Sorbonne Université
Master Mathématiques et Applications
MU4MA015 : Statistique

2 décembre 2022
Aucun document autorisé
Durée : 1 heure

# Contrôle Continu

## Rappels

- Il sera tenu compte du soin apporté à la présentation et à la rédaction. Si vous n'arrivez pas à démontrer un résultat, vous pouvez l'admettre pour la suite de l'exercice.
- Si $T$ suit une loi exponentielle de paramètre $\lambda > 0$, alors $\mathbb{E}[T] = \frac{1}{\lambda}$ et $\mathrm{Var}(T) = \frac{1}{\lambda^2}$.

**Exercice 1** (4 points)
Soit $\theta > -1$ un paramètre inconnu et $X$ de densité

$$f_\theta(x) = (1+\theta)x^\theta \mathbf{1}_{]0,1[}(x).$$

Dans la suite, on suppose disposer d'un échantillon $(X_1, \ldots, X_n)$ i.i.d. de même loi que $X$.

1. Montrer que la variable $Y = -\log X$ suit une loi exponentielle de paramètre $(1+\theta)$. En déduire $\mathbb{E}[Y]$ et $\mathrm{Var}(Y)$.
**Correction.** Pour toute fonction mesurable bornée $\varphi$, le Théorème de Transfert et le changement de variable $y = -\log x$ permettent d'écrire
$$\mathbb{E}[\varphi(Y)] = \mathbb{E}[\varphi(-\log X)] = \int_0^1 \varphi(-\log x) f_\theta(x) dx = \int_0^\infty \varphi(y)(1+\theta)e^{-(1+\theta)y} dy,$$
ce qui revient bien à dire que $Y \sim \mathcal{E}(1+\theta)$. On en déduit que $\mathbb{E}[Y] = \frac{1}{1+\theta}$ et $\mathrm{Var}(Y) = \frac{1}{(1+\theta)^2}$.
2. Déterminer l'estimateur du maximum de vraisemblance $\widehat{\theta}_n$. Démontrer qu'il est fortement consistant et asymptotiquement normal.
**Correction.** Pour tout $\theta > -1$, la log-vraisemblance vaut
$$\ell_n(\theta) = n\log(1+\theta) + \theta \sum_{i=1}^n \log X_i$$
donc
$$\ell_n'(\theta) = \frac{n}{1+\theta} + \sum_{i=1}^n \log X_i.$$
On voit que $\ell_n'(\theta) > 0$ si $\theta < -1 - (\frac{1}{n}\sum_{i=1}^n \log X_i)^{-1}$ et $\ell_n'(\theta) < 0$ si $\theta > -1 - (\frac{1}{n}\sum_{i=1}^n \log X_i)^{-1}$ donc l'EMV est
$$\widehat{\theta}_n = -1 - (\frac{1}{n}\sum_{i=1}^n \log X_i)^{-1}$$

La loi forte des grands nombres et le Théorème de continuité impliquent que

$$\widehat{\theta}_n \xrightarrow[n\to\infty]{ps} -1 - (\mathbb{E}[\log X])^{-1} = -1 - (\mathbb{E}_\theta[-Y])^{-1} = \theta,$$

c'est-à-dire que $\widehat{\theta}_n$ est fortement consistant. Le Théorème Central Limite donne, puisque $\mathbb{E}[\log X] = -\frac{1}{1+\theta}$ et $\mathrm{Var}(\log X) = \mathrm{Var}(Y) = \frac{1}{(1+\theta)^2}$,

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\log X_i - \left(-\frac{1}{1+\theta}\right)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{(1+\theta)^2}\right).$$

Autrement dit, en notant $\phi(t) = -\frac{1}{1+t}$ le $C^1$-difféomorphisme de $]-1,\infty[$ vers $]-\infty,0[$ de réciproque $\phi^{-1}(u) = -1 - \frac{1}{u}$,

$$\sqrt{n}\left(\phi(\widehat{\theta}_n) - \phi(\theta)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{(1+\theta)^2}\right).$$

La méthode Delta permet de conclure que

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, (1+\theta)^2\right).$$

3. Le modèle $(f_\theta)_{\theta>-1}$ est-il régulier ? Si oui, préciser l'information de Fisher pour une seule donnée $I(\theta) = I_1(\theta)$.
**Correction.** Pour tout $x \in ]0,1[$, la fonction $\theta \mapsto f_\theta(x)$ est de classe $C^1$ sur $]-1,\infty[$ donc les deux premiers points de la définition d'un modèle régulier sont satisfaits. Pour le dernier, on a $\ell_\theta(X) = \log(1+\theta) + \theta\log X$ donc le score est la variable aléatoire

$$\ell'_\theta(X) = \log X + \frac{1}{1+\theta} = \log X - \mathbb{E}[\log X].$$

Son moment d'ordre 2 s'en déduit :

$$\mathbb{E}[\ell'_\theta(X)^2] = \mathrm{Var}(\log X) = \frac{1}{(1+\theta)^2},$$

qui est une fonction continue sur $]-1,\infty[$ donc le modèle est régulier, d'information de Fisher $I(\theta) = \frac{1}{(1+\theta)^2}$.

