

CM 5: estimation paramétrique ponctuelle

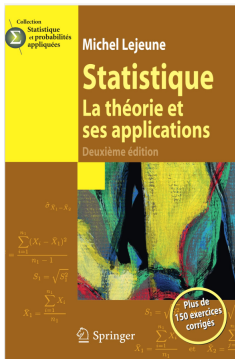
François Le Maître

francois.le-maitre@ube.fr

21 octobre 2025

Référence bibliographique

Ce cours est repris de celui proposé par Thomas Chambrion l'an passé, lui-même librement adapté du chapitre 6 du livre
« Statistique » de M. Lejeune (Ed Springer).



Disponible légalement en pdf à la BU.

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Cadre de l'estimation

- Θ un sous ensemble de $\mathbb{R}^k$ (pour nous, souvent $k = 1$)
- $\mathcal{L}(\theta)$ une famille de loi de probabilité paramétrée par θ
- un certain θ_0 (inconnu) détermine la loi $\mathcal{L}(\theta_0)$ d'un phénomène physique.
- données : un n -échantillon $X_1, X_2, \dots, X_n$ de loi mère $\mathcal{L}(\theta_0)$

Problème d'estimation paramétrique

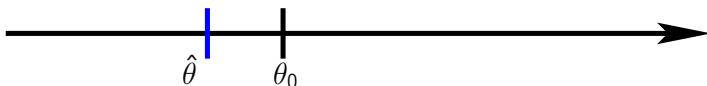
À partir d'une réalisation de $X_1, X_2, \dots, X_n$, retrouver θ_0 .

Problème d'estimation paramétrique ponctuelle

Problème d'estimation paramétrique ponctuelle

À partir d'une réalisation de $X_1, X_2, \dots, X_n$, retrouver θ_0 .

Pour l'instant, nous cherchons **une seule valeur**,



Problème d'estimation paramétrique ponctuelle

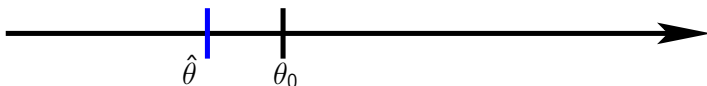
Problème d'estimation paramétrique ponctuelle

À partir d'une réalisation de $X_1, X_2, \dots, X_n$, retrouver θ_0 .

Problème d'estimation paramétrique par intervalle

À partir d'une réalisation de $X_1, X_2, \dots, X_n$, et pour une certaine confiance donnée α , trouver un intervalle I_α tel que $\theta_0 \in I$.

Pour l'instant, nous cherchons **une seule valeur**,



Problème d'estimation paramétrique ponctuelle

Problème d'estimation paramétrique ponctuelle

À partir d'une réalisation de $X_1, X_2, \dots, X_n$, retrouver θ_0 .

Problème d'estimation paramétrique par intervalle

À partir d'une réalisation de $X_1, X_2, \dots, X_n$, et pour une certaine confiance donnée α , trouver un intervalle I_α tel que $\theta_0 \in I$.

Pour l'instant, nous cherchons **une seule valeur**, sans parler d'intervalle de confiance (la partie en bleu clair, pour le CM 6).



Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Notion d'estimateur

Estimateur de θ

Suite $(T_n)_n$ de statistiques de n échantillons $X_1, \dots, X_n$.

Dans l'idéal :

$$T_n(X_1, \dots, X_n) = \theta_0$$

Pas possible de manière certaine en général.

- θ_0 est un paramètre fixé (un nombre, comme $1/2$).
- $T_n(X_1, \dots, X_n)$ est une variable aléatoire, pas constante en général.

Tout ce qu'on peut espérer, c'est que T_n soit « proche » de θ_0 en général.

Tout dépend de θ

L'écart $T_n - \theta$ est une variable aléatoire fonction de θ .

Qualité d'un estimateur : biais

Biais d'un estimateur

$$b_{\theta}(T_n) = \mathbb{E}_{\theta}(T_n) - \theta$$

Estimateur sans biais

T_n est sans biais si $b_{\theta}(T_n) = 0$ pour tout θ dans Θ .

Qualité d'un estimateur : biais

Biais d'un estimateur

$$b_{\theta}(T_n) = \mathbb{E}_{\theta}(T_n) - \theta$$

Estimateur sans biais

T_n est sans biais si $b_{\theta}(T_n) = 0$ pour tout θ dans Θ .

Exemple

Le couple (moyenne empirique, variance de l'échantillon) est un estimateur sans biais pour $\mathcal{N}(\mu, \sigma^2)$ (ici $\Theta = \mathbb{R}^2$, et $\theta = (\mu, \sigma^2)$) (cf. cours précédent).

