

## Séance 4 : Estimateurs maximum de vraisemblance et méthode des moments.

### Exercices

#### Exercice 1

Soient  $X_1, \dots, X_n$  i.i.d. de loi de Poisson de moyenne  $\lambda$ . Déterminez le MLE pour le paramètre  $\lambda$ .

#### Solution :

Comme dans tous les exercices, on cherche la valeur  $\mathbf{T}(\mathbf{X})$  qui maximise le logarithme de la vraisemblance (ceci est bien sûr équivalent à maximiser la vraisemblance...mais à l'avantage d'avoir des dérivées plus aisées à calculer). Ainsi,

$$\begin{aligned} L_\lambda(\mathbf{X}) &= \frac{e^{-n\lambda}}{\prod_i X_i!} \lambda^{\sum_i X_i} \\ \log L_\lambda(\mathbf{X}) &= -n\lambda + \sum_i X_i \log \lambda - \log\left(\prod_i X_i!\right) \\ \partial_\lambda \log L_\lambda(\mathbf{X}) &= -n + \frac{\sum_i X_i}{\lambda} \\ \partial_\lambda \log L_\lambda(\mathbf{X}) = 0 &\Leftrightarrow n = \frac{\sum_i X_i}{\lambda} \\ &\Leftrightarrow \lambda = \frac{1}{n} \sum_{i=1}^n X_i \end{aligned}$$

Vérifions que cette valeur est bien un maximum : prenons la dérivée seconde.

$$\partial_\lambda^2 \log L_\lambda(\mathbf{X}) = 0 - \frac{\sum_i X_i}{\lambda^2},$$

qui est toujours négatif. L'estimateur maximum de vraisemblance est  $\hat{\lambda} = \mathbf{T}(\mathbf{X}) = \bar{X}$ .

#### Exercice 2

Soient  $X_1, \dots, X_n$  i.i.d. Bin(1,  $p$ ). Déterminez le MLE pour le paramètre  $p$ .

#### Solution :

Toujours les mêmes idées :

$$\begin{aligned} L_p(\mathbf{X}) &= p^{\sum_i X_i} (1-p)^{n-\sum_i X_i} \\ \log L_p(\mathbf{X}) &= \log p \sum_i X_i + (n - \sum_i X_i) \log(1-p) \\ \partial_p \log L_p(\mathbf{X}) &= \frac{\sum_i X_i}{p} - \frac{(n - \sum_i X_i)}{(1-p)} \\ \partial_p \log L_p(\mathbf{X}) = 0 &\Leftrightarrow (1-p) \sum_i X_i = p(n - \sum_i X_i) \\ &\Leftrightarrow \sum_i X_i - p \sum_i X_i = np - p \sum_i X_i \\ &\Leftrightarrow p = \frac{1}{n} \sum_i X_i = \bar{X} \end{aligned}$$

Ainsi, l'estimateur maximum de vraisemblance est  $\hat{p} = \bar{X}$ .

**Exercice 3**

Soient  $X_1, \dots, X_n$  i.i.d.  $\mathcal{N}(\mu, \sigma^2)$ . Déterminez le MLE pour le paramètre  $\underline{\theta} = (\mu, \sigma^2)^T$ . En déduire le MLE pour

- le paramètre  $\mu$ ,  $\sigma$  étant supposé connu,
- le paramètre  $\sigma^2$ ,  $\mu$  étant supposé connu.

**Solution :**

On trouve,

$$\begin{aligned} L_{\underline{\theta}}(\underline{\mathbf{X}}) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2} \sum_i (X_i - \mu)^2\right) \\ \log L_{\underline{\theta}}(\underline{\mathbf{X}}) &= C - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 \\ \nabla_{\underline{\theta}} \log L_{\underline{\theta}}(\underline{\mathbf{X}}) &= \begin{pmatrix} \frac{\partial}{\partial \mu} \log L_{\underline{\theta}}(\underline{\mathbf{X}}) \\ \frac{\partial}{\partial \sigma^2} \log L_{\underline{\theta}}(\underline{\mathbf{X}}) \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sigma^2} \sum_i (X_i - \mu) \\ -\frac{n}{\sigma^2} + \frac{1}{2\sigma^4} \sum_i (X_i - \mu)^2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{n}{2\sigma^4} (\bar{X} - \mu) \\ \frac{n}{2\sigma^4} (n^{-1} \sum_i (X_i - \mu)^2 - \sigma^2) \end{pmatrix} \\ \nabla_{\underline{\theta}} \log L_{\underline{\theta}}(\underline{\mathbf{X}}) = \mathbf{0} &\Leftrightarrow \begin{cases} \mu = \bar{X} \\ \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - \sigma^2 = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} \mu = \bar{X} \\ \sigma^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = s^2 \end{cases} \end{aligned}$$

