Regression avancée
Chapitre 2 : la famille exponentielle et le maximum de vraisemblance

Agathe Guilloux
Professeure au LaMME - Université d'Évry - Paris Saclay

Quelques exemples

Exemples de lois continues

- Loi normale (à densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$)
  $\mathcal{N}(\mu, 1)$ avec $\mu \in \mathbb{R}$

$$\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(y-\mu)^2) \text{ sur } \mathbb{R}$$

- Loi exponentielle (à densité par rapport à la mesure de Lebesgue sur $\mathbb{R}_+$)
  $\mathcal{E}(\lambda)$ avec $\lambda > 0$

$$\lambda \exp(-\lambda y) \text{ sur } \mathbb{R}_+$$

- Loi log-normale (à densité par rapport à la mesure de Lebesgue sur $\mathbb{R}_+$)
  $\log \mathcal{N}(\mu, 1)$ avec $\mu \in \mathbb{R}$
  Rappel $Y \sim \log \mathcal{N}(\mu, 1) \Leftrightarrow \log(Y) \sim \mathcal{N}(\mu, 1)$

$$\frac{1}{\sqrt{2\pi} y} \exp(-\frac{1}{2}(\log(y)-\mu)^2) \text{ sur } \mathbb{R}_+^\star$$

Exemples de lois discrètes

- Loi de Bernoulli (à densité par rapport à la mesure de comptage sur $\{0,1\}$) $\beta(p)$ avec $p \in]0,1[$

$$p^y(1-p)^{1-y} \text{ pour } y \in \{0,1\}$$

- Loi de Poisson (à densité par rapport à la mesure de comptage sur $\mathbb{N}$) $\mathcal{P}(\lambda)$ avec $\lambda > 0$

$$\exp(-\lambda)\frac{\lambda^y}{y!} \text{ pour } y \in \mathbb{N}$$

Exemples de lois continues

- Loi normale $\mathcal{N}(\mu, 1)$

$$\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(y-\mu)^2) = \exp\left(y\mu - \frac{1}{2}\mu^2 - \frac{1}{2}y^2 - \log(y) - \log(\sqrt{2\pi})\right)$$

- Loi exponentielle $\mathcal{E}(\lambda)$

$$\lambda \exp(-\lambda y) = \exp\left(- \lambda y + \log(\lambda)\right)$$

- Loi log-normale $\log \mathcal{N}(\mu, 1)$

$$\frac{1}{\sqrt{2\pi} y} \exp(-\frac{1}{2}(\log(y)-\mu)^2)$$
$$= \exp\left(\log(y)\mu - \frac{1}{2}\mu^2 - \frac{1}{2}(\log y)^2 - \log(y) - \log(\sqrt{2\pi})\right)$$

Exemples de lois discrètes

- Loi de Bernoulli $\beta(p)$

$$p^y(1-p)^{1-y} = \exp\left(y\log\left(\frac{p}{1-p}\right) + \log(1-p)\right)$$

- Loi de Poisson $\mathcal{P}(\lambda)$

$$\exp(-\lambda)\frac{\lambda^y}{y!} = \exp\left(y\log(\lambda) - \lambda - \log(y!)\right)$$

La famille exponentielle naturelle : définitions

Qu'ont en commun ces lois

On peut récrire leurs densités sous la forme

$$\exp(\eta T(y) - A(\eta) - c(y))$$

ou presque ....

Prélinimaire d'analyse

Proposition

Soient $\mathcal{Y}$ un espace mesurable, ν une mesure positive sur $\mathcal{Y}$, T une fonction mesurable sur $\mathcal{Y}$ (supposée non ν-p.p. constante) et c une fonction mesurable. On définit $\mathcal{H} = \left\{\eta \in \mathbb{R} : \int_{\mathcal{Y}} \exp\left(\eta T(y) - c(y)\right) d\nu(y) < \infty\right\}$ et l'on suppose que $\mathring{\mathcal{H}} \neq \emptyset$ alors

la fonction $\eta \mapsto \int_{\mathcal{Y}} \exp\left(\eta T(y) - c(y)\right) d\nu(y)$ est infiniment différentiable sur $\mathring{\mathcal{H}}$ et de dérivée k-ième

$$\int_{\mathcal{Y}} T^k(y) \exp\left(\eta T(y) - c(y)\right) d\nu(y).$$

Famille exponentielle naturelle

Famille exponentielle naturelle

On se donne une mesure positive ν sur $\mathcal{Y}$, une fonction T (non ν p.p. constante) telles que $\mathring{\mathcal{H}} \neq \emptyset$ alors on définit la famille exponentielle naturelle comme la famille des lois de densités g_η pour $\eta \in \mathring{\mathcal{H}}$ par rapport à ν de la forme

$$g_\eta(y) = \exp\Big(\eta T(y) - A(\eta) - c(y)\Big).$$

Propriété

On a forcément

$$A(\eta) = \log\bigg(\int_{\mathcal{Y}} \exp\Big(\eta T(y) - c(y)\Big)\, d\nu(y)\bigg).$$

Exemples

- Loi normale $\mathcal{N}(\mu, 1)$

$$\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(y-\mu)^2) = \exp\left(y\mu - \frac{1}{2}\mu^2 - \frac{1}{2}y^2 - \log(y) - \log(\sqrt{2\pi})\right)$$

- Loi log-normale $\log \mathcal{N}(\mu, 1)$

$$\frac{1}{\sqrt{2\pi}y} \exp(-\frac{1}{2}(\log(y)-\mu)^2)$$
$$= \exp\left(\log(y)\mu - \frac{1}{2}\mu^2 - \frac{1}{2}(\log y)^2 - \log(y) - \log(\sqrt{2\pi})\right)$$

Premières propriétés

Dans une famille exponentielle naturelle (sur $\mathcal{Y}$, pour une mesure positive ν et une fonction T), on a

$$\mathbb{E}_\eta\big(T(Y)\big) = A^{'}(\eta)$$
$$\mathbb{V}_\eta\big(T(Y)\big) = A^{''}(\eta)$$

pour tout $\eta \in \overset{\circ}{\mathcal{H}}$.

