EXAMEN STATISTIQUE - 1MF2E

Lundi 24 mars 2025 (8h-9h45)

Partiel sans document (Une feuille A4 recto-verso autorisée)

Exercice 1 : Tests Statistiques (10 **points)**

On considère n observations $x_1, ..., x_n$ issues d'un vecteur $(X_1, ..., X_n)$ de n variables aléatoires indépendantes de mêmes lois de densités

$$f(x_i; a) = \begin{cases} (a-1)x_i^{-a} & \text{si } x_i > 1 \\ 0 \text{ sinon} \end{cases}$$

On désire utiliser les observations $x_1, ..., x_n$ pour déterminer si $a = a_0 > 1$ ou si $a = a_1 > a_0$. On considère donc le test d'hypothèses

$$H_0 : a = a_0, \quad H_1 : a = a_1 \qquad \text{avec } a_1 > a_0 > 1.$$

1. (2pt) Déterminer la statistique T_n du test de Neyman Pearson et la région critique associée. En admettant que $E[\ln X_i] = \frac{1}{a-1}$, vérifier que la région critique du test est en accord avec cette espérance mathématique.

2. (1pt) En admettant que $Y_i = \ln X_i$ suit une loi gamma de paramètres 1 et $a-1$, i.e., $Y_i \sim \mathcal{G}(1, a-1)$, déterminer la loi approchée de T_n sous les deux hypothèses H_0 et H_1 provenant du théorème central limite.

3. (2pts) On note G la fonction de répartition d'une loi du normale $\mathcal{N}(0,1)$. Exprimer le risque de première espèce α en fonction du seuil du test de Neyman Pearson noté S_α, de G, n et de a_0. En déduire la valeur du seuil S_α en fonction de $G^{-1}(\alpha)$ et de n et a_0.

4. (2pts) Déterminer les caractéristiques opérationnelles du récepteur (courbes COR) pour ce test et montrer qu'elles ne dépendent que de n et de $r(a_0, a_1) = \frac{a_1-1}{a_0-1}$ (et bien sûr des fonctions G et G^{-1}). Analyser les performances du test en fonction de $r(a_0, a_1)$ et tracer l'allure approximative des courbes COR pour différentes valeurs de a_1 lorsque $a_0 = 2$.

5. (3pts) On désire vérifier que les observations suivent la loi de densité $f(x_i; a)$ avec $a = 2$ à l'aide d'un test de Kolmogorov.

   - Déterminer la fonction de répartition de la loi de densité $f(x_i; a)$ notée F lorsque $a = 2$.
   - On observe l'échantillon de taille $n = 4$ suivant : $x_1 = 2$, $x_2 = 3$, $x_3 = 4$ et $x_4 = \frac{3}{2}$. Le tableau suivant résume les quantités nécessaires pour effectuer le test

$x_{(i)}$	2	2.5	3	4
$F\left(x_{(i)}\right)$	0.5000	0.6000	0.6667	0.7500
$\hat{F}\left(x_{(i)}^-\right)$	0	0.25	0.50	0.75
$\hat{F}\left(x_{(i)}^+\right)$	0.25	0.50	0.75	1
$E_i^- = \lvert F\left(x_{(i)}\right) - \hat{F}\left(x_{(i)}^-\right)\rvert$	0.5000	0.3500	0.1667	0.0000
$E_i^+ = \lvert F\left(x_{(i)}\right) - \hat{F}\left(x_{(i)}^+\right)\rvert$	0.2500	0.1000	0.0833	0.2500

     où $(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)})$ est l'échantillon ordonné.
     Expliquer ce que représentent $\hat{F}\left(x_{(i)}^-\right)$ et $\hat{F}\left(x_{(i)}^+\right)$.
   - Rappeler la région critique du test de Kolmogorov. Pour $\alpha = 0.01$ et $n = 4$, on a $S_{0.01} = 0.7342$. Que peut-on en conclure ?

