Statistique/Modèle linéaire

Partiel du 14 novembre 2001
Corrigé

Exercice 1

1. *Montrer que $L(\theta, d)$ est toujours positive et qu'elle ne s'annule que pour $d = \theta$:* La fonction $f(x) = \exp(x) - x - 1$ admet comme derivée $\exp(x) - 1$, strictement négative pour $x < 0$ et strictement positive pour $x > 0$.

2. Par un développement en série entière de $\exp(x)$, on voit que
$$f(x) = \frac{x^2}{2} + \mathrm{o}(x^2)$$
où $x^{-2}\mathrm{o}(x^2)$ tend vers 0 quand x tend vers 0. Par conséquent,
$$\lim_{c \to 0} \frac{\mathrm{L}(\theta, d)}{c^2} = \frac{(d-\theta)^2}{2}\,.$$

3. *L'estimation de Bayes sous le coût LINEX*, $\delta^\pi(x)$, est donné par
$$\arg\min \mathbb{E}^\pi \left[e^{c(d-\theta)} - c(d-\theta) - 1 \middle| x \right]$$
soit, par dérivation,
$$c\, e^{c\delta^\pi(x)}\, \mathbb{E}^\pi \left[e^{-c\theta} \middle| x \right] = c$$
soit encore
$$\delta^\pi(x) = \frac{-1}{c} \log\left\{ \mathbb{E}^\pi [e^{-c\theta}|x] \right\}\,. \tag{1}$$

4. Dans le cas où $x \sim \mathcal{N}(\theta, 1)$ et $\pi(\theta) = 1$, la loi a posteriori de θ est la loi normale $\mathcal{N}(x, 1)$, pour laquelle
$$\begin{aligned}
\mathbb{E}[e^{-c\theta}|x] &= \int e^{-c\theta}\, e^{-(x-\theta)^2/2} \frac{d\theta}{\sqrt{2\pi}} \\
&= \int e^{-x^2/2 + (x-c)\theta - \theta^2/2} \frac{d\theta}{\sqrt{2\pi}} \\
&= e^{-x^2/2}\, e^{(x-c)^2/2} = e^{c^2/2 - xc}
\end{aligned}$$

transformée de Laplace pour la loi normale. Donc

$$\delta^\pi(x) = \frac{-1}{c}(c^2/2 - xc) = x - \frac{c}{2}\,.$$

5. *Comparer le risque de l'estimateur de Bayes avec celui de* $\delta_0(x) = x$ *:* On a

$$\begin{aligned} R(\theta, \delta^\pi) &= \mathbb{E}_\theta\left[e^{c(X-c/2-\theta)}\right] - c\mathbb{E}_\theta\left[X - c/2 - \theta\right] - 1 \\ &= \mathbb{E}_\theta\left[e^{cX}\right] e^{-c^2/2-c\theta} + c^2/2 - 1 \\ &= e^{c\theta+c^2/2-c^2/2-c\theta} + c^2/2 - 1 \\ &= c^2/2 \end{aligned}$$

   et

$$R(\theta, \delta_0) = \mathbb{E}_\theta\left[e^{c(X-\theta)}\right] - c\mathbb{E}_\theta\left[X - \theta\right] - 1 = e^{c^2/2} - 1$$

   donc le risque de δ_0 est toujours plus grand (par l'argument de la question 1).

6. Lorsque $x \sim \mathcal{G}(\alpha, \theta)$ et $\pi(\theta) = 1/\theta$ sur $\mathbb{R}_+$,

$$\begin{aligned} \pi(\theta|x) &\propto f(x|\alpha,\theta)\pi(\theta) \\ &\propto \theta^\alpha x^{\alpha-1} e^{-\theta x}\,\frac{1}{\theta} \\ &\propto \theta^{\alpha-1} e^{-\theta x}\,, \end{aligned}$$

   loi $\mathcal{G}(\alpha, x)$. Et

$$\begin{aligned} \mathbb{E}[e^{-c\theta}|x] &= \int_{\mathbb{R}_+} e^{-c\theta}\frac{\theta^{\alpha-1}x^\alpha e^{-\theta x}}{\Gamma(\alpha)}\,d\theta \\ &= \int_{\mathbb{R}_+} \frac{\theta^{\alpha-1}x^\alpha e^{-\theta(c+x)}}{\Gamma(\alpha)}\,d\theta \\ &= \left(\frac{x}{c+x}\right)^\alpha \end{aligned}$$

   Donc

$$\delta^\pi(x) = -\frac{\alpha}{c}\,\log\left(\frac{x}{c+x}\right)\,,$$

   équivalent à αx pour c petit.

Exercice 2

1. *Donner la vraisemblance de cet échantillon et montrer qu'elle factorise en $\overline{x}$, moyenne empirique des x_i :* On a la vraisemblance

$$\begin{aligned} L(\theta|x_1,\dots,x_n) &= \prod_{i=1}^{n} e^{-(x_i-\theta)^2/2\sigma^2}\frac{1}{\sqrt{2\pi}} \\ &\propto \exp\frac{-1}{2\sigma^2}\left[n(\overline{x}-\theta)^2+s_n^2\right]^2 \end{aligned}$$

où s^2 est la somme des carrés des écarts à la moyenne. On se ramène donc à la loi de la statistique exhaustive $\overline{X}$, $\mathcal{N}(\theta,\sigma^2/n)$.