Exemples

- Loi normale $\mathcal{N}(\mu, 1)$

$$\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(y-\mu)^2) = \exp\left(y\mu - \frac{1}{2}\mu^2 - \frac{1}{2}y^2 - \log(y) - \log(\sqrt{2\pi})\right)$$

On a $T = \mathrm{id}$, $\eta = \mu$, $A(\eta) = \mu^2/2$ et

$$A^{'}(\eta) = \eta$$
$$A^{''}(\eta) = 1.$$

- Loi log-normale $\log \mathcal{N}(\mu, 1)$

$$\frac{1}{\sqrt{2\pi y}} \exp(-\frac{1}{2}(\log(y)-\mu)^2)$$
$$= \exp\left(\log(y)\mu - \frac{1}{2}\mu^2 - \frac{1}{2}(\log y)^2 - \log(y) - \log(\sqrt{2\pi})\right)$$

On a $T = \log$, $\eta = \mu$, $A(\eta) = \mu^2/2$ et

$$A^{'}(\eta) = \eta$$
$$A^{''}(\eta) = 1.$$

Estimation au maximum de vraisemblance

Echantillon et vraisemblance

On a un échantillon de v.a. i.i.d. $Y_1, \ldots, Y_n$ de densité $g_{\eta^\star}$ (dans une famille exponetielle naturelle) avec $\eta^\star \in \overset{\circ}{\mathcal{H}}$ et inconnu.

Log-vraisemblance et score

La log-vraisemblance (divisée par $1/n$) de l'échantillon en tout point $\eta \in \overset{\circ}{\mathcal{H}}$ est donnée par

$$L(Y_1, \ldots, Y_n; \eta) = \frac{1}{n} \sum_{i=1}^{n} \log \Big(g_\eta(Y_i) \Big) = \frac{1}{n} \sum_{i=1}^{n} T(Y_i) - A(\eta) - \frac{1}{n} \sum_{i=1}^{n} c(Y_i).$$

Propriété de la log-vraisemblance

La log-vraisemblance (divisée par $1/n$) $\eta \in \mathring{\mathcal{H}} \mapsto L(Y_1, \ldots, Y_n; \eta)$ est une fonction strictement concave sur $\mathring{\mathcal{H}}$. Donc s'il existe $\hat{\eta} \in \mathring{\mathcal{H}}$ tel que

$$L^{'}(Y_1, \ldots, Y_n; \eta) = 0 \Leftrightarrow \frac{1}{n} \sum_{i=1}^{n} T(Y_i) = A^{'}(\hat{\eta}),$$

c'est l'unique estimateur au maximum de vraisemblance de $\eta^{\star}$.

Définition

Famille exponentielle générale

On se donne une mesure positive ν sur $\mathcal{Y}$, une fonction T (non ν p.p. constante) et une fonction Q $\mathcal{C}^1$ strictement monotone. On définit alors la famille de densités

$$f_\theta(y) = \exp\Big(Q(\theta)\,T(y) - \alpha(\theta) - c(y)\Big).$$

Remarque : on retrouve la famille exponentielle naturelle en posant

$$\eta = Q(\theta) \text{ et } A(\eta) = \alpha\Big(Q^{-1}(\eta)\Big).$$

Exemples

- Loi exponentielle $\mathcal{E}(\lambda)$

$$\lambda \exp(-\lambda y) = \exp\big(-\lambda y + \log(\lambda)\big).$$

On pose $T = \text{id}$, $Q(\lambda) = -\lambda$, et $\alpha(\lambda) = -\log(\lambda)$.

- Loi de Bernoulli $\beta(p)$

$$p^y(1-p)^{1-y} = \exp\Big(y \log\big(\frac{p}{1-p}\big) + \log(1-p)\Big).$$

On pose $T = \text{id}$, $Q(p) = p/(1-p)$, et $\alpha(p) = -\log(1-p)$.

- Loi de Poisson $\mathcal{P}(\lambda)$

$$\exp(-\lambda)\frac{\lambda^y}{y!} = \exp\Big(y \log(\lambda) - \lambda - \log(y!)\Big)$$

On pose $T = \text{id}$, $Q(\lambda) = \log(\lambda)$, et $\alpha = \text{id}$.

Propriétés asymptotiques de l'EMV

Pour l'estimation de $A^{'}(\eta*)$, on choisit

$$\widehat{A'(\eta*)} = A^{'}(\hat{\eta}) = \frac{1}{n}\sum_{i=1}^{n} T(Y_i).$$

- la loi des grands nombres indique que $\widehat{A'(\eta*)} \xrightarrow{p.s.} \mathbb{E}(T(Y_i)) = A^{'}(\eta*)$.
- la théorème central limite implique que

$$\sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^{n} T(Y_i) - A^{'}(\eta*)}{\sqrt{A''(\eta*)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

- avec le lemme de Slutsky, on obtient

$$\sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^{n} T(Y_i) - A^{'}(\eta*)}{\sqrt{A''(\hat{\eta})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Si on veut des résultats pour η^*, il faut utiliser le théorème d'application continue et la delta-méthode.

Outline