Exercice 2 : Estimation (10 **points)**

On considère n observations $x_1, ..., x_n$ issues d'un vecteur $(X_1, ..., X_n)$ de n variables aléatoires indépendantes de même loi de densité

$$f(x_i; a) = \begin{cases} (a-1)x_i^{-a} & \text{si } \ x_i > 1 \\ 0 \text{ sinon} \end{cases}$$

avec $a > 1$.

1. (1pt) Montrer que l'estimateur du vraisemblance du paramètre a noté $\widehat{a}_{\text{MV}}$ est défini par

$$\widehat{a}_{\text{MV}} = 1 + \frac{n}{\sum_{i=1}^n \ln X_i}.$$

2. (2pts) On rappelle que si Y est une variable aléatoire réelle de loi gamma de paramètres ν et θ, notée $Y \sim \mathcal{G}(\nu, \theta)$, alors $Z = \frac{1}{Y}$ suit une loi inverse-gamma de paramètres ν et θ, i.e., $Z \sim \mathcal{IG}(\nu, \theta)$. Montrer que la variable aléatoire $Y_i = \ln X_i$ suit une loi gamma dont on déterminera les paramètres. Quelle est la loi de $Y = \sum_{i=1}^n \ln X_i$? En déduire que $\widehat{a}_{\text{MV}} = 1 + nZ$ où Z est une variable aléatoire de loi inverse gamma $\mathcal{IG}(n, a-1)$.

3. (3pts) Déterminer le bias de l'estimateur $\widehat{a}_{\text{MV}}$ et en déduire un estimateur sans biais du paramètre a noté $\widetilde{a}_{\text{MV}}$. Déterminer la variance de $\widetilde{a}_{\text{MV}}$ et en déduire que cet estimateur est convergent.

4. (1pt) Déterminer la borne de Cramer-Rao pour un estimateur non biaisé du paramètre a. L'estimateur $\widetilde{a}_{\text{MV}}$ est-il l'estimateur efficace du paramètre a ?

5. (3pts) On suppose désormais que le paramètre a est muni d'une loi a priori de densité

$$f(a) = \begin{cases} e^{1-a} & \text{si } \ a > 1 \\ 0 \text{ sinon} \end{cases}$$

   Déterminer l'estimateur du maximum a posteriori du paramètre a noté $\widehat{a}_{\text{MAP}}$. À l'aide de la densité de la variable $W = a|x_1, ..., x_n$ (a sachant $x_1, ..., x_n$), connue à une constante multiplicative près, montrer que la loi de $U = W - 1$ est une loi gamma dont on déterminera les paramètres (on remarquera que la densité de W est non nulle sur l'intervalle $]1, \infty[$). En déduire l'estimateur MMSE de a noté $\widehat{a}_{\text{MMSE}}$.