4. L'estimateur $\widehat{\theta}_n$ est-il asymptotiquement efficace ?
**Correction.** Au total, nous avons établi que

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, (1+\theta)^2\right) = \mathcal{N}\left(0, \frac{1}{I(\theta)}\right),$$

ce qui assure que $\widehat{\theta}_n$ est asymptotiquement efficace.

**Exercice 2** (6 points)
Soient $\lambda > 0$ et $n \geq 1$ entier. On considère le modèle bayésien suivant

$$\boldsymbol{\theta} \sim \Pi = \mathcal{E}(\lambda),$$
$$\mathbf{X} = (X_1, \ldots, X_n) \mid \boldsymbol{\theta} \sim \mathcal{P}(\boldsymbol{\theta})^{\otimes n},$$

où $\mathcal{E}(\lambda)$ est la loi exponentielle de paramètre $\lambda$, et $\mathcal{P}(\theta)$ la loi de Poisson de paramètre $\theta$.

1. Déterminer la loi a posteriori $\Pi(\cdot \mid \mathbf{X})$.
**Correction.** Pour tout $\theta > 0$, la densité a posteriori est donnée par

$$\begin{aligned}\pi(\theta \mid \mathbf{X}) &= \frac{\pi(\theta)\prod_{i=1}^n p_\theta(X_i)}{f(\mathbf{X})} \\ &\propto e^{-(\lambda+n)\theta}\theta^{n\overline{X}_n}\mathbf{1}_{\theta>0}\,.\end{aligned}$$

On reconnaît le terme général de la densité d'une loi $\Gamma(n\overline{X}_n + 1, \lambda + n)$.

On considère la fonction de perte $\ell$ définie par

$$\forall \theta > 0\,,\ \forall t > 0\,,\ \ell(\theta, t) = \frac{1}{\theta}(t-\theta)^2\,.$$

2. Déterminer un estimateur de Bayes pour $\Pi$ et la perte $\ell$. On le notera $T^\star$.
**Correction.** On cherche à minimiser le risque a posteriori. Pour un estimateur $T$, on a

$$\begin{aligned}\rho(\Pi, T \mid \mathbf{X}) &= \int_\Theta \frac{1}{\theta}(T(\mathbf{X})-\theta)^2 \mathrm{d}\Pi(\theta \mid \mathbf{X}) \\ &= \frac{(\lambda+n)^{n\overline{X}_n+1}}{\Gamma(n\overline{X}_n+1)}\int_0^{+\infty}\frac{1}{\theta}(T(\mathbf{X})-\theta)^2\theta^{n\overline{X}_n}e^{-(\lambda+n)\theta}\mathrm{d}\theta \\ &= \frac{(\lambda+n)^{n\overline{X}_n+1}}{\Gamma(n\overline{X}_n+1)}\int_0^{+\infty}(T(\mathbf{X})-\theta)^2\theta^{n\overline{X}_n-1}e^{-(\lambda+n)\theta}\mathrm{d}\theta\end{aligned}$$

Si $\overline{X}_n = 0$, ce qui équivaut à dire que $\mathbf{X} = \mathbf{0} := (0, \dots, 0)$, la seule possibilité pour que cette quantité ne soit pas infinie est que $T(\mathbf{0}) = 0$. Si $\overline{X}_n > 0$, on remarque que, modulo des constantes multiplicatives, l'intégrale peut être vue comme le risque a posteriori de $T$ pour la perte quadratique et une loi a posteriori $\Gamma(n\overline{X}_n, \lambda + n)$. Le cours nous dit alors qu'un estimateur de Bayes est donné par

$$T^\star(\mathbf{X}) = \frac{n\overline{X}_n}{\lambda+n}\,,$$

qui vérifie bien $T^\star(\mathbf{0}) = 0$.

3. Montrer que la fonction de risque de $T^\star$ (ou risque ponctuel) est donnée par

$$\forall \theta > 0\,,\ \mathbf{R}(\theta, T^\star) = \frac{n+\lambda^2\theta}{(n+\lambda)^2}\,.$$

**Correction.** Pour tout $\theta > 0$, on a

$$\begin{aligned}\mathbf{R}(\theta, T^\star) &= \mathbb{E}_\theta\left[\frac{1}{\theta}\left(T^\star(\mathbf{X})-\theta\right)^2\right] \\ &= \frac{1}{\theta}\mathbb{E}_\theta\left[\left(\frac{n}{n+\lambda}(\overline{X}_n-\theta) - \frac{\lambda\theta}{\lambda+n}\right)^2\right] \\ &= \frac{1}{\theta}\left(\frac{n^2}{(n+\lambda)^2}\mathbb{E}_\theta[(\overline{X}_n-\theta)^2] + \frac{\lambda^2\theta^2}{(\lambda+n)^2}\right) \\ &= \frac{1}{\theta}\left(\frac{n^2}{(n+\lambda)^2}\frac{\theta}{n} + \frac{\lambda^2\theta^2}{(\lambda+n)^2}\right) \\ &= \frac{n+\lambda^2\theta}{(\lambda+n)^2}\,,\end{aligned}$$

où l'on a utilisé que $\mathbb{E}_\theta[\overline{X}_n] = \theta$ et $\mathrm{Var}_\theta(\overline{X}_n) = \theta/n$.

