ISFA- 2ème année (M1)
STATISTIQUE INFÉRENTIELLE

FICHE DE TD N° 2

Exercice 1.

On considère un modèle d'échantillonnage avec $n > 1$ et pour lequel P_λ est une loi de poisson de paramètre $\lambda > 0$. On a donc les X_i iid avec $\forall i = 1, \ldots, n$, $X_i \sim$ Poisson(λ). On pose :

$$T_n = \sum_{i=1}^{n} X_i \text{ et } T'_n = X_1.$$

1. Montrer que la statistique T_n est exhaustive pour le paramètre λ et que T'_n ne l'est pas.
2. Soit $k \in \mathbb{N}$. On veut estimer $p(\lambda) = \mathrm{P}_\lambda(X_i = k)$. Calculer la moyenne et la variance de N_k/n où :

$$N_k = \sum_{i=1}^{n} \mathbf{1}_{\{\mathbf{X_i=k}\}}$$

est le nombre de X_i égaux à k.

3. Calculer $E\left(\frac{N_k}{n}/T_n\right)$. C'est l'unique estimateur sans biais de $p(\lambda)$ fonction de N_k/n.

Exercice 2. Soit $X = (X_1, \ldots, X_n)$ un n-échantillon de loi P_θ. Donner la vraisemblance du modèle, vérifier sa régularité et calculer l'information de Fisher dans le cas où :

1. P_θ est une loi de poisson de paramètre θ.
2. P_θ est une loi de Pareto de paramètre $\alpha > 1$ et $\beta > 0$, fixé de densité

$$f(x, \theta, \beta) = \frac{\theta - 1}{\beta}\left(\frac{\beta}{x}\right)^{\theta} \mathbf{1}_{x \geq \beta}$$

3. P_θ est une loi exponentielle de paramètre $\theta > 0$
4. P_θ est une loi uniforme sur $[0, \theta], \theta > 0$.

Exercice 3.

Soit $X = (X_1, \ldots, X_n)$ un n-échantillon de loi P_θ.

1. On suppose que P_θ est une loi de Poisson de paramètre $|\theta|$ si $-1 \leq \theta < 0$ et P_θ est une loi de Bernoulli de paramètre θ si $0 \leq \theta \leq 1$. Calculer la loi conditionnelle de X sachant $n\overline{X}$. En déduire que $n\overline{X}$ n'est pas exhaustive.
2. Vérifier que le modèle est dominé et en appliquant le théorème de factorisation, donner une statistique exhaustive pour $\theta \in [0, 1]$.

3. On suppose maintenant que P_θ est une loi uniforme discrète sur $\{1, 2, \ldots, \theta\}$, où $\theta \in \mathbb{N}^*$.
   a. Vérifier que le modèle statistique est dominé et donner une vraisemblance du modèle.
   b. Déterminer l'estimateur du maximum de vraisemblance de θ.

Exercice 4.

L'organisateur d'une exposition s'intéresse au rythme d'arrivées de groupes de visiteurs à partir des observations faites au cours des premières journées. Il constate que le temps séparant l'arrivée de deux groupes successifs peut être assimilé à une variable X de loi uniforme sur $[0, \theta]$ et que ces temps inter-arrivées sont indépendantes Il souhaite estimer θ à partir de l'observation d'un n-échantillon $X = (X_1, \ldots, X_n)$ de ces inter-arrivées.

1. Vérifier que le modèle statistique est dominé et donner une vraisemblance du modèle.
2. Déterminer l'estimateur du maximum de vraisemblance T_n de θ et en déduire sa loi.
3. Déterminer β tel que $W_n = \beta T_n$ soit un estimateur sans biais de θ.

Exercice 5.

Soit $X = (X_1, \ldots, X_n)$ un n-échantillon de loi P_θ.

1. On suppose que P_θ est une loi de Poisson de paramètre θ. On veut estimer θ^2.
   a. Vérifier que le modèle statistique est dominé et donner une vraisemblance du modèle.
   b. Déterminer l'estimateur du maximum de vraisemblance de θ^2. Est-il sans biais ?
2. On suppose que P_θ est une loi exponentielle de paramètre $1/\theta$.
   a. Vérifier que le modèle statistique est dominé et donner une vraisemblance du modèle.
   b. Déterminer l'estimateur du maximum de vraisemblance de θ.

