

# Méthode des moments

- Supposons que l'échantillon tiré soit représentatif de la population
- Pour obtenir des estimateurs pour les paramètres inconnus de la population, on égalise les "moments" de l'échantillon ("empirique") à ceux de la population ("théorique")
- k-ème moment  $k \in \{1, 2, 3, \dots, \infty\}$ 
  - Population ("théorique") :  $m_k = \mathbb{E}_\theta(X^k) = m_k(\theta)$ . Comme la loi de  $X$  dépend de  $\theta$ ,  $m_k = m_k(\theta)$
  - Echantillon ("empirique") :  $\hat{m}_k = \frac{1}{n} \sum_{i=1}^n X_i^k \rightarrow m_k(\theta)$
  - L'estimateur des moments s'obtient on égalisant  $m_k$  et  $\hat{m}_k$ , ce qui donne des équation(s) pour  $\theta \in \mathbb{R}^p$
  - Par exemple  $m_1(\theta) = \mathbb{E}_\theta(X) = \sum_{i=1}^n X_i/n = \bar{X}_n$
- On a donc besoin d'autant de moments (supposés finies!) que de paramètres inconnus

**Exemple** Soient  $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} U(0, \theta)$ . Estimer  $\theta$ .

**Exemple** Soient  $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ . Estimer  $\mu$  et  $\sigma^2$ .

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

## Méthode des moments

**Exemple** Soient  $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} U(0, \theta)$ . Estimer  $\theta$ .

$k=1$

$$m_1(\theta) = E_{\theta}[X_i] = \dots = \frac{\theta}{2}$$

$$\hat{m}_1 = \bar{X}_n \Rightarrow \text{Equation} \quad \frac{\theta}{2} = \bar{X}_n$$

$$\text{Donc } \boxed{\hat{\theta}_{\text{mom}} = 2\bar{X}_n}$$

$$k=2 \Rightarrow \hat{\theta}_{\text{mom}} = \sqrt{\frac{3}{2} \sum_{i=1}^n X_i^2}$$

# Méthode des moments

**Exemple** Soient  $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ . Estimer  $\mu$  et  $\sigma^2$ .

$k=1$   
 $k=2$

$$\begin{cases} \bar{X}_n = \bar{\mu}_1 = \mathbb{E}_{\mu, \sigma^2}[X_i] = \mu \\ \frac{1}{n} \sum_{i=1}^n X_i^2 = \bar{\mu}_2 = \mathbb{E}_{\mu, \sigma^2}[X_i^2] = \text{var} + (\mathbb{E} X_i)^2 = \sigma^2 + \mu^2 \end{cases}$$

$$\boxed{\beta_{\text{mom}} = \bar{X}_n}$$

$$\sigma_{\text{mom}}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left( \beta_{\text{mom}} \right)^2$$

$$\text{var}(\bar{Y}) = \mathbb{E}(\bar{Y}^2) - (\mathbb{E}\bar{Y})^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \sigma_{\text{mom}}^2$$

# Méthode du maximum de vraisemblance

**Définition:** Soient  $x_1, \dots, x_n$  des données supposées être une réalisation d'un échantillon aléatoire  $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} f(x; \theta)$ . La **vraisemblance** pour  $\theta$  est la fonction

$$L(\theta) = f(x_1; \theta) \times f(x_2; \theta) \times \dots \times f(x_n; \theta) = \prod_{i=1}^n f(x_i; \theta).$$

$f$  est la fonction de densité ou de masse (on suppose qu'elle existe)

**Définition:** L'**estimateur du maximum de vraisemblance** (maximum likelihood)  $\hat{\theta}_{\text{ML}}$  d'un paramètre  $\theta$  est celui qui maximise la fonction de vraisemblance parmi tous les  $\theta$  possibles :

$$L(\hat{\theta}_{\text{ML}}) \geq L(\theta) \quad \text{pour tout } \theta$$

Il est plus facile de maximiser  $\ell(\theta) := \log L(\theta)$ , souvent en résolvant  $d\ell(\theta)/d\theta = 0$ , et vérifiant qu'il s'agit bien d'un maximum (par exemple si la deuxième dérivée est négative  $d^2\ell(\theta)/d\theta^2 < 0$ )