Qualité d'un estimateur : variance

Erreur quadratique moyenne

$$eqm_{\theta}(T_n) = \mathbb{E}_{\theta}[(T_n - \theta)^2] = b_{\theta}(T_n)^2 + V_{\theta}(T_n)$$

Qualité d'un estimateur : variance

Erreur quadratique moyenne

$$eqm_{\theta}(T_n) = \mathbb{E}_{\theta}[(T_n - \theta)^2] = b_{\theta}(T_n)^2 + V_{\theta}(T_n)$$

Preuve :

$$\begin{aligned} eqm_{\theta}(T_n) &= \mathbb{E}_{\theta}[(T_n - \theta)^2] \\ &= \mathbb{E}_{\theta}(T_n^2 - 2\theta T_n + \theta^2) \\ &= \mathbb{E}_{\theta}(T_n)^2 - 2\theta \mathbb{E}_{\theta}(T_n) + \theta^2 + \mathbb{E}_{\theta}(T_n^2) - \mathbb{E}_{\theta}(T_n)^2 \\ &= b_{\theta}(T_n)^2 + V_{\theta}(T_n) \end{aligned}$$

Qualité d'un estimateur : variance

Erreur quadratique moyenne

$$eqm_{\theta}(T_n) = \mathbb{E}_{\theta}[(T_n - \theta)^2] = b_{\theta}(T_n)^2 + V_{\theta}(T_n)$$

Preuve :

$$\begin{aligned} eqm_{\theta}(T_n) &= \mathbb{E}_{\theta}[(T_n - \theta)^2] \\ &= \mathbb{E}_{\theta}(T_n^2 - 2\theta T_n + \theta^2) \\ &= \mathbb{E}_{\theta}(T_n)^2 - 2\theta \mathbb{E}_{\theta}(T_n) + \theta^2 + \mathbb{E}_{\theta}(T_n^2) - \mathbb{E}_{\theta}(T_n)^2 \\ &= b_{\theta}(T_n)^2 + V_{\theta}(T_n) \end{aligned}$$

Estimateur UMVUE

Rappel

$$eqm_{\theta}(T_n) = \mathbb{E}_{\theta}[(T_n - \theta)^2] = b_{\theta}(T_n)^2 + V_{\theta}(T_n)$$

Estimateur UMVUE

Rappel

$$eqm_{\theta}(T_n) = \mathbb{E}_{\theta}[(T_n - \theta)^2] = b_{\theta}(T_n)^2 + V_{\theta}(T_n)$$

Estimateur UMVUE

L'estimateur T_n^* est UMVUE pour θ (*uniformly minimum variance unbiased estimator*) si

- $b_{\theta}(T_n^*) = 0$ pour tout θ
- pour tout estimateur T_n sans biais pour θ ,

$$V_{\theta}(T_n^*) \leq V_{\theta}(T_n) \quad \forall \theta \in \Theta$$

Estimateur UMVUE

Rappel

$$eqm_{\theta}(T_n) = \mathbb{E}_{\theta}[(T_n - \theta)^2] = b_{\theta}(T_n)^2 + V_{\theta}(T_n)$$

Estimateur UMVUE

L'estimateur T_n^* est UMVUE pour θ (*uniformly minimum variance unbiased estimator*) si

- $b_{\theta}(T_n^*) = 0$ pour tout θ
- pour tout estimateur T_n sans biais pour θ ,

$$V_{\theta}(T_n^*) \leq V_{\theta}(T_n) \quad \forall \theta \in \Theta$$

Un UMVUE, c'est bien mais

- ça n'existe pas toujours ;
- quand ça existe, c'est souvent dur à trouver.

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Classe exponentielle (paramètre en dimension 1)

Définition (cas des lois à densité)

La famille $\mathcal{L}(\theta)$ (avec $\theta \in \mathbb{R}$) est une famille exponentielle si pour tout θ , la densité de la loi $\mathcal{L}(\theta)$ est

$$f(x, \theta) = a(\theta)b(x) \exp(c(\theta)d(x))$$

Définition (cas des lois discrètes)

La famille $\mathcal{L}(\theta)$ (avec $\theta \in \mathbb{R}$) est une famille exponentielle si pour tout θ , si X suit la loi $\mathcal{L}(\theta)$,

$$\mathbb{P}(X = x) = f(x, \theta) = a(\theta)b(x) \exp(c(\theta)d(x))$$

Exemple : loi binomiale $\mathcal{B}(n, p)$ (n connu)

$$f(k, p) = C_n^k p^k (1-p)^{n-k} = (1-p)^n \frac{n!}{k!(n-k)!} \exp \left(\ln \left(\frac{p}{1-p} \right) k \right)$$

- $\theta = p, k = x$
- $a(\theta) = (1-p)^n$
- $b(k) = \frac{n!}{k!(n-k)!}$
- $c(\theta) = \ln \left(\frac{p}{1-p} \right)$
- $d(k) = k$

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Estimateurs UMVUE pour les classes exponentielles

Théorème (Pitman–Koopman–Darmois)

Si la famille de la loi mère appartient à la classe exponentielle avec paramètre de dimension 1 ($\theta \in \mathbb{R}$) et s'il existe une statistique T_n fonction de $\sum_{i=1}^n d(X_i)$ qui soit sans biais pour θ , alors elle est unique et elle est UMVUE pour θ .

