# *Statistique Mathématique*, Feuille 2

**Familles exponentielles**

7) Vérifier si les lois suivantes sont des membres de la famille exponentielle de plein rang. Le cas échéant, trouver la statistique naturelle et le paramètre naturel.

*a*) La loi beta $B(\alpha, \beta)$.
*b*) La loi exponentielle $E(\lambda)$.
*c*) La loi « Gaussienne inverse » de densité $f(x; \mu, \lambda)$ donnée en 3*f*).
*d*) La loi $N(\theta, \theta^2)$.
*e*) La loi de Cauchy $C(\mu, \sigma)$.
*f*) La loi Logistique $Log(\mu, \sigma)$.
*g*) La loi Multinormale $N_p(\boldsymbol{\mu}, \Sigma)$
*h*) La loi lognormale de densité $f(x; \theta) = \dfrac{\theta_1}{\sqrt{2\pi}\,x} \exp\left\{-\tfrac{1}{2}\left\{\theta_2 + \theta_1 \log x\right\}^2\right\}$ si $x > 0$, $\theta_1 > 0$, $\theta_2 \in \mathbb{R}$.
*i*) La loi *des valeurs extrêmes* de densité $f(x; \theta) = \exp\left\{-\exp\left[-\dfrac{(x-\theta_1)}{\theta_2}\right]\right\}$, $x \in \mathbb{R}$, $\theta_1 \in \mathbb{R}$, $\theta_2 > 0$.
*j*) La loi de Laplace $L(\mu, \sigma)$.

8) On considère le modèle statistique pour $n$ observations $X_i$ *iid* de loi normale, de paramètre $\theta = (\mu, \sigma^2)$ où $\mu$ est l'espérance de la loi et $\sigma^2$ dénote sa variance. Trouver, en utilisant le fait que cette loi est membre de la famille exponentielle, la valeur de

$$\mathbb{E}_\theta\left(\sum_{i=1}^n X_i\right),\quad \mathbb{E}_\theta\left(\sum_{i=1}^n X_i^2\right),\quad \mathbb{V}_\theta\left(\sum_{i=1}^n X_i\right),\quad \mathbb{V}_\theta\left(\sum_{i=1}^n X_i^2\right) \quad \text{et} \quad \mathbb{C}ov_\theta\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$$

En déduire $\mathbb{E}_\theta\left(S^2\right)$, $\mathbb{V}_\theta\left(S^2\right)$ et $\mathbb{C}ov_\theta\left(\bar{X}, S^2\right)$ où $\bar{X} = n^{-1}\sum_{i=1}^n X_i$ et $S^2 = \dfrac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2 = \dfrac{n}{n-1}\left(\overline{X^2} - \bar{X}^2\right)$ où de façon générale $\overline{X^r} = n^{-1}\sum_{i=1}^n X_i^r$

9) On considère le modèle statistique de Rayleigh pour $n$ observations $X_i$ *iid* de densité

$$f(x; \theta) = \frac{x}{\theta^2}\exp\left\{-\frac{x^2}{2\theta^2}\right\} \text{ pour } x > 0 \text{ et où } \theta \in \Theta = ]0, \infty).$$

Déterminer si cette loi est membre de la famille exponentielle et trouver

$$\mathbb{E}_\theta\left(\sum_{i=1}^n X_i^2\right),\quad \mathbb{V}_\theta\left(\sum_{i=1}^n X_i^2\right).$$

10) Vérifier que la loi multinomiale $M(1, (p_1, \ldots, p_{q-1}))$ est membre de la famille exponentielle et de plein rang. Considérer $n$ vecteurs aléatoires $\boldsymbol{X}_i$ *iid* $\sim M(1, (p_1, \ldots, p_{q-1}))$. Déterminer l'espérance et les covariances des composantes de $\sum_{i=1}^n \boldsymbol{X}_i$. (prendre d'abord $q = 3$)

11) Montrer que pour une famille exponentielle, l'espace naturel des paramètres est un ensemble convexe (suggestion : utiliser l'inégalité de Hölder)

**Statistiques exhaustives.**

12) Soit $X_1, X_2$ des $v.a.iid$ de loi $P(\theta)$ où $\theta > 0$. Montrer que la loi de $X_1$ étant donné que $T = X_1 + X_2 = t$ est la loi binomiale $N(t, 1/2)$ qui ne dépend pas de $\theta$. Montrer que la loi du couple ($X_1$, $X_2$) étant donné que $T = t$ est $M(t, (1/2))$. En déduire que $T$ est une statistique exhaustive.