Après vérification de maximalité (via la matrice Hessienne), on conclut que l'estimateur maximum de vraisemblance est

$$\hat{\underline{\theta}} = \begin{pmatrix} \bar{X} \\ s^2 \end{pmatrix}.$$

Le cas où l'un des deux paramètres est connu est laissé en exercice.

**Exercice 4**

Soit  $X_1, \dots, X_n$  un échantillon simple issu d'une population uniforme sur  $\{1, 2, \dots, N\}$ . Déterminez, à partir de cet échantillon, le MLE pour  $N$ .

**Solution :**

Le paramètre d'intérêt est  $N$ . On trouve

$$\begin{aligned} \mathbb{P}[X = x] &= \begin{cases} \frac{1}{N} & \text{si } x \in \{1, \dots, N\} \\ 0 & \text{sinon} \end{cases} \\ L_N(\underline{\mathbf{X}}) &= \prod_{i=1}^n \frac{1}{N} \mathbb{I}_{\{1, \dots, N\}}(X_i) \\ &= \frac{1}{N^n} \prod_{i=1}^n \mathbb{I}_{\{1, \dots, N\}}(X_i) \\ &= \frac{1}{N^n} \mathbb{I}_{\{1, \dots, N\}}(X_{(1)}) \mathbb{I}_{\{1, \dots, N\}}(X_{(n)}) \end{aligned}$$

Cette quantité est maximisée en  $N$  lorsqu'elle est non-nulle et que  $N^n$  est petit, c'est-à-dire lorsque  $X_{(n)} \leq N$  et  $N$  petit. On trouve donc

$$\hat{N} = X_{(n)} = X_{max}.$$

### Exercice 5

Soit  $X_1, \dots, X_n$  un échantillon aléatoire simple issu d'une population de densité

$$f_\theta(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}(x-\gamma)} & \text{si } x > \gamma \\ 0 & \text{sinon} \end{cases}$$

où  $\theta > 0$ . Déterminez une estimation des paramètres  $\theta$  et  $\gamma$

- par la méthode du MLE,
- par la méthode des moments.

### Solution :

$$\begin{aligned} L_{\theta,\gamma}(\mathbf{X}) &= \prod_{i=1}^n \frac{1}{\theta} \exp(-\theta^{-1}(X_i - \gamma)) \mathbb{I}_{[\gamma, +\infty[}(X_i) \\ &= \frac{1}{\theta^n} \exp\left(-\theta^{-1} \sum (X_i - \gamma)\right) \mathbb{I}_{[\gamma, +\infty[}(X_{(1)}) \end{aligned}$$

En se limitant à  $X_{(1)} > \gamma$ ,

$$\begin{aligned} \log L_{\theta,\gamma}(\mathbf{X}) &= -n \log \theta - \frac{1}{\theta} \sum_i (X_i - \gamma) \\ \partial_\theta \log L_{\theta,\gamma}(\mathbf{X}) &= -n \frac{1}{\theta} + \frac{1}{\theta^2} \sum_i (X_i - \gamma) \\ \partial_\theta \log L_{\theta,\gamma}(\mathbf{X}) = 0 &\Leftrightarrow -n\theta + \sum_i X_i - n\gamma = 0 \\ &\Leftrightarrow \theta = \bar{X} - \gamma \\ \partial_\gamma \log L_{\theta,\gamma}(\mathbf{X}) &= \frac{n}{\theta} \end{aligned}$$

Cette dernière quantité n'est jamais nulle. Souhaitant maximiser la vraisemblance, on remarque qu'à  $\theta$  fixé, la vraisemblance est une fonction croissante de  $\gamma$ . Quand  $\gamma$  prend sa valeur maximale, la vraisemblance sera maximale. Or,  $\gamma \leq X_{(1)}$ . On trouve alors

$$\hat{\gamma} = X_{(1)} = X_{min} \text{ et } \hat{\theta} = \bar{X} - X_{(1)}.$$