Estimateurs UMVUE pour les classes exponentielles

Théorème (Pitman–Koopman–Darmois)

Si la famille de la loi mère appartient à la classe exponentielle avec paramètre de dimension 1 ($\theta \in \mathbb{R}$) et s'il existe une statistique T_n fonction de $\sum_{i=1}^n d(X_i)$ qui soit sans biais pour θ , alors elle est unique et elle est UMVUE pour θ .

En clair : si on trouve T_n qui ne dépend que de $\sum_{i=1}^n d(X_i)$ et sans biais pour θ , c'est **un** bon estimateur.

Exemple

On veut évaluer le paramètre p d'une loi de Bernoulli.

- la loi de Bernoulli est dans la classe exponentielle
- $d(x) = x$

Exemple

On veut évaluer le paramètre p d'une loi de Bernoulli.

- la loi de Bernoulli est dans la classe exponentielle
- $d(x) = x$

Si on arrive à trouver une statistique

- qui ne dépend que de
$$d(X_1) + d(X_2) + \dots + d(X_n) = X_1 + X_2 + \dots + X_n$$
- et d'espérance p

ce sera l'UMVUE cherché.

Exemple

On veut évaluer le paramètre p d'une loi de Bernoulli.

- la loi de Bernoulli est dans la classe exponentielle
- $d(x) = x$

Si on arrive à trouver une statistique

- qui ne dépend que de
$$d(X_1) + d(X_2) + \dots + d(X_n) = X_1 + X_2 + \dots + X_n$$
- et d'espérance p

ce sera l'UMVUE cherché.

$$T_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Retour sur l'estimation du paramètre p d'une loi de Bernoulli

- $X_1, \dots, X_n$ n échantillon de loi $\mathcal{B}(p)$
- Notation $S_n = X_1 + X_2 + \dots + X_n$
- On veut estimer p
- On propose l'estimateur $T_n = \frac{S_n + 1}{n + 2}$.

Retour sur l'estimation du paramètre p d'une loi de Bernoulli

- $X_1, \dots, X_n$ n échantillon de loi $\mathcal{B}(p)$
- Notation $S_n = X_1 + X_2 + \dots + X_n$
- On veut estimer p
- On propose l'estimateur $T_n = \frac{S_n + 1}{n + 2}$.

$$E(T_n) = \frac{np + 1}{n + 2} \quad V(T_n) = \frac{V(S_n)}{(n + 2)^2} = \frac{np(1 - p)}{(n + 2)^2}$$

Retour sur l'estimation du paramètre p d'une loi de Bernoulli

- $X_1, \dots, X_n$ n échantillon de loi $\mathcal{B}(p)$
- Notation $S_n = X_1 + X_2 + \dots + X_n$
- On veut estimer p
- On propose l'estimateur $T_n = \frac{S_n + 1}{n + 2}$.

$$E(T_n) = \frac{np + 1}{n + 2} \quad V(T_n) = \frac{V(S_n)}{(n + 2)^2} = \frac{np(1 - p)}{(n + 2)^2}$$

$$eqm_p(T_n) = \left(\frac{np + 1}{n + 2} - p \right)^2 + \frac{np(1 - p)}{(n + 2)^2} = \left(\frac{1 - 2p}{n + 2} \right)^2 + \frac{np(1 - p)}{(n + 2)^2}$$

Retour sur l'estimation du paramètre p d'une loi de Bernoulli

- $X_1, \dots, X_n$ n échantillon de loi $\mathcal{B}(p)$
- Notation $S_n = X_1 + X_2 + \dots + X_n$
- On veut estimer p
- On propose l'estimateur $T_n = \frac{S_n + 1}{n + 2}$.

$$E(T_n) = \frac{np + 1}{n + 2} \quad V(T_n) = \frac{V(S_n)}{(n + 2)^2} = \frac{np(1 - p)}{(n + 2)^2}$$

$$eqm_p(T_n) = \left(\frac{np + 1}{n + 2} - p \right)^2 + \frac{np(1 - p)}{(n + 2)^2} = \left(\frac{1 - 2p}{n + 2} \right)^2 + \frac{np(1 - p)}{(n + 2)^2}$$

Comparaison avec l'estimateur UMVUE S_n/n

Pour p suffisamment proche de $\frac{1}{2}$, $eqm_p(T_n) < eqm_p\left(\frac{S_n}{n}\right)$.

