



I N S E A





4

0

2. Pourquoi la moyenne a posteriori, non pas la maximum a posteriori?



Idée: développer un critère pour mesurer

D'abord, en analyse fréquentiste:

“worst-case” analysis

1. On suppose qu'il existe un vrai paramètre θ tel que les données $X \sim \mathbb{P}_\theta(X)$.

2. On détermine une fonction T (estimateur) appliquée aux données pour estimer θ par $T(X)$.

3. On peut quantifier sa qualité en calculant une fonction de perte (loss function) $\mathcal{L}(\textcolor{blue}{T}(X), \textcolor{red}{\theta})$.

4. Choix classiques de \mathcal{L} : la perte quadratique $\mathcal{L}(a, b) = (a - b)^2$ ou en valeur absolue: $\mathcal{L}(a, b) = |a - b|$.

5. $\mathcal{L}(\textcolor{blue}{T}(X), \textcolor{red}{\theta})$ dépend des données, on calcule une perte moyenne dite ‘Risque’: $\mathcal{R}(\textcolor{blue}{T}, \textcolor{red}{\theta}) \stackrel{\text{def}}{=} \mathbb{E}_X [\mathcal{L}(\textcolor{blue}{T}(X), \textcolor{red}{\theta})]$

6. Ce Risque dépend de θ , on veut un estimateur T qui soit bon pour tous les θ .

7. On quantifie sa **pire performance** par le plus grand risque $\mathcal{R}(\textcolor{blue}{T}) \stackrel{\text{def}}{=} \max_{\textcolor{red}{\theta}} \mathcal{R}(\textcolor{blue}{T}, \textcolor{red}{\theta})$.

8. Le meilleur estimateur T est celui qui minimise ce **plus grand risque** $\min_T \max_{\theta} \mathcal{R}(T, \theta)$.

“critère mini-max”

Risque fréquentiste minmax

2. Pourquoi la moyenne a posteriori, non pas le maximum ou la médiane a posteriori ?

Idée: développer un critère pour comparer ces estimateurs

D'abord, en analyse fréquentiste:

1. On suppose qu'il existe un vrai paramètre θ tel que les données $X \sim \mathbb{P}_\theta(X)$.
2. On détermine une fonction T (estimateur) appliquée aux données pour estimer θ par $T(X)$.
3. On peut quantifier sa qualité en calculant une fonction de perte (loss function) $\mathcal{L}(T(X), \theta)$.
4. Choix classiques de \mathcal{L} : la perte quadratique $\mathcal{L}(a, b) = (a - b)^2$ ou en valeur absolue: $\mathcal{L}(a, b) = |a - b|$.
5. $\mathcal{L}(T(X), \theta)$ dépend des données, on calcule une perte moyenne dite "Risque": $\mathcal{R}(T, \theta) \stackrel{\text{def}}{=} \mathbb{E}_X [\mathcal{L}(T(X), \theta)]$
6. Ce Risque dépend de θ , on veut un estimateur T qui soit bon pour tous les θ .
7. On quantifie sa **pire performance** par le plus grand risque $\mathcal{R}(T) \stackrel{\text{def}}{=} \max_{\theta} \mathcal{R}(T, \theta)$. **"worst-case" analysis**
8. Le meilleur estimateur T est celui qui minimise ce **plus grand risque** $\min_T \max_{\theta} \mathcal{R}(T, \theta)$.

"critère mini-max"



1. Introduction

2. Les Bayésiens vs Les fréquentistes

3. Rappels de probabilités (exemples)

4. Loi a posteriori et modèles conjugués

5. Estimateur de Bayes



Le fréquentiste cherche l'estimateur T qui minimise:

$$\max_{\theta} \mathcal{R}(T, \theta) = \max_{\theta} \mathbb{E}_X [\mathcal{L}(T(X), \theta)]$$

