École Normale Supérieure
FIMFA — Statistique 2016–2017

# Partiel

1. Pour $\theta$ in $\mathbb{R}$, la loi $P_\theta$ est définie par sa densité $p_\theta(x) = \frac{1}{2}\exp(-|x-\theta|)$.
   (a) (1 point) Calculer l'espérance et la variance de $P_\theta$.
   (b) (1 point) Proposer un estimateur de $\theta$ fondé sur la méthode des moments. Etudier sa consistance, sa normalité asymptotique.
   (c) (1 point) La fonction score notée $\dot{\ell}_n$ est définie comme la dérivée par rapport à $\theta$ de la log vraisemblance quand cette dérivée est bien définie, 0 dans les autres cas. Si elles sont bien définies, que valent $\mathbb{E}_\theta[\dot{\ell}_n(\theta)]$ et $\mathbb{E}_\theta[\dot{\ell}_n(\theta)^2]$ ?
   (d) (2 points) L'estimateur au maximum de vraisemblance est-il bien défini dans ce modèle ? Si oui, étudier sa consistance et éventuellement sa normalité asymptotique.
   (e) (1 point) On veut tester $H_0 : \theta \leq 0$ contre $H_1 : \theta > \delta > 0$. Proposer un test de niveau asymptotique $\alpha > 0$ pour séparer ces deux hypothèses.
2. Dans un modèle exponentiel minimal, on considère deux paramètres $\theta$ et $\theta_n = \theta + h/\sqrt{n}$ où $h$ verifie $\sum_{i=1}^d h_i = 0$.
   (a) (1 point) Calculer la distance de Hellinger entre $P_\theta$ et $P_{\theta_n}$. Donner un équivalent de $H^2(P_\theta, P_{\theta_n})$ lorsque $n \to \infty$.
   (b) (2 points) Quelle est la loi limite de $\ell_n(\theta_n) - \ell_n(\theta)$ sous $P_\theta$, sous $P_{\theta_n}$ ?
   (c) (2 points) Quelle est la puissance optimale limite de tests de niveau $\alpha$ entre $P_\theta$ et $P_{\theta_n}$ ?
   (d) (1 point) Sous $P_\theta$, quelle est la limite en loi de $I(\theta)^{-1}\frac{1}{\sqrt{n}}\nabla\ell_n(\theta) + \sqrt{n}(\widehat{\theta}_n - \theta)$ ?
3. Modèles de mélange.
   On observe $n$ tirages indépendants $X_1, \ldots, X_n$ selon une loi inconnue $P$ sur $\{0, \ldots, d\}$.
   La loi $P$ est un mélange de deux lois binomiales de paramètres $(d, q_0)$ et $(d, q_1)$ avec $q_0 < q_1$ dans des proportions $\pi, 1-\pi$ avec $\pi \in ]0,1[$.
   (a) (1 point) Dans le modèle de mélange, calculer, $\mathbb{E}X, \mathbb{E}X^2, \mathbb{E}X^3$ où $X \sim P$.
   (b) (½ point) Le modèle de mélange est-il identifiable ?
   (c) (1 point) Dans le modèle de mélange proposer une suite d'estimateurs consistants de $q_0, q_1, \pi$.
   (d) (1 point) Etudier l'éventuelle normalité asymptotique de l'estimateur.
   (e) (1 point) Calculer l'information de Fisher dans le modèle de mélange.
   (f) (1 point) Comparaison de la variance de l'estimateur et de l'inverse de l'information de Fisher (vous pouvez comparer les matrices de covariance au sens de l'ordre semi-défini positif ou vous intéresser aux variances des estimateurs de $\pi, q_0, q_1$).
   (g) (1 point) Proposer un estimateur facile à calculer et dont la variance asymptotique coïncide avec l'inverse de l'information de Fisher.
4. On dispose de deux échantillons i.i.d. collectés indépendamment l'un de l'autre $X_1, \ldots, X_n$ et $Y_1, ..., Y_m$. Les $X_i$ sont distribués selon $\mathcal{N}(\mu_1, \sigma_1^2)$, les $Y_i$ sont distribués selon $\mathcal{N}(\mu_2, \sigma_2^2)$.
   (a) (1 point) Proposer un intervalle de confiance de niveau $\alpha \in (0,1)$ pour le rapport $\sigma_2^2/\sigma_1^2$.
   (b) (1 point) Proposer un test de niveau $\alpha$ pour séparer $H_O : \sigma_1^2 = \sigma_2^2$ contre $H_1 : \sigma_2^2/\sigma_1^2 > 1$.
   (c) (1 point) Si on suppose $\sigma_1 = \sigma_2$, proposer un intervalle de confiance de niveau $1-\alpha$, pour $\mu_1 - \mu_2$.
   (d) (1 point) Si on suppose $\sigma_1 = \rho\sigma_2$ ($\rho$ connu), proposer un intervalle de confiance de niveau $1-\alpha$, pour $\mu_1 - \mu_2$.
5. La loi $P_{\gamma,\mu,\sigma}$ (avec $\gamma, \sigma > 0$) est la loi de $\gamma + \exp(\sigma X + \mu)$ où $X \sim \mathcal{N}(0,1)$. On dispose d'un échantillon i.i.d. $X_1, \ldots, X_n$ de $P_{\gamma,\mu,\sigma}$.
   (a) (1 point) Calculer, espérance, variance, médiane de $P_{\gamma,\mu,\sigma}$.
   (b) (1 point) Proposer un estimateur des moments pour $(\gamma, \mu, \sigma)$, étudier sa consistance, et éventuellement sa normalité asymptotique.
   (c) (2 points) La maximisation de la vraisemblance fournit elle une méthode d'estimation consistante dans ce modèle ?

| Question : | 1 | 2 | 3 | 4 | 5 | Total |
|---|---|---|---|---|---|---|
| Points : | 6 | 6 | 6½ | 4 | 4 | 26½ |
| Score : | | | | | | |