Retour sur l'estimation du paramètre p d'une loi de Bernoulli

- $X_1, \dots, X_n$ n échantillon de loi $\mathcal{B}(p)$
- Notation $S_n = X_1 + X_2 + \dots + X_n$
- On veut estimer p
- On propose l'estimateur $T_n = \frac{S_n + 1}{n + 2}$.

$$E(T_n) = \frac{np + 1}{n + 2} \quad V(T_n) = \frac{V(S_n)}{(n + 2)^2} = \frac{np(1 - p)}{(n + 2)^2}$$

$$eqm_p(T_n) = \left(\frac{np + 1}{n + 2} - p \right)^2 + \frac{np(1 - p)}{(n + 2)^2} = \left(\frac{1 - 2p}{n + 2} \right)^2 + \frac{np(1 - p)}{(n + 2)^2}$$

Comparaison avec l'estimateur UMVUE S_n/n

Pour p suffisamment proche de $\frac{1}{2}$, $eqm_p(T_n) < eqm_p\left(\frac{S_n}{n}\right)$.

...mais l'estimateur a un biais !

Bilan provisoire

- Si la loi est dans la classe exponentielle **et** si on sait faire les calculs, le problème d'estimation est réglé théoriquement.

Bilan provisoire

- Si la loi est dans la classe exponentielle **et** si on sait faire les calculs, le problème d'estimation est réglé théoriquement.
- L'estimateur UMVUE est **un** bon estimateur.

Bilan provisoire

- Si la loi est dans la classe exponentielle **et** si on sait faire les calculs, le problème d'estimation est réglé théoriquement.
- L'estimateur UMVUE est **un** bon estimateur.
- Ennui : ça donne juste une valeur (c'est ce qu'on demandait !)

Bilan provisoire

- Si la loi est dans la classe exponentielle **et** si on sait faire les calculs, le problème d'estimation est réglé théoriquement.
- L'estimateur UMVUE est **un** bon estimateur.
- Ennui : ça donne juste une valeur (c'est ce qu'on demandait !)
- En pratique, on fera souvent autrement.

Bilan provisoire

- Si la loi est dans la classe exponentielle **et** si on sait faire les calculs, le problème d'estimation est réglé théoriquement.
- L'estimateur UMVUE est **un** bon estimateur.
- Ennui : ça donne juste une valeur (c'est ce qu'on demandait !)
- En pratique, on fera souvent autrement.
- Dans la suite, on présente brièvement trois méthodes assez largement utilisées en pratique.

Bilan provisoire

- Si la loi est dans la classe exponentielle **et** si on sait faire les calculs, le problème d'estimation est réglé théoriquement.
- L'estimateur UMVUE est **un** bon estimateur.
- Ennui : ça donne juste une valeur (c'est ce qu'on demandait !)
- En pratique, on fera souvent autrement.
- Dans la suite, on présente brièvement trois méthodes assez largement utilisées en pratique.
- Les deux dernières permettront d'obtenir un intervalle.

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Principe

Moments empiriques

Moment empirique d'ordre r ($r \in \mathbb{N}$) :

$$M_r = \frac{X_1^r + X_2^r + \cdots + X_n^r}{n}$$

Principe

Moments empiriques

Moment empirique d'ordre r ($r \in \mathbb{N}$) :

$$M_r = \frac{X_1^r + X_2^r + \cdots + X_n^r}{n}$$

Convergence des moments empiriques

Si la loi mère admet des moments finis $\mathbb{E}(|X|^s)$ jusqu'à l'ordre r , alors les moments empiriques jusqu'à l'ordre r convergent presque sûrement vers les moments correspondants de la loi mère.

Preuve : loi forte des grands nombres.

Principe

Moments empiriques

Moment empirique d'ordre r ($r \in \mathbb{N}$) :

$$M_r = \frac{X_1^r + X_2^r + \cdots + X_n^r}{n}$$

Convergence des moments empiriques

Si la loi mère admet des moments finis $\mathbb{E}(|X|^s)$ jusqu'à l'ordre r , alors les moments empiriques jusqu'à l'ordre r convergent presque sûrement vers les moments correspondants de la loi mère.

Preuve : loi forte des grands nombres.

Principe de la méthode des moments

Exprimer θ en fonction des moments de la loi mère, et approcher ces moments par les moments empiriques.

Inversion

Pour utiliser la méthode des moments :

- On calcule les premiers moments de la loi mère (en général, le premier ou les deux premiers) en fonction de θ .

$$\mathbb{E}(X) = \mu(\theta), \quad \mathbb{E}(X^2) = \mu_2(\theta), \quad \dots$$

Autant de moments que la dimension de θ .