**Exemple**  $x_1, \dots, x_n$  réalisations d'une loi  $\exp(\lambda)$  avec  $\lambda > 0$ . Estimer  $\lambda$

## Exemple : maximum de vraisemblance

**Exemple** Supposons que  $x_1, \dots, x_n$  soient des réalisations i.i.d. d'une loi exponentielle,

$$f(x; \lambda) = \lambda e^{-\lambda x}, \quad x \geq 0, \quad \lambda > 0.$$

Trouver  $\hat{\lambda}_{ML}$ .

$$L(\lambda) = \prod_{i=1}^n f(x_i; \lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

$$l(\lambda) = \log L(\lambda) = n \log \lambda - \lambda \sum_{i=1}^n x_i = n \log \lambda - \lambda n \bar{x}_n$$

$$l'(\lambda) = \frac{n}{\lambda} - n \bar{x}_n = 0 \iff \lambda = \frac{1}{\bar{x}_n}$$

$$l''(\lambda) = -\frac{n}{\lambda^2} < 0 \implies \boxed{\lambda_{ML} = \frac{1}{\bar{x}_n}}$$

# Erreur quadratique moyenne

**Définition:** L'**erreur quadratique moyenne** de l'estimateur  $\hat{\theta}$  de  $\theta$  est

$$\text{EQM}_\theta(\hat{\theta}) = \mathbb{E}_\theta\{(\hat{\theta} - \theta)^2\} = \dots = \text{Var}_\theta(\hat{\theta}) + [b_\theta(\hat{\theta})]^2,$$

où  $b_\theta(\hat{\theta}) = \mathbb{E}_\theta(\hat{\theta}) - \theta$  est le **biais** de  $\hat{\theta}$

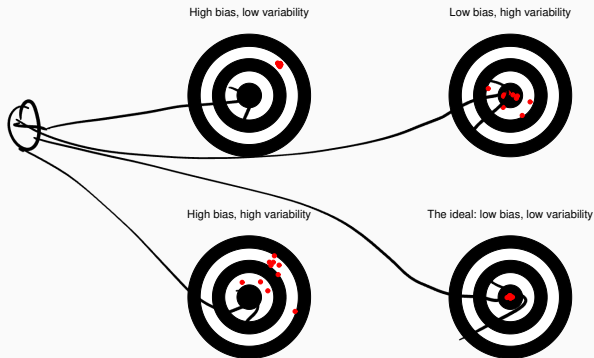
- La distribution de  $\hat{\theta}$  dépend de celle des  $X_i$  et donc de  $\theta$
- Si  $\hat{\theta}$  et  $\hat{\theta}'$  sont deux estimateurs du même paramètre  $\theta$  et  $\text{EQM}_\theta(\hat{\theta}) \leq \text{EQM}_\theta(\hat{\theta}')$ , on préfère  $\hat{\theta}$
- si  $b_\theta(\hat{\theta}) < 0$ , alors  $\hat{\theta}$  sous-estime  $\theta$
- si  $b_\theta(\hat{\theta}) > 0$ , alors  $\hat{\theta}$  sur-estime  $\theta$
- si  $b_\theta(\hat{\theta}) \equiv 0$ , alors  $\hat{\theta}$  est **non biaisé**, et  $\text{EQM}_\theta(\hat{\theta}) = \text{var}(\hat{\theta})$

$$\begin{aligned} \mathbb{E}[\bar{X}_n] &= \mu \\ \text{Var}(\bar{X}_n) &= \frac{\sigma^2}{n} \end{aligned}$$