Exercice 6.

Soit $X = (X_1, \ldots, X_n)$ un n-échantillon de loi $\mathcal{N}(0, \theta), \theta > 0$.

1. On veut estimer θ par la méthode du maximum de vraisemblance. On pose

$$T_n(X_1, \ldots, X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

   a. Soit $p(\theta, x_1, \ldots, x_n)$ la fonction de vraisemblance. Calculer :

$$2\left(\ln p(\theta, x_1, \ldots, x_n) - \ln p(T_n(x_1, \ldots, x_n), x_1, \ldots, x_n)\right).$$

   b. En utilisant l'inégalité $\ln x \leq x - 1$, montrer que T_n est l'estimateur du maximum de vraisemblance de θ.
   c. Montrer que T_n est sans biais et calculer $R(T_n, \theta)$.
2. Calculer l'information de Fisher pour l'estimation de θ

3. L'estimateur T_n est-il efficace de θ ?

4. On veut maintenant estimer θ par la méthode de Bayes relatif à la loi à priori ayant pour densité :
$$g(\theta) = \frac{\lambda^a}{\Gamma(a)} \frac{1}{\theta^{a+1}} \exp\left(-\frac{\lambda}{\theta}\right) \mathbb{1}_{\{\theta>0\}},$$
où $\lambda > 0, a > 1$ et $\Gamma(a+1) = a\Gamma(a)$. Les estimateurs de Bays sont définis par :
$$T_{\lambda,a} = \frac{nT_n + 2\lambda}{2a+n-2}.$$
Montrer qu'ils sont asymptotiquement sans biais.

5. On veut maintenant comparer l'estimateur du maximum de vraisemblance avec les estimateurs bayésiens.

   a. Peut-on les comparer au sens minimax ?

   b. Montrer que la limite de $R(T_{\lambda,a}, \theta) - R(T_n, \theta)$ lorsque $\theta \to 0$ est strictement positive.

   c. Pour tous λ et a trouver un point $\theta^*_{\lambda,a}$ où $R(T_{\lambda,a}, \theta^*_{\lambda,a}) < R(T_n, \theta^*_{\lambda,a})$.

   d. Conclure.

Exercice 7.

1. Montrer que si T_n est un estimateur efficace de θ, alors $kT_n + b$ est aussi un estimateur efficace de $\theta, \forall k \in \mathbb{R}^*, \forall b \in \mathbb{R}$. On considère la loi normale $\mathcal{N}(\mu, \sigma^2)$.

2. On suppose σ^2 connue et l'on considère le modèle paramétrique réel d'échantillonnage suivant :
$$\left(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathcal{N}(\mu, \sigma^2), \mu \in \Theta = \mathbb{R}\right)^n.$$
L'estimateur $\overline{X}$ est-il un estimateur efficace de μ ?

3. Sans supposer μ connue, on pose $\theta = \sigma^2$ et l'on considère le modèle paramétrique réel d'échantillonnage suivant :
$$\left(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathcal{N}(\mu, \sigma^2), \sigma^2 \in \Theta = \mathbb{R}_+^*\right)^n.$$
L'estimateur $S^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2$ est-il un estimateur efficace de σ^2 ?

4. Écrire la vraisemblance du modèle. Vérifier que le couple $(\overline{X}, S^2)$ forment une statistique exhaustive pour le paramètre (μ, σ^2).

Exercice 2. Soit $X = (X_1, \dots, X_n)$ un n-échantillon de loi P_θ. Donner la vraisemblance du modèle, vérifier sa régularité et calculer l'information de Fisher dans le cas où :

1. P_θ est une loi de poisson de paramètre θ.
2. P_θ est une loi de Pareto de paramètre $\alpha > 1$ et $\beta > 0$, fixé de densité
$$f(x, \theta, \beta) = \frac{\theta - 1}{\beta}\left(\frac{\beta}{x}\right)^\theta \mathbf{1}_{x \geq \beta}$$
3. P_θ est une loi exponentielle de paramètre $\theta > 0$
4. P_θ est une loi uniforme sur $[0, \theta]$, $\theta > 0$.