- On inverse le système précédent (on résout en θ) :

$$\theta = F(\mathbb{E}(X), \mathbb{E}(X^2), \dots)$$

- Si F est continue, F sur les moments empiriques converge vers θ .

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Exemple : loi normale

$\mathcal{N}(\mu, \sigma^2)$. Ici $\theta = (\mu, \sigma^2)$ de dimension deux. Donc on utilisera les deux premiers moments de la loi.

- $\mathbb{E}(X) = \mu, \mathbb{E}(X^2) = V(X) + \mathbb{E}(X)^2 = \sigma^2 + \mu^2$

Exemple : loi normale

$\mathcal{N}(\mu, \sigma^2)$. Ici $\theta = (\mu, \sigma^2)$ de dimension deux. Donc on utilisera les deux premiers moments de la loi.

- $\mathbb{E}(X) = \mu, \mathbb{E}(X^2) = V(X) + \mathbb{E}(X)^2 = \sigma^2 + \mu^2$
- On inverse

$$\begin{cases} \mathbb{E}(X) &= \mu \\ \mathbb{E}(X^2) &= \sigma^2 + \mu^2 \end{cases} \Leftrightarrow \begin{cases} \mu &= \mathbb{E}(X) \\ \sigma^2 &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{cases}$$

Exemple : loi normale

$\mathcal{N}(\mu, \sigma^2)$. Ici $\theta = (\mu, \sigma^2)$ de dimension deux. Donc on utilisera les deux premiers moments de la loi.

- $\mathbb{E}(X) = \mu, \mathbb{E}(X^2) = V(X) + \mathbb{E}(X)^2 = \sigma^2 + \mu^2$
- On inverse

$$\begin{cases} \mathbb{E}(X) &= \mu \\ \mathbb{E}(X^2) &= \sigma^2 + \mu^2 \end{cases} \Leftrightarrow \begin{cases} \mu &= \mathbb{E}(X) \\ \sigma^2 &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{cases}$$

- Conclusion : on peut estimer l'espérance par la moyenne empirique $\bar{X}$ et la variance par la variance empirique $\tilde{S}^2$.

Exemple : loi normale

$\mathcal{N}(\mu, \sigma^2)$. Ici $\theta = (\mu, \sigma^2)$ de dimension deux. Donc on utilisera les deux premiers moments de la loi.

- $\mathbb{E}(X) = \mu, \mathbb{E}(X^2) = V(X) + \mathbb{E}(X)^2 = \sigma^2 + \mu^2$
- On inverse

$$\begin{cases} \mathbb{E}(X) &= \mu \\ \mathbb{E}(X^2) &= \sigma^2 + \mu^2 \end{cases} \Leftrightarrow \begin{cases} \mu &= \mathbb{E}(X) \\ \sigma^2 &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{cases}$$

- Conclusion : on peut estimer l'espérance par la moyenne empirique $\bar{X}$ et la variance par la variance empirique $\tilde{S}^2$.

Exemple : loi normale

$\mathcal{N}(\mu, \sigma^2)$. Ici $\theta = (\mu, \sigma^2)$ de dimension deux. Donc on utilisera les deux premiers moments de la loi.

- $\mathbb{E}(X) = \mu, \mathbb{E}(X^2) = V(X) + \mathbb{E}(X)^2 = \sigma^2 + \mu^2$
- On inverse

$$\begin{cases} \mathbb{E}(X) &= \mu \\ \mathbb{E}(X^2) &= \sigma^2 + \mu^2 \end{cases} \Leftrightarrow \begin{cases} \mu &= \mathbb{E}(X) \\ \sigma^2 &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{cases}$$

- Conclusion : on peut estimer l'espérance par la moyenne empirique $\bar{X}$ et la variance par la variance empirique $\tilde{S}^2$.

Pour cet exemple

- ça marche (convergence)
- mais $\mathbb{E}(\tilde{S}^2) = \frac{n}{n-1}\sigma^2 \neq \sigma^2$. $\tilde{S}$ est un estimateur biaisé de σ^2 .

Qualités et défauts de la méthode

Points positifs

- Facile à utiliser
- Quand on peut l'utiliser, ça marche (convergence asymptotique)
- Méthode très utilisée quand on ne sait pas quoi faire d'autre

Qualités et défauts de la méthode

Points positifs

- Facile à utiliser
- Quand on peut l'utiliser, ça marche (convergence asymptotique)
- Méthode très utilisée quand on ne sait pas quoi faire d'autre

Points négatifs

- Pas toujours utilisable
- Aucun sens pour les petits échantillons
- Estimateurs en général peu performants
- Très difficile de passer ensuite à des intervalles de confiance

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Principe

Fonction de vraisemblance

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta)f(x_2, \theta) \cdots f(x_n, \theta)$$

Principe

Fonction de vraisemblance

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta) f(x_2, \theta) \cdots f(x_n, \theta)$$

Idée

Considérer les x_i donnés par les mesures et $L(x_1, x_2, \dots, x_n, \theta)$ comme une fonction de θ . On cherche le θ (noté θ^{MV}) qui maximise L .