Nous souhaitons maintenant obtenir des estimateurs via la méthode des moments. Il faut pour cela calculer  $\mathbb{E}[X]$  et  $\mathbb{E}[X^2]$ . Ceci peut se faire par intégration classique ou de la manière suivante. Pour rappel,

$$\begin{aligned} Y \sim \text{Exp}(\lambda) &\Leftrightarrow f_\lambda(x) = \lambda e^{-\lambda x} \mathbb{I}_{[x>0]}, \\ \mathbb{E}[Y] &= \frac{1}{\lambda} \quad \text{et} \quad \text{Var}(Y) = \frac{1}{\lambda^2}. \end{aligned}$$

On remarque que  $X - \gamma \sim \text{Exp}(\theta^{-1})$ . Ainsi,

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}[X - \gamma] + \gamma = \theta + \gamma \\ \mathbb{E}[X^2] &= \text{Var}(X) + \mathbb{E}[X]^2 \\ &= \theta^2 + (\theta + \gamma)^2\end{aligned}$$

Il suffit alors de résoudre le système d'équations

$$\begin{cases} \frac{1}{n} \sum_{i=1}^n X_i &= \hat{\theta} + \hat{\gamma} \\ \frac{1}{n} \sum_{i=1}^n X_i^2 &= \hat{\theta}^2 + (\hat{\theta} + \hat{\gamma})^2 \end{cases}$$

Ceci livre

$$\hat{\theta} = s_{\mathbf{X}} = \sqrt{\left(\frac{1}{n} \sum_i X_i^2\right) - \bar{X}^2} \text{ et } \hat{\gamma} = \bar{X} - s_{\mathbf{X}}.$$

### Exercice 6

- Soient  $X_1, \dots, X_n$  i.i.d.  $U[0, \theta]$ . Déterminez  $\hat{\theta}$ , le MLE pour le paramètre  $\theta$ . Quelle est la loi de  $\hat{\theta}$ ?
- Même question si on prend  $X_1, \dots, X_n$  i.i.d.  $U[-\theta, \theta]$ . Quelle est la loi de  $\hat{\theta}$ ?
- Soient  $X_1, \dots, X_n$  i.i.d.  $U[a, b]$ . Déterminez  $\hat{a}, \hat{b}$ , d'abord par la méthode du maximum de vraisemblance, puis en ayant recours à la méthode des moments.

### Solution :

Le seul cas un peu intéressant est celui où  $X_1, \dots, X_n$  sont i.i.d  $U[-\theta, \theta]$ . Dans ce cas, on trouve

$$\begin{aligned}L_{\theta}(\mathbf{X}) &= \prod_{i=1}^n \frac{1}{2\theta} \mathbb{I}_{[-\theta, \theta]}(X_i) \\ &= \frac{1}{(2\theta)^n} \mathbb{I}(X_{(1)} \geq -\theta) \mathbb{I}(X_{(n)} \leq \theta)\end{aligned}$$

Cette quantité est à maximiser en  $\theta$ . On peut supposer les deux indicatrices satisfaites en prenant  $\theta$  suffisamment grand. Ce maximum sera atteint si  $\frac{1}{(2\theta)^n}$  est grand, i.e. si  $\theta$  est petit. Il faut donc choisir le plus petit  $\theta$  tel que  $X_{(1)} \geq -\theta$  et  $X_{(n)} \leq \theta$ , c'est-à-dire

$$\hat{\theta} = \max(X_{(n)}, -X_{(1)}).$$

La loi de  $\hat{\theta}_1 = X_{(n)}$  dans le cas  $U[0, \theta]$  peut être trouvée facilement. Celle-ci est donnée par

$$F^{\hat{\theta}_1}(x) = \begin{cases} 0 & \text{pour } x < 0 \\ (F^X(x))^n = \left(\frac{x}{\theta}\right)^n & \text{pour } 0 \leq x \leq \theta \\ 1 & \text{pour } x > \theta \end{cases}$$

Cherchons alors la loi de  $\hat{\theta}_2 = \max(X_{(n)}, -X_{(1)})$ . Dans le cas  $x \in ]0, \theta]$ ,

$$\begin{aligned}F^{\hat{\theta}_2}(x) &= \mathbb{P}[\hat{\theta}_2 \leq x] \\ &= \mathbb{P}[X_{(1)} \geq -x \cap X_{(n)} \leq x] \\ &= \mathbb{P}[-x \leq X_1 \leq x \cap \dots \cap -x \leq X_n \leq x] \\ &= (\mathbb{P}[-x \leq X \leq x])^n \\ &= \left(\frac{x}{\theta}\right)^n\end{aligned}$$