**Exemple** Soient  $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ . La médiane  $M_n$  et la moyenne  $\bar{X}_n$  sont (à peu près) non-biaisés pour  $\mu$  mais  $\text{var}(M_n) > \text{var}(\bar{X}_n)$ . Lequel des estimateurs  $\bar{X}_n$  et  $M_n$  de  $\mu$  est préférable? Et si des valeurs aberrantes peuvent apparaître?

$$\mathbb{E}[M_n] \approx \mu \quad \text{var}(M_n) > \frac{\sigma^2}{n}$$

# Biais et variance



• réalisations  
de  $\hat{\theta}$

- $\theta$  = “bulle centrale”, supposée être la vraie valeur
- fléchettes rouges = réalisations de  $\hat{\theta}$  qui estime  $\theta$

**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  $\hat{\sigma}^a := a\hat{\sigma}_n^2$ .

## Biais et variance : exemple

**Exemple** Soient  $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ ,  $\hat{\mu}_n = \bar{X}_n$ , 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  $\hat{\mu}_n$ . Trouver les valeurs de  $a$  qui minimisent le biais, la variance, et la EQM, pour  $\hat{\sigma}^a := a\hat{\sigma}_n^2$ .

$$\mathbb{E}_{\mu, \sigma^2}[\hat{\mu}_n] = \mathbb{E}_{\mu, \sigma^2}[X_i] = \mu \rightarrow \text{non-biaisé}$$

$$\text{var}_{\mu, \sigma^2}[\hat{\mu}_n] = \frac{1}{n} \text{var}_{\mu, \sigma^2}[X_i] = \frac{\sigma^2}{n}$$

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i^2 - 2X_i\bar{X}_n + \bar{X}_n^2)$$

$$= \frac{1}{n} \left[ \left( \sum_{i=1}^n X_i^2 \right) - 2\bar{X}_n \sum_{i=1}^n X_i + n\bar{X}_n^2 \right] = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$$

## 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  $\hat{\sigma}^a := a\hat{\sigma}_n^2$ .

$$\begin{aligned} \text{var}_{\mu, \sigma^2}(a \hat{\sigma}_n^2) &= \text{var}_{\mu, \sigma^2}\left(\frac{\sum_{i=1}^n (X_i - \bar{X}_n)^2}{n} a\right) \\ &= \left(\frac{a}{n}\right)^2 \text{var}_{\mu, \sigma^2}\left(\sum_{i=1}^n (X_i - \bar{X}_n)^2\right) = \frac{a^2 2\sigma^4(n-1)}{n^2} \end{aligned}$$

Minimiser  $\text{var} \hat{\sigma}^a \rightarrow a_{\text{minvar}} = 0$

$$s_n^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x}_n)^2. \quad E(s_n^2) = E\left(\frac{1}{n} \sum_{i=1}^n x_i^2\right) - E(\bar{x}_n^2)$$

$$\Rightarrow \underbrace{\text{Var}}_{\mu^2} x_i + \underbrace{\left(E x_i\right)^2}_{\mu^2} - \left(\underbrace{\text{Var}}_{\frac{\sigma^2}{n}} \bar{x}_n + \underbrace{\left(E \bar{x}_n\right)^2}_{\mu^2}\right)$$

$$= \sigma^2 \left(1 - \frac{1}{n}\right) \cdot \frac{1}{n} \sigma^2 = \sigma^2 \left(1 - \frac{1}{n}\right) - \sigma^2$$

$$E(s_n^2) = \sigma^2 \left[ \left(\frac{(n-1)}{n}\right)^2 + \frac{2(n-1)}{n^2} \right]$$

$$\min \text{var} \rightarrow a=0$$

$$\min \text{bias}^2 \rightarrow a = \frac{n}{n-1}$$

$$\rightarrow \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

est non-biaisé pour  $\sigma^2$

$$\min \text{EQM}: \frac{\partial \text{EQM}}{\partial a} = \frac{6^4}{n^2} [4a(n-1)]$$

$$2[(a(n-1)-n)(n-1)] = 0 \Leftrightarrow a = \frac{n}{n-1}$$