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Exemple : loi normale $\mathcal{N}(\mu, \sigma^2)$

Ici, $\theta = (\mu, \sigma^2)$.

$$\begin{aligned} L(x_1, x_2, \dots, x_n, \mu, \sigma^2) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \dots \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}} \\ &= \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \exp\left(-\frac{(x_1-\mu)^2 + \dots + (x_n-\mu)^2}{2\sigma^2}\right) \end{aligned}$$

Maximiser L revient à maximiser $2 \ln L$, donc on va maximiser

$$F(\mu, \sigma) = -2n \ln(\sigma) - \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

Exemple : loi normale $\mathcal{N}(\mu, \sigma^2)$

On va maximiser

$$F(\mu, \sigma) = -2n \ln(\sigma) - \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\frac{\partial F}{\partial \mu} = \frac{2}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

Exemple : loi normale $\mathcal{N}(\mu, \sigma^2)$

On va maximiser

$$F(\mu, \sigma) = -2n \ln(\sigma) - \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\frac{\partial F}{\partial \mu} = \frac{2}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\frac{\partial F}{\partial \sigma} = -\frac{2n}{\sigma} + 2\frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2 = \tilde{S}^2 (\neq S^2)$$

Exemple : loi normale $\mathcal{N}(\mu, \sigma^2)$

On va maximiser

$$F(\mu, \sigma) = -2n \ln(\sigma) - \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\frac{\partial F}{\partial \mu} = \frac{2}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\frac{\partial F}{\partial \sigma} = -\frac{2n}{\sigma} + 2 \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2 = \tilde{S}^2 (\neq S^2)$$

Attention

L'estimateur du maximum de vraisemblance peut avoir un biais !

Convergence

Information de Fisher

$$I(\theta) = \mathbb{E}_{\theta} \left[\left(\frac{\partial}{\partial \theta} \ln f(X, \theta) \right)^2 \right]$$

Convergence

Information de Fisher

$$I(\theta) = \mathbb{E}_{\theta} \left[\left(\frac{\partial}{\partial \theta} \ln f(X, \theta) \right)^2 \right]$$

En pratique : $I(\theta)$ donné dans les tables (Lejeune p. 96).

Convergence de l'estimateur du maximum de vraisemblance

Sous des conditions suffisantes de régularité, (toujours vérifiées pour nous en pratique), $\sqrt{n}(\hat{\theta}_n - \theta)$ converge en loi vers $\mathcal{N}\left(0, \frac{1}{I(\theta)}\right)$.

Qualités et défauts de la méthode

Points positifs

- facile à utiliser
- implémentée dans tous les logiciels sérieux
- fournira des intervalles de confiance/tests
- extrêmement efficace pour les grands échantillons (disons $n \geq 30$)

Qualités et défauts de la méthode

Points positifs

- facile à utiliser
- implémentée dans tous les logiciels sérieux
- fournira des intervalles de confiance/tests
- extrêmement efficace pour les grands échantillons (disons $n \geq 30$)

Points négatifs

- Possiblement biaisée
- Pas utilisable sur les petits échantillons.

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Principe de l'approche bayésienne

Idée fondamentale

- Considérer tout ce qu'on ne connaît pas comme des variables aléatoires.
- Exprimer la loi de ce qu'on l'observe en fonction des paramètres inconnus.
- Utiliser le théorème de Bayes (probabilités conditionnelles) pour « retourner » les probabilités conditionnelles et obtenir la loi des paramètres en fonction des observations.

Rappel : théorème de Bayes

Théorème de Bayes

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)}{\mathbb{P}(B)} \mathbb{P}(A)$$

Version continue pour les densités :

Théorème de Bayes (version continue)

$$f(\theta|x) = \frac{f(x, \theta)}{\int_{\theta} f(x, \theta) \pi(\theta) d\theta} \pi(\theta)$$

Mise en place

- on choisit une loi a priori sur θ (exemple : loi uniforme) de densité π
- on écrit la densité de la loi de $(x_1, x_2, \dots, x_n, \theta)$

$$f(x_1, \theta)f(x_2, \theta) \cdots f(x_n, \theta)\pi(\theta)$$

- on récupère, par le théorème de Bayes, la loi de θ connaissant $x_1, \dots, x_n$.
- Estimateur bayésien : espérance de θ pour cette loi a posteriori.
- on peut en déduire des intervalles de confiance

Qualités et défauts de la méthode

Points positifs

- utilisable dans le cas des petits échantillons
- très à la mode (intelligence artificielle)
- fournira des intervalles de probabilité pour θ

Qualités et défauts de la méthode

Points positifs

- utilisable dans le cas des petits échantillons
- très à la mode (intelligence artificielle)
- fournira des intervalles de probabilité pour θ

Points négatifs

- Dépend de la densité *a priori* choisie
- **Nécessite une expertise** pour ce choix.
- Calculs difficiles à la main, nécessitent des moyens informatiques.