LOIS DE PROBABILITÉ DISCRÈTES

$p_k = P[X = k] \qquad p_{1,\ldots,m} = P[X_1 = k_1, \ldots, X_m = k_m]$

LOI	**Probabilités**	Moyenne	Variance	Fonction Caractéristique
Uniforme	$p_k = \frac{1}{n}$ $k \in \{1, \ldots, n\}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\dfrac{e^{it}\left(1 - e^{itn}\right)}{n\left(1 - e^{it}\right)}$
Bernoulli	$p_1 = P[X = 1] = p$ $p_0 = P[X = 0] = q$ $p \in [0, 1] \quad q = 1 - p$	p	pq	$pe^{it} + q$
Binomiale $B(n, p)$	$p_k = \binom{n}{k} p^k q^{n-k}$ $p \in [0, 1] \quad q = 1 - p$ $k \in \{0, 1, \ldots, n\}, \binom{n}{k} = \frac{n!}{k!(n-k)!}$	np	npq	$\left(pe^{it} + q\right)^n$
Binomiale négative	$p_k = \binom{n+k-1}{n-1} p^n q^k$ $p \in [0, 1] \quad q = 1 - p$ $k \in \mathbb{N}, \binom{n}{k} = \frac{n!}{k!(n-k)!}$	$n\dfrac{q}{p}$	$n\dfrac{q}{p^2}$	$\left(\dfrac{p}{1 - qe^{it}}\right)^n$
Multinomiale	$p_{1,\ldots,m} = \frac{n!}{k_1! \ldots k_m!} p_1^{k_1} \ldots p_m^{k_m}$ $p_j \in [0, 1] \quad q_j = 1 - p_j$ $k_j \in \{0, 1, \ldots, n\}$ $\sum_{j=1}^m k_j = n \quad \sum_{j=1}^m p_j = 1$	np_j	Variance : $np_j q_j$ Covariance : $-np_j p_k$	$\left(\sum_{j=1}^m p_j e^{it}\right)^n$
Poisson $P(\lambda)$	$p_k = e^{-\lambda} \dfrac{\lambda^k}{k!}$ $\lambda > 0 \quad k \in \mathbb{N}$	λ	λ	$\exp\left[\lambda\left(e^{it} - 1\right)\right]$
Géometrique	$p_k = pq^{k-1}$ $p \in [0, 1] \quad q = 1 - p$ $k \in \mathbb{N}^*$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	$\dfrac{pe^{it}}{1 - qe^{it}}$

LOIS DE PROBABILITÉ CONTINUES

LOI	**Densité de probabilité**	Moyenne	Variance	Fonction Caractéristique
Uniforme	$f(x) = \frac{1}{b-a}$ $x \in]a,b[$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\dfrac{e^{itb} - e^{ita}}{it(b-a)}$
Gamma $\mathcal{G}(\nu,\theta)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\theta x} x^{\nu-1}$ $\theta > 0,\ \nu > 0$ $x \geq 0$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\dfrac{\nu}{\theta}$	$\dfrac{\nu}{\theta^2}$	$\dfrac{1}{\left(1 - i\frac{t}{\theta}\right)^\nu}$
Inverse gamma $\mathcal{IG}(\nu,\theta)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\frac{\theta}{x}} \frac{1}{x^{\nu+1}}$ $\theta > 0,\ \nu > 0$ $x \geq 0$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\frac{\theta}{\nu-1}$ si $\nu > 1$	$\frac{\theta^2}{(\nu-1)^2(\nu-2)}$ si $\nu > 2$	(*)
Première loi de Laplace	$f(x) = \frac{1}{2} e^{-\lvert x \rvert}, \quad x \in \mathbb{R}$	0	2	$\dfrac{1}{1+t^2}$
Normale univariée $\mathcal{N}\left(m,\sigma^2\right)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Normale multivariée $\mathcal{N}_p(\boldsymbol{m},\boldsymbol{\Sigma})$	$f(x) = K e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{m})}$ $K = \frac{1}{\sqrt{(2\pi)^p \det(\Sigma)}}$ $x \in \mathbb{R}^p$	$\boldsymbol{m}$	$\boldsymbol{\Sigma}$	$e^{i\boldsymbol{u}^T\boldsymbol{m} - \frac{1}{2}\boldsymbol{u}^T\boldsymbol{\Sigma}\boldsymbol{u}}$
Khi2 χ^2_ν $\Gamma\left(\frac{\nu}{2},\frac{1}{2}\right)$	$f(x) = k e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}$ $k = \dfrac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)}$ $\nu \in \mathbb{N}^*,\ x \geq 0$	ν	2ν	$\dfrac{1}{(1-2it)^{\frac{\nu}{2}}}$
Cauchy $c_{\lambda,\alpha}$	$f(x) = \dfrac{1}{\pi\lambda\left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0,\ \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda\lvert t \rvert}$
Beta $B(a,b)$	$f(x) = k x^{a-1}(1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0$ $x \in]0,1[$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)