ECN 6338 Cours 3, annexe

Version avec le modèle poissonnien

William McCausland

2025-09-18

Éléments de l'analyse maximum de vraisemblance

- Quantités pertinentes :
  - θ, un vecteur de paramètres inconnus,
  - $y = (y_1, \ldots, y_T)$, un vecteur aléatoire des variables observables,
  - y°, le vecteur observé.
- Fonctions pertinentes :
  - $f(y|\theta)$, la densité conditionnelle des données (modèle),
  - $\mathcal{L}(\theta; y) = f(y|\theta)$, la vraisemblance,
  - $\mathcal{L}(\theta; y^\circ) = f(y^\circ|\theta)$, la vraisemblance réalisée.

Le modèle poissonnien

- Supposez que les y_i sont iid Poisson avec moyenne $\theta > 0$.
- La fonction de masse de probabilité de y_i est

$$f(y_i|\theta) = e^{-\theta}\frac{\theta^{y_i}}{y_i!}.$$

- On observe le vecteur aléatoire $y = (y_1, \ldots, y_n)$; la fonction de masse de probabilité de y est

$$f(y|\theta) = \prod_{i=1}^{n} f(y_i|\theta) = \prod_{i=1}^{n} e^{-\theta}\frac{\theta^{y_i}}{y_i!} = \left[\prod_{i=1}^{n}\frac{1}{y_i!}\right] e^{-n\theta}\theta^{\sum_{i=1}^{n} y_i}.$$

- Pour simplifier un facteur qui importe peu,

$$c \equiv \left[\prod_{i=1}^{n}\frac{1}{y_i!}\right].$$

Deux intérpretations de la même expression

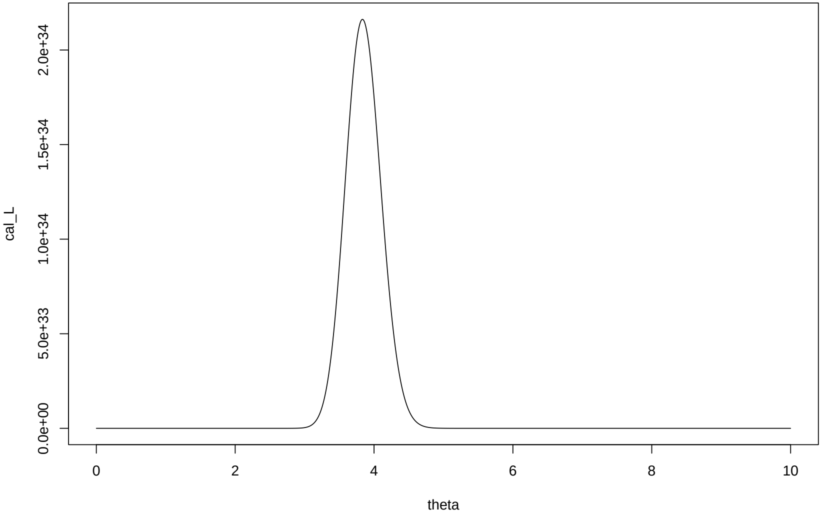
- L'expression :

$$f(y|\theta) = ce^{-n\theta}\theta^{\sum_{i=1}^{n} y_i} = \mathcal{L}(\theta; y).$$

- Deux intérpretations :
  - Fonction de masse de probabilité $f(y|\theta)$.
  - Fonction de vraisemblance $\mathcal{L}(\theta; y)$.
- $f(y|\theta)$ donne, pour θ fixe, les probabilités relatives des séquences possibles $(y_1, \ldots, y_n)$.
- $\mathcal{L}(\theta; y)$ donne, pour y fixe (notamment $y = y^\circ$) une note (ou évaluation) à chaque valeur θ pour la qualité de sa prévision des données observées.
- Soit $L(\theta; y) = \log \mathcal{L}(\theta; y)$, la log-vraisemblance.

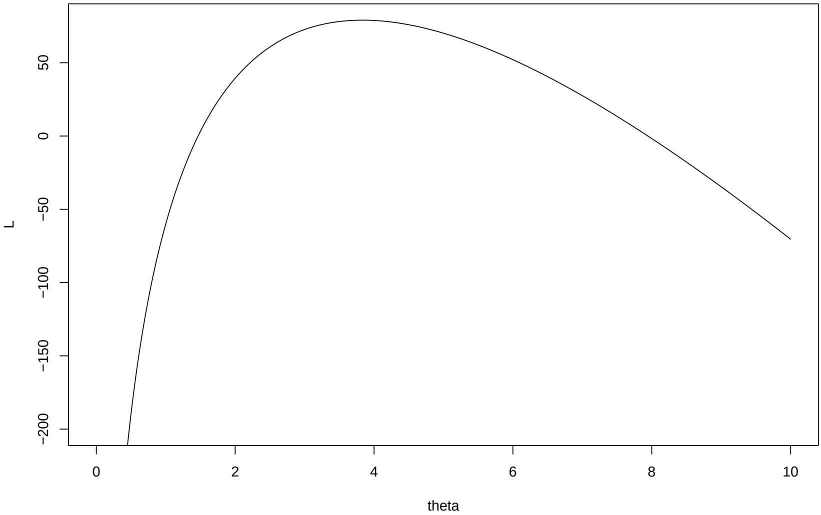
Vraisemblance poissonnienne pour $n = 60$, $\sum_{i=1}^{n} y_i = 230$

```
n = 60; somme_y = 230; theta = seq(0, 10, by=0.001)
cal_L = exp(-n*theta) * theta^somme_y
plot(theta, cal_L, type='l')
```