Plan

- 1 Estimation
 - Problèmes d'estimation
 - Estimateurs
- 2 Classe exponentielle
 - Définition
 - Estimateurs pour une classe exponentielle
- 3 Méthode des moments
 - Principe
 - Exemple
- 4 Maximum de vraisemblance
 - Principe
 - Exemple
- 5 Approche bayésienne
 - Principe
 - Exemple

Exemple : lancer de dé

Question

On lance un dé 100 fois, on obtient 16 « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

Exemple : lancer de dé

Question

On lance un dé 100 fois, on obtient 16 « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Recherche du paramètre p d'une loi de Bernoulli $\mathcal{B}(p)$.

Exemple : lancer de dé

Question

On lance un dé 100 fois, on obtient 16 « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Recherche du paramètre p d'une loi de Bernoulli $\mathcal{B}(p)$.
- Loi de Bernoulli dans la classe exponentielle.

Exemple : lancer de dé

Question

On lance un dé 100 fois, on obtient 16 « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Recherche du paramètre p d'une loi de Bernoulli $\mathcal{B}(p)$.
- Loi de Bernoulli dans la classe exponentielle.
- On va utiliser les trois méthodes

Exemple : lancer de dé

Question

On lance un dé 100 fois, on obtient 16 « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Recherche du paramètre p d'une loi de Bernoulli $\mathcal{B}(p)$.
- Loi de Bernoulli dans la classe exponentielle.
- On va utiliser les trois méthodes
 - Méthode des moments

Exemple : lancer de dé

Question

On lance un dé 100 fois, on obtient 16 « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Recherche du paramètre p d'une loi de Bernoulli $\mathcal{B}(p)$.
- Loi de Bernoulli dans la classe exponentielle.
- On va utiliser les trois méthodes
 - Méthode des moments
 - Maximum de vraisemblance

Exemple : lancer de dé

Question

On lance un dé 100 fois, on obtient 16 « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Recherche du paramètre p d'une loi de Bernoulli $\mathcal{B}(p)$.
- Loi de Bernoulli dans la classe exponentielle.
- On va utiliser les trois méthodes
 - Méthode des moments
 - Maximum de vraisemblance
 - Approche bayésienne

Méthode des moments

Analyse du problème :

- X_i : $i^{\text{ème}}$ sortie (1 si « 6 », 0 sinon)
- $S_n = X_1 + \dots + X_n$
- S_n suit une loi binomiale $\mathcal{B}(n, p)$ (n connu)
- donnée : $S_{100} = 16$

$$E(S_n) = np \Rightarrow E\left(\frac{S_n}{n}\right) = p$$

Estimateur des moments : $T = \frac{S_n}{n}$ (UMVUE)

Méthode des moments

Analyse du problème :

- X_i : $i^{\text{ème}}$ sortie (1 si « 6 », 0 sinon)
- $S_n = X_1 + \dots + X_n$
- S_n suit une loi binomiale $\mathcal{B}(n, p)$ (n connu)
- donnée : $S_{100} = 16$

$$E(S_n) = np \Rightarrow E\left(\frac{S_n}{n}\right) = p$$

Estimateur des moments : $T = \frac{S_n}{n}$ (UMVUE)

Méthode des moments

$$p\hat{p} = \frac{16}{100} = \frac{4}{25}$$

Maximum de vraisemblance

Fonction de vraisemblance

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta) f(x_2, \theta) \cdots f(x_n, \theta)$$

Maximum de vraisemblance

Fonction de vraisemblance

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta)f(x_2, \theta) \cdots f(x_n, \theta)$$

$$F(p) = L(x_1, x_2, \dots, x_{100}, p) = p^{16}(1 - p)^{84}$$

Maximum de vraisemblance

Fonction de vraisemblance

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta)f(x_2, \theta) \cdots f(x_n, \theta)$$

$$F(p) = L(x_1, x_2, \dots, x_{100}, p) = p^{16}(1 - p)^{84}$$

Dérivée en p :

Maximum de vraisemblance

Fonction de vraisemblance

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta)f(x_2, \theta) \cdots f(x_n, \theta)$$

$$F(p) = L(x_1, x_2, \dots, x_{100}, p) = p^{16}(1 - p)^{84}$$

Dérivée en p :

$$\begin{aligned} F'(p) &= 16p^{15}(1 - p)^{84} - 84p^{16}(1 - p)^{83} \\ &= p^{15}(1 - p)^{83}(16(1 - p) - 84p) \\ &= p^{15}(1 - p)^{83}(16 - 100p) \end{aligned}$$