4. Que vaut le risque de Bayes $\mathbf{R}_B(\Pi)$ ?
**Correction.** Le risque de Bayes s'obtient en prenant l'espérance de $\frac{n+\lambda^2\boldsymbol{\theta}}{(\lambda+n)^2}$ pour $\boldsymbol{\theta} \sim \mathcal{E}(\lambda)$ :
$$\mathbf{R}_B(\Pi) = \mathbb{E}\left[\frac{n+\lambda^2\boldsymbol{\theta}}{(\lambda+n)^2}\right] = \frac{n+\lambda^2\mathbb{E}[\boldsymbol{\theta}]}{(\lambda+n)^2} = \frac{1}{\lambda+n},$$
puisque $\mathbb{E}[\boldsymbol{\theta}] = 1/\lambda$.

5. Que vaut le risque maximal de l'estimateur $\overline{X}_n$ (pour la perte $\ell$) ?
**Correction.** Pour $\theta > 0$, on a
$$\mathbf{R}(\theta, \overline{X}_n) = \frac{1}{\theta}\mathbb{E}_\theta\left[(\overline{X}_n - \theta)^2\right] = \frac{1}{n}\cdot$$
La fonction de risque de $\overline{X}_n$ est donc constante et $\mathbf{R}_{\max}(\overline{X}_n) = \frac{1}{n}$.

6. Montrer que $\overline{X}_n$ est minimax (toujours pour la perte $\ell$).
**Correction.** En considérant la suite de lois a priori $(\Pi_k)_{k\geq 1}$ avec $\Pi_k = \mathcal{E}(1/k)$, on a, par la question 6,
$$\lim_{k\to+\infty} \mathbf{R}_B(\Pi_k) = \lim_{k\to+\infty} \frac{1}{1/k+n} = \frac{1}{n} = \mathbf{R}_{\max}(\overline{X}_n).$$
Comme pour tout $k \geq 1$, on a $\mathbf{R}_B(\Pi_k) \leq \mathbf{R}_M$, on en déduit que $\overline{X}_n$ est minimax.

7. Grâce à la formule donnée à la question 3, on veut construire un estimateur Monte Carlo de $\mathbf{R}_B(\Pi)$ et afficher son évolution en fonction de la taille de l'échantillon Monte Carlo. Par quoi faut-il remplacer `???` et `%%%` pour que le script suivant produise le graphique de la Figure 1 (aucune justification n'est demandée) ?

```
Lambda = 2  # Paramètre de la loi a priori
n = 30  # Taille du vecteur aléatoire X
N = 1000  # Taille de l'échantillon Monte Carlo
prior = stats.expon(scale=1/Lambda)  # Loi a priori

sample_theta = ???  # N réalisations de la loi a priori
risk = (n + Lambda**2 * sample_theta) / (n + Lambda)**2
est_MC = %%%  # Tableau des valeurs de l'estimateur Monte Carlo en fonction de N

plt.plot(np.arange(1, N+1), est_MC, label='Estimateur Monte Carlo')
plt.axhline(1/(n+Lambda), color='r', label='$\\mathbf{{R}}_{{B}}(\\Pi)$')
plt.xlabel('$N$')
plt.legend()
```

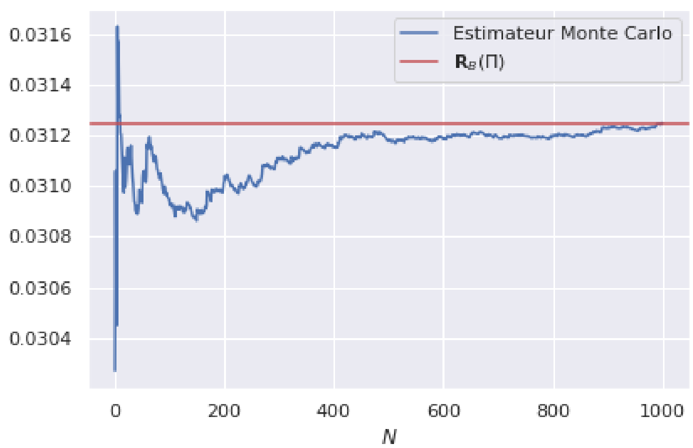


Figure 1 – Estimation Monte Carlo de $\mathbf{R}_B(\Pi)$.