Log vraisemblance poissonnienne, $n = 60$, $\sum_{i=1}^{n} y_i = 230$

```r
L = -n*theta + somme_y*log(theta)
plot(theta, L, type='l', ylim=c(-200, max(L)))
```



Maximum de la vraisemblance poissonnienne

- Vraisemblance : $\mathcal{L}(\theta; y) = ce^{-n\theta}\theta^{\sum_{i=1}^n y_i}$.
- Log vraisemblance : $L(\theta; y) = \log c - n\theta + \left(\sum_{i=1}^n y_i\right) \log \theta$.
- Deux dérivées de la log vraisemblance :

$$\frac{\partial L(\theta; y)}{\partial \theta} = -n + \frac{\sum_{i=1}^n y_i}{\theta}$$

$$\frac{\partial^2 L(\theta; y)}{\partial \theta^2} = -\frac{\sum_{i=1}^n y_i}{\theta^2} < 0.$$

- La valeur $\hat{\theta}$ (souvent vue comme une variable aléatoire) qui maximise la vraisemblance et la log-vraisemblance est

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

- Pour $n = 60$ et $\sum_{i=1}^n y_i = 230$, $\hat{\theta} = \frac{23}{6} \approx 3.833$.

Information de Fisher et variance de la vraisemblance poissonnienne

- Un cas rare où les calculs analytiques sont faisables.
- La matrice d'information de Fisher : ($E[y_i] = \theta$, $\mathrm{Var}[y_i] = \theta$)

$$\mathcal{I}(\theta) = E_{y|\theta}\left[-\frac{\partial^2 L}{\partial\theta^2}\right] = E_{y|\theta}\left[\frac{\sum_{i=1}^n y_i}{\theta^2}\right] = \frac{n\theta}{\theta^2} = \frac{n}{\theta}.$$

- La variance de $\hat{\theta}$ (exacte, pas asymptotique) :

$$\mathrm{Var}[\hat{\theta}] = \mathrm{Var}\left[\frac{\sum_{i=1}^n y_i}{n}\right] = \frac{1}{n^2} n\mathrm{Var}[y_i] = \frac{\theta}{n}.$$

- La variance dépend de θ inconnu, mais elle est souvent peu sensible à θ : pour $n = 60$, $\mathrm{Var}[\hat{\theta}]$ est de $(0.2528)^2$ pour $\theta = 23/6 \approx 3.833$, $(0.2236)^2$ pour $\theta = 3$ et $(0.2739)^2$ pour $\theta = 4.5$.

La loi *a posteriori* pour l'exemple poissonnien

- Si $y_i \sim \mathrm{iid}\,\mathrm{Po}(\theta)$, $f(y|\theta) = ce^{-n\theta}\theta^{\sum_{i=1}^n y_i}$.
- Mettons qu'on choisit la loi *a priori* $\theta \sim \mathrm{Ga}(\alpha, \beta)$ sur $[0, \infty)$:

$$f(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}.$$

- La densité conjointe est

$$f(\theta, y) = f(\theta)f(y|\theta) = c\frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \theta^{\alpha+\sum_{i=1}^n y_i - 1} \cdot e^{-(\beta+n)\theta}.$$

- La loi *a posteriori* doit être $\theta \sim \mathrm{Ga}(\alpha + \sum_{i=1}^n y_i, \beta + n)$.
- La vraisemblance marginale est $f(\theta, y)/f(\theta|y)$:

$$f(y) = c\frac{\beta^\alpha}{(\beta+n)^{\alpha+\sum_{i=1}^n y_i}} \cdot \frac{\Gamma(\alpha + \sum_{i=1}^n y_i)}{\Gamma(\alpha)}.$$

Densités *a prior* (rouge) et *a posterior* (noir), exemple poissonnien

```r
n = 60; somme_y = 230; alpha=2; beta=0.4
x = seq(0, 10, by=0.02)
plot(x, dgamma(x, alpha+somme_y, beta+n), type='l')
lines(x, dgamma(x, alpha, beta), col='red')
```