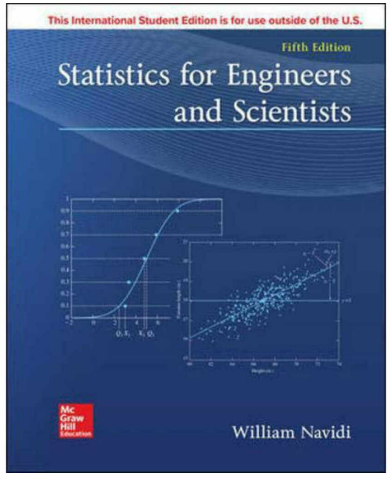
Introduction aux statistiques pour l'Ingénieur
Estimateurs et estimation

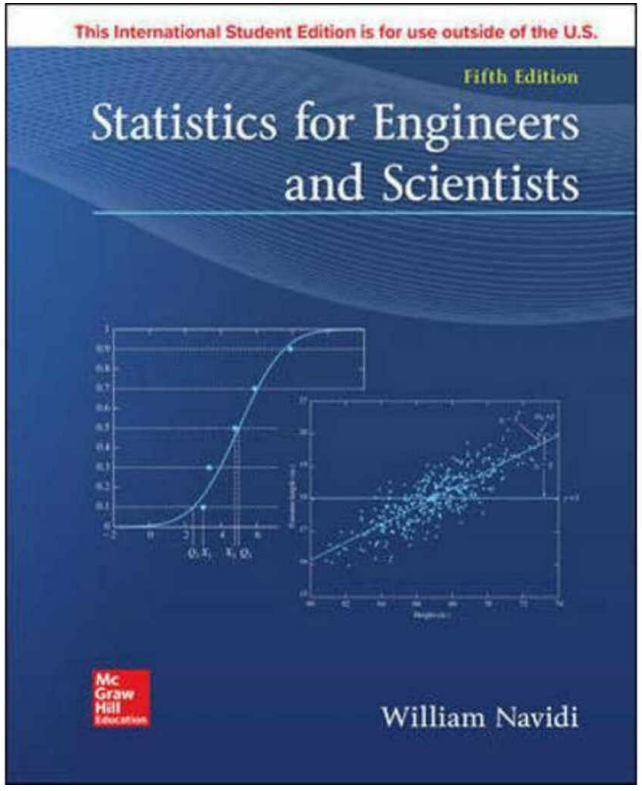
Stéphane Canu
asi.insa-rouen.fr/enseignants/~scanu
scanu@insa-rouen.fr



ITI 3, INSA Rouen Normandie, mars 2023

Lecture road map

1. Estimateurs



https://moodle.insa-rouen.fr/course/view.php?id=93

Motivations : le besoin d'estimateurs

- on souhaite connaitre de temps que met un client pour passer à la caisse
- on fait l'hypothèse que se temps suit une loi exponentielle de paramètre λ, et donc de densité

$$f(x) = \lambda e^{-\lambda x}$$

- pour inférer λ, on construit un échantillon $(X_1, \ldots, X_i, \ldots, X_n)$

Définition : Estimateur

Un estimateur $\widehat{\theta}$ d'un paramètre θ est une statistique, c'est à dire une fonction d'un échantillon

Deux questions associées :

- comment construire un estimateur ?
- comment prouver qu'un estimateur est préférable à un autre ?

comment construire un estimateur ?

- la méthode du maximum de vraisemblance
- la méthode des moments
- des techniques d'améliorations des estimateurs (θ est un vecteur)

Cadre formel

Modèle statistique

- Variable aléatoire X
- Loi parente $\mathbb{P}_\theta(x)$
- Paramètre $\theta \in \Theta$
- $\theta^\star$ la « vraie valeur » du paramètre (inconnue !)

Le but du statisticien

- A partir d'un échantillon i.i.d. $(X_1, \ldots, X_i, \ldots, X_n)$
- Construire un estimateur $\widehat{\theta}(X_1, \ldots, X_i, \ldots, X_n)$ de $\theta^\star$

Difficultés :

- l'estimateur $\widehat{\theta}(X_1, \ldots, X_i, \ldots, X_n)$ est une variable aléatoire
- le paramètre $\theta^\star$ est un réel
- Quel critère ?

