Examen de MATH F-309
Janvier 2018

Nom : **Prénom :** **Section :**

Question 1 (4 points)

Soient $\mathbf{X}_1, \dots, \mathbf{X}_n$ i.i.d. $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\bar{\mathbf{X}} = (\mathbf{X}_1 + \cdots + \mathbf{X}_n)/n$ et soit $\mathbf{a} \in \mathbf{R}^p$.

1. Quelle est la loi de $\mathbf{a}'\bar{\mathbf{X}}$?
2. Donnez la définition d'un vecteur aléatoire $\mathbf{V}$ qui suit une loi de Wishart de paramètres p, m et $\boldsymbol{\Sigma}$.
3. Montrez que

$$\mathbf{W} = \sum_{k=1}^{n} (\mathbf{X}_k - \bar{\mathbf{X}})(\mathbf{X}_k - \bar{\mathbf{X}})'$$

   suit une loi de Wishart avec comme paramètres p, $n-1$ et $\boldsymbol{\Sigma}$.

Question 2 (5 points)

Soit $\mathbf{X} = (X_1, X_2, X_3)'$, où X_1 est une variable aléatoire suivant une loi de Poisson de paramètre $\lambda > 0$, $X_2 = X_1 + 1$ et X_3 est une variable aléatoire indépendante de X_1 et suivant une loi de Poisson de même paramètre λ.

1. Déterminer $\boldsymbol{\Sigma} = \mathrm{Var}(\mathbf{X})$.
2. Déterminer les composantes principales en fonction de λ.
3. Quelle est la part de la variabilité totale expliquée par la première composante principale.

Rappel : X suit une loi de Poisson de paramètre $\lambda > 0$, si $\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$, pour tout entier $k \geq 0$.

Question 3 (4 points)

Soient $\mathbf{X}_1, \dots, \mathbf{X}_n$ i.i.d. $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

1. Quels sont les estimateurs maximum de vraisemblance pour $\boldsymbol{\mu}$ et $\boldsymbol{\Sigma}$?
2. Donnez la preuve du point 1) ci-dessus.

Question 4 (5 points)

Soit $\mathbf{X} = (X_1, X_2)'$ un vecteur aléatoire de loi $\mathcal{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ où

$$\boldsymbol{\mu} = \begin{pmatrix} a \\ 0 \end{pmatrix} \quad \text{et} \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad 0 \leq \rho < 1.$$

1. Calculer la fonction caractéristique de $\mathbf{X}$.
2. Calculer la probabilité que $X_1 > \rho + a$ sachant que $X_2 = 1$.
3. On considère le probléme de classification avec deux populations : $a = 0$ pour la popultation π_1 et $a = 1$ pour la population π_2 (sans coûts de mauvaise classification et sans a priori). Dans quelle population classer l'observation $\mathbf{x} = (1, -1)'$?

Question 5 (2 points)

Soit un échantillon i.i.d. $\mathbf{X}_1, \ldots, \mathbf{X}_n$ de vecteurs aléatoires de loi Fisher-von Mises(FvM) (paramètre de position $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$ et de concentration $\kappa > 0$). Pour rappel, un vecteur aléatoire $\mathbf{X}$ de loi FvM$(\boldsymbol{\theta}, \kappa)$ admet une densité de la forme $\mathbf{x} \to c_\kappa \exp(\kappa \mathbf{x}'\boldsymbol{\theta})$ (par rapport à la mesure de Lebesgue sur $\mathcal{S}^{p-1}$), où c_κ est une constante de normalisation. Soit

$$\hat{\boldsymbol{\theta}} = \frac{\bar{\mathbf{X}}}{\|\bar{\mathbf{X}}\|},$$

où $\bar{\mathbf{X}} := n^{-1} \sum_{i=1}^n \mathbf{X}_i$. Obtenir la loi asymptotique de $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$.