Maximum de vraisemblance

Fonction de vraisemblance

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta)f(x_2, \theta) \cdots f(x_n, \theta)$$

$$F(p) = L(x_1, x_2, \dots, x_{100}, p) = p^{16}(1 - p)^{84}$$

Dérivée en p :

$$\begin{aligned} F'(p) &= 16p^{15}(1 - p)^{84} - 84p^{16}(1 - p)^{83} \\ &= p^{15}(1 - p)^{83}(16(1 - p) - 84p) \\ &= p^{15}(1 - p)^{83}(16 - 100p) \end{aligned}$$

Maximum de F en $\hat{p} = \frac{16}{100} = \frac{4}{25}$.

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Théorème de Bayes

$$f_{P|S_{100}=16}(p) = \frac{C_{100}^{16} p^{16} (1-p)^{84} \cdot \mathbf{1}}{\int_{q=0}^1 C_{100}^{16} q^{16} (1-q)^{84} \cdot \mathbf{1} dq}$$

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Théorème de Bayes

$$f_{P|S_{100}=16}(p) = \frac{C_{100}^{16} p^{16} (1-p)^{84} \cdot 1}{\int_{q=0}^1 C_{100}^{16} q^{16} (1-q)^{84} \cdot 1 dq}$$

$$f_{P|S_{100}=16}(p) = 135931923533728279650 p^{16} (1-p)^{84}$$

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Théorème de Bayes

$$f_{P|S_{100}=16}(p) = \frac{C_{100}^{16} p^{16} (1-p)^{84} \cdot 1}{\int_{q=0}^1 C_{100}^{16} q^{16} (1-q)^{84} \cdot 1 dq}$$

$$f_{P|S_{100}=16}(p) = 135931923533728279650 p^{16} (1-p)^{84}$$

$$\mathbb{E}(P|S_{100} = 16) = \int_{p=0}^1 p \cdot f_{P|S_{100}=16}(p) dp = \frac{1}{6}$$

Estimation de p avec un seul tirage

Mise en garde : les miracles n'existent pas

Statistiques avec peu de données = pas grand chose à espérer.

Estimation de p avec un seul tirage

Mise en garde : les miracles n'existent pas

Statistiques avec peu de données = pas grand chose à espérer.

Question

On lance une fois le dé, on obtient « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

Estimation de p avec un seul tirage

Mise en garde : les miracles n'existent pas

Statistiques avec peu de données = pas grand chose à espérer.

Question

On lance une fois le dé, on obtient « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Méthode des moments : **exclue** ($n < 30$)

Estimation de p avec un seul tirage

Mise en garde : les miracles n'existent pas

Statistiques avec peu de données = pas grand chose à espérer.

Question

On lance une fois le dé, on obtient « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Méthode des moments : **exclue** ($n < 30$)
- Maximum de vraisemblance : **exclue** ($n < 30$)

Estimation de p avec un seul tirage

Mise en garde : les miracles n'existent pas

Statistiques avec peu de données = pas grand chose à espérer.

Question

On lance une fois le dé, on obtient « 6 ». Estimer la probabilité de sortie d'un « 6 » sur ce dé.

- Méthode des moments : **exclue** ($n < 30$)
- Maximum de vraisemblance : **exclue** ($n < 30$)
- Seule approche possible pour nous : approche bayésienne

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Théorème de Bayes

$$f_{P|S_1=1}(p) = \frac{C_1^1 p^1 (1-p)^0 \cdot \mathbf{1}}{\int_{q=0}^1 C_1^1 q^1 (1-q)^0 \cdot \mathbf{1} dq}$$

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous **choisissons** $\mathcal{U}([0, 1])$.

Théorème de Bayes

$$f_{P|S_1=1}(p) = \frac{C_1^1 p^1 (1-p)^0 \cdot \mathbf{1}}{\int_{q=0}^1 C_1^1 q^1 (1-q)^0 \cdot \mathbf{1} dq}$$

$$f_{P|S_1=1}(p) = 2p$$

Approche bayésienne

- P probabilité de sortie d'un « 6 » vue comme une variable aléatoire
- loi de probabilité a priori sur P . Nous choisissons $\mathcal{U}([0, 1])$.

Théorème de Bayes

$$f_{P|S_1=1}(p) = \frac{C_1^1 p^1 (1-p)^0 \cdot 1}{\int_{q=0}^1 C_1^1 q^1 (1-q)^0 \cdot 1 dq}$$

$$f_{P|S_1=1}(p) = 2p$$

$$\mathbb{E}(P|S_1 = 1) = \int_{p=0}^1 p \cdot f_{P|S_1=1}(p) dp = \frac{2}{3}$$