Deux critères différents

Le risque statistique : écart moyen eentre $\widehat{\theta}$ et θ

$$R_{\widehat{\theta}}(\theta) = \mathbb{E}\big((\widehat{\theta}(X_1, \ldots, X_n) - \theta)^2\big)$$

La divergence de Kullback-Leibler (KL)

$$KL(P||Q) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)} = \int_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}$$

$$KL(P||Q) = \mathbb{E}_P \log \frac{P(x)}{Q(x)}$$

La divergence de Kullback-Leibler : écart entre $\mathbb{P}_{\theta^\star}(x)$ et $\mathbb{P}_{\widehat{\theta}}(x)$

$$KL\big(\mathbb{P}_{\theta^\star}(x)||\mathbb{P}_{\widehat{\theta}}(x)\big) = \mathbb{E}_{\theta^\star} \log \mathbb{P}_{\theta^\star}(X) - \mathbb{E}_{\theta^\star} \log \mathbb{P}_{\widehat{\theta}}(X)$$

La divergence de Kullback-Leibler (KL)

$$KL\big(\mathbb{P}_{\theta^\star}(x)||\mathbb{P}_{\widehat{\theta}}(x)\big) = \mathbb{E}_{\theta^\star} \log \mathbb{P}_{\theta^\star}(X) - \mathbb{E}_{\theta^\star} \log \mathbb{P}_{\widehat{\theta}}(X)$$

$$\begin{aligned}
\min_{\widehat{\theta}} KL\big(\mathbb{P}_{\theta^\star}(x)||\mathbb{P}_{\widehat{\theta}}(x)\big) &\iff \min_{\widehat{\theta}} \mathbb{E}_{\theta^\star} \log \mathbb{P}_{\widehat{\theta}}(X) \\
&\iff \min_{\widehat{\theta}} -\frac{1}{n}\sum_{i=1}^{n} \log \mathbb{P}_{\widehat{\theta}}(X_i) \\
&\iff \max_{\widehat{\theta}} \frac{1}{n}\sum_{i=1}^{n} \log \mathbb{P}_{\widehat{\theta}}(X_i) \\
&\iff \frac{\partial \sum_{i=1}^{n} \log \mathcal{L}_{X_i}(\widehat{\theta})}{\partial \widehat{\theta}} = 0
\end{aligned}$$

Le principe du maximum de vraisemblance

Une recette pour construire un estimateur $\widehat{\theta}(X_1, \ldots, X_i, \ldots, X_n)$ de $\theta^\star$

1. Modèle statistique : v.a. X de loi parente $\mathbb{P}_\theta(x)$ et paramètre $\theta \in \Theta$
   A partir d'un échantillon i.i.d. $(X_1, \ldots, X_i, \ldots, X_n)$
2. la fonction de vraisemblance : $L(\theta) = \prod_{i=1}^{n} \mathbb{P}_{\widehat{\theta}}(X_i)$
3. la log vraisemblance $\mathcal{L}(\theta) = \sum_{i=1}^{n} \log \mathbb{P}_{\widehat{\theta}}(X_i)$
4. on dérive $d\mathcal{L}(\theta) = \dfrac{\partial \sum_{i=1}^{n} \log \mathcal{L}_{X_i}(\widehat{\theta})}{\partial \widehat{\theta}}$
5. on résoud
$$d\mathcal{L}(\widehat{\theta}) = 0$$

Exemple : la loi normale

Loi normale

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\frac{(x-\mu)^2}{2\sigma^2}$$

$$\begin{aligned} L(s_n, \theta) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\frac{(x_i-\mu)^2}{2\sigma^2} \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp -\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2} \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp -\frac{\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2}{2\sigma^2} \end{aligned}$$

$$\begin{aligned} \mathcal{L}(s_n, \theta) &= \log L(s_n, \theta) \\ &= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2}{2\sigma^2} \end{aligned}$$

$$\partial\mathcal{L}(s_n, \theta) = -\frac{-2\sum_{i=1}^{n} x_i + 2n\mu}{2\sigma^2}$$