13) Montrer la partie suivante du Théorème 3.3 : Soit $\mathcal{M}$ un modèle exponentiel de plein rang de dimension $q$ de statistique naturelle $\varphi$. Alors $\varphi$ est exhaustive minimale.

14) Pour les lois de l'exercice 7) déterminer, si elles existent, les statistiques exhaustives et les statistiques exhaustives minimales.

15) Soit $(\mathcal{X}, \{\mathbb{P}_\theta, \theta \in \Theta\})$ le modèle statistique d'un échantillon aléatoire $X_1, \ldots, X_n$ *iid* de loi $U[\theta - 1/2, \theta + 1/2]$. Montrer que la statistique $(X_{(1)}, X_{(n)})$ n'est pas complète.

**Compléments**

16) Montrer que le vecteur des statistiques d'ordre $(X_{(1)}, \ldots, X_{(n)})$ est toujours une statistique exhaustive. (prendre $n = 2$ pour commencer)

17) Généralisation 1) du Théorème de Rao-Blackwell. Soit $\phi \in \mathcal{E}$, un estimateur SB de $g(\theta)$ et soit $\psi$, une statistique exhaustive. Comme dans le théorème de Rao-Blackwell, soit $\varphi = \mathbb{E}_\theta[\ \phi \mid \psi = \psi(x)]$. Pour le problème d'estimer $g(\theta)$, considérons une fonction de coût convexe $c(\theta, \varphi(\boldsymbol{x}))$ (plutôt que quadratique). En utilisant l'inégalité de Jensen (si $g$ est une fonction convexe, $g(\mathbb{E}(X)) \leq \mathbb{E}(g(X))$, montrer qu'on a toujours $R(\theta, \varphi) \leq R(\theta, \phi)$.

18) Généralisation 2) du Théorème de Rao-Blackwell. Soit $\phi \in \mathcal{E}$, un estimateur SB de $g(\theta) \in \mathbb{R}^p$ et soit $\psi$, une statistique exhaustive. Comme dans le théorème de Rao-Blackwell, soit $\varphi = \mathbb{E}_\theta[\ \phi \mid \psi = \psi(x)]$ avec maintenant $\varphi \in \mathbb{R}^p$. Pour d'estimer $g(\theta)$, considérons la fonction de coût

$$c(\theta, \varphi(\boldsymbol{x})) = \sum_{i=1}^{p} \sum_{j=1}^{p} \upsilon_i \upsilon_j (g_i(\theta) - \varphi_i)(g_j(\theta) - \varphi_j) .$$

Montrer qu'on a $R(\theta, \varphi) \leq R(\theta, \phi)$. Soit $\Gamma_\varphi$ la matrice d'éléments $\mathbb{E}_\theta[(g_i(\theta) - \varphi_i)(g_j(\theta) - \varphi_j)]$ et de même pour la matrice $\Gamma_\phi$. Montrer que l'inégalité $R(\theta, \varphi) \leq R(\theta, \phi)$ du théorème de Rao-Blackwell implique alors que, pour tout $\boldsymbol{\upsilon} \in \mathrm{R}^p$, $\boldsymbol{\upsilon}^\mathsf{T}(\Gamma_{\phi\cdot} - \Gamma_\varphi)\boldsymbol{\upsilon} \geq 0$. Ceci montre que la matrice de variance-covariance de $\varphi$ est plus petite que celle de $\phi$ dans le sens de la relation d'ordre sur les matrices d'ordre $p$ symétriques et définies-positives donnée par $\mathrm{A} \leq \mathrm{B}$ si, pour tout $\boldsymbol{\upsilon} \in \mathrm{R}^p$, $\boldsymbol{\upsilon}^\mathsf{T}(\mathrm{B} - \mathrm{A})\boldsymbol{\upsilon} \geq 0$.