Les cas  $x \leq 0$  et  $x \geq \theta$  sont triviaux. On remarque que

$$\hat{\theta}_2 \stackrel{\mathcal{L}}{=} \hat{\theta}_1.$$

**Exercice 7**

Soit  $X_1, \dots, X_n$  un échantillon aléatoire simple issu d'une population de densité

$$f_\theta(x) = \begin{cases} \frac{\theta}{1-\theta} x^{\frac{2\theta-1}{1-\theta}} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

où  $1/2 < \theta < 1$ . Déterminez le MLE pour  $\theta$ .

**Solution :**

Immédiatement, en supposant  $X_1, \dots, X_n$  dans le support de  $f_\theta$ ,

$$\begin{aligned} L_\theta(\underline{\mathbf{X}}) &= \prod_{i=1}^n f_\theta(X_i) \\ &= \prod_{i=1}^n \frac{\theta}{1-\theta} X_i^{\frac{2\theta-1}{1-\theta}} \mathbb{I}_{0 < X < 1}(X_i) \\ &= \left(\frac{\theta}{1-\theta}\right)^n \left(\prod_{i=1}^n X_i\right)^{\frac{2\theta-1}{1-\theta}} \prod_{i=1}^n \mathbb{I}_{0 < X < 1}(X_i) \\ \log L_\theta(\underline{\mathbf{X}}) &= n \log \theta - n \log(1-\theta) + \frac{2\theta-1}{1-\theta} \left(\sum_{i=1}^n \log X_i\right) \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) &= \frac{n}{\theta} + \frac{n}{1-\theta} + \left(\sum_{i=1}^n \log X_i\right) \frac{2(1-\theta) + (2\theta-1)}{(1-\theta)^2} \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) = 0 &\Leftrightarrow \theta = \frac{1}{1 - \frac{1}{n} \sum_i \log(X_i)} \end{aligned}$$

**Exercice 8**

Les éléments d'une population possèdent un caractère  $X$  qui suit une loi de densité

$$f_\theta(x) = \frac{\theta+1}{2} (1-|x|)^\theta, \quad x \in (-1, 1),$$

où on suppose le paramètre  $\theta > -1$ . On en extrait un échantillon simple  $X_1, \dots, X_n$ . Déterminez l'estimateur du maximum de vraisemblance  $\hat{\theta}$  de  $\theta$ .

**Solution :**

Comme d'habitude, en supposant que les indicatrices sont vérifiées (le support de dépendant pas de  $\theta$ ),

$$\begin{aligned} L_\theta(\underline{\mathbf{X}}) &= \prod_{i=1}^n f_\theta(X_i) \\ &= \left(\frac{\theta+1}{2}\right)^n \left(\prod_{i=1}^n (1-|X_i|)\right)^\theta \mathbb{I}_{\{X \geq -1\}}(X_{(1)}) \mathbb{I}_{\{X \leq 1\}}(X_{(n)}) \\ \log L_\theta(\underline{\mathbf{X}}) &= n \log \left(\frac{\theta+1}{2}\right) + \theta \sum_{i=1}^n \log(1-|X_i|) \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) &= \frac{n}{\theta+1} + \sum_{i=1}^n \log(1-|X_i|) \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) = 0 &\Leftrightarrow \theta = -\frac{n}{\sum \ln(1-|X_i|)} - 1 \end{aligned}$$

**Exercice 9**

On considère  $n$  observations indépendantes  $(X_1, \dots, X_n)$  de loi  $pU[0, a] + (1-p)U[0, b]$ , c'est-à-dire que  $X_i$  suit la loi uniforme sur  $[0, a]$  avec probabilité  $p$  et la loi uniforme sur  $[0, b]$  avec probabilité  $1-p$ . On suppose que  $a < b$  avec  $a$  et  $b$  connus et fixés.

- Déterminez la fonction de répartition et la densité de  $X_1$ .
- Considérons  $N_a$  la variable aléatoire égale au nombre d'individus  $X_i$  compris entre 0 et  $a$ . Quelle est la loi de  $N_a$ ? En déduire son espérance et sa variance.
- Déterminez l'estimateur du maximum de vraisemblance  $\hat{p}$  du paramètre  $p$ .

