^2 - (\mu^2 + \sigma^2/n) = \sigma^2(1 - \frac{1}{n}).$$

Ainsi

$$b_{\mu, \sigma^2}[a\hat{\sigma}^2] = a\mathbb{E}_{\mu, \sigma^2}[\hat{\sigma}^2] - \sigma^2 = \sigma^2(a - \frac{1}{n} - 1)$$