2. Pour le test de H_0 : $\theta=\theta_0$ contre H_1 : $\theta=\theta_1$, le théorème de Neyman–Pearson conduit à la procédure optimale suivante : accepter H_0 si et seulement si

$$\frac{L(\theta_0|\overline{x})}{L(\theta_1|\overline{x})} > c \tag{2}$$

et

$$\begin{aligned} \frac{L(\theta_0|\overline{x})}{L(\theta_1|\overline{x})} &= \frac{e^{-n(\overline{x}-\theta_0)^2/2\sigma^2}}{e^{-n(\overline{x}-\theta_1)^2/2\sigma^2}} \\ &= \frac{e^{n\overline{x}\theta_0/\sigma^2-n\theta_0^2/2\sigma^2}}{e^{nx\theta_1/\sigma^2-n\theta_1^2/2\sigma^2}} \\ &= e^{n\overline{x}(\theta_0-\theta_1)/\sigma^2-n(\theta_0^2-\theta_1^2)/2\sigma^2} \end{aligned}$$

donc, remplaçant c par k^n dans (2), on a bien que la procédure de Neyman–Pearson est d'accepter H_0 si et seulement si

$$\overline{x} < \frac{(2\sigma^2\log k)-\theta_0^2+\theta_1^2}{2(\theta_1-\theta_0)}\,. \tag{3}$$

3. *Le choix d'un niveau α pour le test de H_0 contre H_1 conduit à une borne (3) indépendante de θ_1 :* en effet, la borne est déterminée sous

H_0 uniquement par

$$\begin{aligned}
&P_{\theta_0}\left(\overline{X} > \frac{(2\sigma^2 \log k) - \theta_0^2 + \theta_1^2}{2(\theta_1 - \theta_0)}\right) = 1 - \alpha \\
&= P_{\theta_0}\left(\frac{\sqrt{n}(\overline{X} - \theta_0)}{\sigma} > \sqrt{n}\frac{(2\sigma^2 \log k) - \theta_0^2 + \theta_1^2 - 2\theta_0(\theta_1 - \theta_0)}{2\sigma(\theta_1 - \theta_0)}\right) \\
&= 1 - \Phi\left(\sqrt{n}\frac{(2\sigma^2 \log k) - \theta_0^2 + \theta_1^2 - 2\theta_0(\theta_1 - \theta_0)}{2\sigma(\theta_1 - \theta_0)}\right) ,
\end{aligned}$$

soit encore

$$\sqrt{n}\frac{(2\sigma^2 \log k) - \theta_0^2 + \theta_1^2 - 2\theta_0(\theta_1 - \theta_0)}{2\sigma(\theta_1 - \theta_0)} = q_\alpha$$

ce qui implique que

$$\frac{(2\sigma^2 \log k) - \theta_0^2 + \theta_1^2}{2(\theta_1 - \theta_0)} = \theta_0 + \frac{\sigma q_\alpha}{\sqrt{n}} ,$$

effectivement indépendent de $\theta_1 > \theta_0$.

4. La question 3 implique que le test précédent est aussi uniformément plus puissant pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta > \theta_0$ puisque la borne ne dépend pas de θ_1. Il est aussi uniformément plus puissant pour tester $H_0 : \theta = \theta_2$ contre $H_1 : \theta > \theta_2$ pour tout $\theta_2 < \theta_0$ en adaptant le niveau de α à

$$1 - \Phi\left(q_\alpha + \frac{\sqrt{n}(\theta_0 - \theta_2)}{\sigma}\right) .$$

On obtient donc en conclusion un test uniformément plus puissant pour tester $H_0 : \theta < \theta_0$ contre $H_1 : \theta > \theta_0$, tel que

$$P_\theta(\text{Rejet}) \leq 1 - \alpha$$

pour tous les $\theta \leq \theta_0$.

5. La procédure précédente fournit un test au niveau $\alpha = P_{\theta_0}(\text{Rejet})$ pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$. Par symétrie, la procédure qui accepte H_0 si et seulement si

$$\sqrt{n}\,\frac{\overline{x} - \theta_0}{\sigma} > q_{1-\alpha} ,$$

ce qui revient à tester $H_0 : \ \theta = \theta_0$ contre $H_1 : \ \theta = \theta_1$ pour $\theta_1 < \theta_0$, est aussi un test au niveau α. Les fonctions puissances de ces deux tests se croisent en θ_0 où elles valent toutes deux α : la première croit en θ, la seconde décroit. Elles ne sont donc pas comparables au sens de la fonction puissance.

En déduire qu'il n'existe pas de test uniformément plus puissant pour ce problème : S'il existait une telle procédure, sa fonction puissance devrait être, sur tout $\mathbb{R}$, supérieure à celle des deux procédures précédentes, ce qui est impossible puisque la première est uniformément la plus puissante pour tester $H_0 : \ \theta = \theta_0$ contre $H_1 : \ \theta > \theta_0$ et la seconde pour tester $H_0 : \ \theta = \theta_0$ contre $H_1 : \ \theta < \theta_0$.

6. *Le problème de test inverse*

$$H_0 : \ \theta \neq \theta_0 \quad \text{contre} \quad H_1 : \ \theta = \theta_0 \,.$$

admet une solution optimale car la puissance n'est plus une fonction mais un point,

$$P_{\theta_0}\left(\varphi(\overline{X}) = 0\right) ,$$

qu'il est possible de maximiser sous la *contrainte de niveau*

$$\sup_{\theta \neq \theta_0} P_\theta(\varphi(\overline{X}) = 0) \leq \alpha \,.$$

Par continuité de la fonction puissance (en θ), on voit que le maximum est atteint en α et qu'une procédure optimale est

$$\varphi^*(\overline{x}) = \begin{cases} 1 & \text{si } |\overline{x} - \theta_0| > c, \\ 0 & \text{si } |\overline{x} - \theta_0| < c, \end{cases}$$

où

$$c = c(\alpha) = \frac{\sigma q_{1-(1-\alpha)/2}}{\sqrt{n}}$$