Exercice 2:

3) P_θ loi exponentielle paramètre θ.

Vraisemblance :

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n \exp\left(-\theta \sum_{i=1}^{n} x_i\right)$$

Régularité :

(H1), (H2) et (H3) vérifiées de façon triviale (H4) vérifiée aussi mais + chiant

Informat° :

$$\frac{\partial}{\partial \theta} \ln L = \frac{n}{\theta} - \sum_{i=1}^{n} x_i$$

$$\to I_n(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln L\right)^2\right] = \mathbb{E}\left[\left(\frac{n}{\theta} - \sum x_i\right)^2\right] = \dots$$

$$\to I_n(\theta) = V\left(\frac{\partial}{\partial\theta}\ln L\right) = V\left(\frac{n}{\theta} - \sum x_i\right) = \sum V(x_i) = \frac{n}{\theta^2}$$

$$\frac{\partial^2}{\partial \theta^2} \ln L = -\frac{n}{\theta^2}$$

$$\Rightarrow I_n(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L\right] = \mathbb{E}\left[\frac{n}{\theta^2}\right] = \frac{n}{\theta^2}$$

Estimateur du max de vraisemblance :

Comme on a (H1) à (H4), l'EMV $\hat{\theta}_n$ de θ est tq

$$\begin{cases} \frac{n}{\hat{\theta}_n} - \sum x_i = 0 \\ -\frac{n}{\hat{\theta}_n^2} < 0 \end{cases} \Rightarrow \hat{\theta}_n = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{X}_n}$$

4) $X_1, \dots, X_n$ iid tq $X_i \sim \mathcal{U}[0,\theta]$ i.e de densité $f_\theta(x) = \frac{1}{\theta} 1_{[0,\theta]}(x)$

Vraisemblance :

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \frac{1}{\theta} 1_{[0,\theta]}(x_i) = \frac{1}{\theta^n} \underbrace{\prod_{i=1}^{n} 1_{[0,\theta]}(x_i)}_{= \begin{cases} 1 \text{ si tous les } x_i \in [0,\theta] \\ 0 \text{ sinon.} \end{cases}}$$

$$= \frac{1}{\theta^n} 1_{(\min x_i \geq 0)} 1_{(\max x_i \leq \theta)}$$

Régularité :

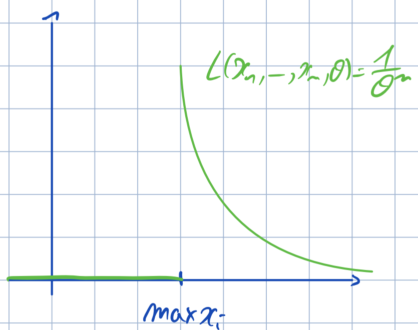
On a $\mathcal{X} = [0,\theta]$ dépend de $\theta \Rightarrow$ (H1) pas vérifiée

→ pas d'informat°, pas efficacité, pas de normalité asympt, pas de modèles expo

→ On a quand m̂ l'EMV, biais, convergence et exhaustivité

Estimateur du max de vraisemblance :

Comme on n'a pas (H1) etc, on trouve l'EMV $\hat{\theta}_n$ de θ en maximisant direct^t la vraisemblance
On observe $x_1, \dots, x_n$ fixés



Remarque : (culture générale mais hors sujet).

On peut quand m̂ calculer l'info (en faisant comme si on n'avait plus $1_{(\dots)}$)

$\ln L(x_1, \dots, x_n, \theta) = -n \ln \theta$	$\ln f(x) = -\ln \theta$
$\frac{\partial}{\partial \theta} \ln L = -\frac{n}{\theta}$	$\frac{\partial}{\partial \theta} \ln f(x) = -\frac{1}{\theta}$
$I_n(\theta) = \mathbb{E}[(\frac{\partial}{\partial \theta} \ln L)^2] = \frac{n^2}{\theta^2}$	$I(\theta) = \frac{1}{\theta^2}$

$$I_n(\theta) = \frac{n^2}{\theta^2} \neq n I(\theta) = \frac{n}{\theta^2}$$