**Solution :**

En notant  $U_1$  (resp.  $U_2$ ) l'événement " $X_1$  provient de la loi uniforme sur  $[0, a]$  (resp.  $[0, b]$ )", la fonction de répartition de  $X_1$  est donnée par

$$\begin{aligned}
 F^{X_1}(x) &= \mathbb{P}[X_1 \leq x] \\
 &= \begin{cases} 0 & \text{pour } x < 0 \\ \mathbb{P}[X_1 \leq x | U_1] \mathbb{P}[U_1] + \mathbb{P}[X_1 \leq x | U_2] \mathbb{P}[U_2] & \text{pour } 0 \leq x \leq a \\ \mathbb{P}[X_1 \leq x | U_1] \mathbb{P}[U_1] + \mathbb{P}[X_1 \leq x | U_2] \mathbb{P}[U_2] & \text{pour } a \leq x \leq b \\ 1 & \text{pour } x > b \end{cases} \\
 &= \begin{cases} 0 & \text{pour } x < 0 \\ p\frac{x}{a} + \frac{(1-p)x}{b} & \text{pour } 0 \leq x \leq a \\ p + (1-p)\frac{x}{b} & \text{pour } a \leq x \leq b \\ 1 & \text{pour } x > b \end{cases}
 \end{aligned}$$

La fonction de densité est alors donnée par

$$f^{X_1}(x) = \begin{cases} 0 & \text{pour } x < 0 \\ \frac{p}{a} + \frac{(1-p)}{b} & \text{pour } 0 \leq x \leq a \\ \frac{1-p}{b} & \text{pour } a \leq x \leq b \\ 0 & \text{pour } x > b \end{cases}$$

Soit maintenant  $N_a$  le nombre d'observations entre 0 et  $a$ . Les variables aléatoires étant supposées indépendantes, chacune de celle-ci appartiendra à l'intervalle avec probabilité  $F^{X_1}(a)$ . Ainsi,

$$N_a \sim \text{Bin} \left( n, \tilde{p} = p + \frac{(1-p)a}{b} \right).$$

Des résultats connus nous donnent espérance et variance en bonus :

$$\mathbb{E}[N_a] = n\tilde{p} \quad \text{et} \quad \text{Var}(N_a) = n\tilde{p}(1 - \tilde{p}).$$

Cherchons alors l'estimateur maximum de vraisemblance pour  $p$ ,  $a$  et  $b$  étant supposé connus. On trouve

$$\begin{aligned}
 L_p(\underline{\mathbf{X}}) &= \prod_{i=1}^n f^X(X_i) \\
 &= \left( \frac{p}{a} + \frac{(1-p)}{b} \right)^{N_a(\underline{\mathbf{X}})} \left( \frac{1-p}{b} \right)^{(n-N_a(\underline{\mathbf{X}}))} \\
 \log L_p(\underline{\mathbf{X}}) &= N_a(\underline{\mathbf{X}}) \log \left( \frac{p}{a} + \frac{(1-p)}{b} \right) + (n - N_a(\underline{\mathbf{X}})) \log \left( \frac{1-p}{b} \right) \\
 \partial_p \log L_p(\underline{\mathbf{X}}) &= N_a(\underline{\mathbf{X}}) \left( \frac{b-a}{pb + (1-p)a} \right) + (n - N_a(\underline{\mathbf{X}})) \left( \frac{-1}{(1-p)} \right) \\
 \partial_p \log L_p(\underline{\mathbf{X}}) = 0 &\Leftrightarrow N_a(\underline{\mathbf{X}}) \left( \frac{b-a}{pb + (1-p)a} \right) + (n - N_a(\underline{\mathbf{X}})) \left( \frac{-1}{(1-p)} \right) = 0 \\
 &\Leftrightarrow \dots \\
 &\Leftrightarrow p = \frac{\frac{N_a(\underline{\mathbf{X}})b}{n} - a}{b-a} = 1 - \frac{\left( \frac{(n-N_a(\underline{\mathbf{X}})b}{n} \right)}{b-a}
 \end{aligned}$$

Remarque : Cet estimateur de  $p$  est bien borné par 1. Il peut par contre être négatif (voir le cas  $N_a(\underline{\mathbf{X}}) = 0$ ). On posera dans ce cas  $p = 0$ . En pratique, ce  $p$  sera-t-il souvent négatif ?