EXAMEN STATISTIQUE - 1SN

Mardi 4 Décembre 2018

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

Exercice 1 : Estimation (10 points)

On considère n observations $x_1, ..., x_n$ issues d'un échantillon $(X_1, ..., X_n)$ distribué suivant la même loi de densité

$$f(x_i;\theta) = \frac{1}{\theta^p \Gamma(p) x_i^{p+1}} \exp\left(-\frac{1}{\theta x_i}\right) \mathbb{I}_{\mathbb{R}^+}(x_i),$$

où $\Gamma(.)$ est la fonction gamma et $\mathbb{I}_{\mathbb{R}^+}(x)$ est la fonction indicatrice sur $\mathbb{R}^+$. On cherche à estimer le paramètre θ à partir des observations $x_1, ..., x_n$ (on supposera que p est connu dans les 4 premières questions de cet exercice).

1. Montrer que la vraisemblance de $(x_1, ..., x_n)$ admet un unique maximum global pour une valeur de θ que l'on déterminera. En déduire l'estimateur du maximum de vraisemblance du paramètre θ noté $\widehat{\theta}_{\text{MV}}$.

2. Montrer que $Y_i = \frac{1}{X_i}$ suit une loi gamma $\Gamma\left(\frac{1}{\theta}, p\right)$. On notera qu'on peut à l'aide des tables déterminer la moyenne et la variance de la variable aléatoire Y_i.

3. L'estimateur $\widehat{\theta}_{\text{MV}}$ est-il un estimateur sans biais et convergent du paramètre θ ?

4. Déterminer la borne de Cramer-Rao pour un estimateur non biaisé du paramètre θ. L'estimateur $\widehat{\theta}_{\text{MV}}$ est-il l'estimateur efficace du paramètre θ ?

5. On suppose désormais que p est inconnu. Expliquer pourquoi il semble difficile d'estimer les deux paramètres p et θ à l'aide de la méthode du maximum de vraisemblance. Au vu de cette difficulté, on décide d'appliquer la méthode des moments aux variables aléatoires $Y_1, ..., Y_n$. Déterminer des estimateurs des moments des paramètres p et θ notés $\hat{p}_{\text{MO}}$ et $\hat{\theta}_{\text{MO}}$. Que pensez vous des propriétés de ces estimateurs $\hat{p}_{\text{MO}}$ et $\hat{\theta}_{\text{MO}}$?

Exercice 2 : Test Statistique (10 points)

On considère n variables aléatoires indépendantes $X_1, ..., X_n$ suivant la même loi normale $\mathcal{N}(\mu, \sigma_1^2)$ et m variables aléatoires $Y_1, ..., Y_m$ suivant la même loi normale $\mathcal{N}(\mu, \sigma_2^2)$, où $\mu \in \mathbb{R}$ est un paramètre inconnu et σ_1^2, σ_2^2 sont deux paramètres connus. On supposera de plus que les vecteurs $(X_1, ..., X_n)^T$ et $(Y_1, ..., Y_m)^T$ sont indépendants. Une situation pratique dans laquelle on peut avoir ces deux ensembles d'observations $(X_1, ..., X_n)^T$ et $(Y_1, ..., Y_m)^T$ correspond par exemple au cas où $(X_1, ..., X_n)^T$ sont obtenues par un premier capteur dont l'incertitude dépend de la variance σ_1^2 et où $(Y_1, ..., Y_m)^T$ sont obtenues par un second capteur dont l'incertitude dépend de la variance σ_2^2. L'objectif de cet exercice est d'étudier un test statistique basé sur les vecteurs $(X_1, ..., X_n)^T$ et $(Y_1, ..., Y_m)^T$ qui permet de déterminer si $\mu = 0$ ou si $\mu = \mu_1 < 0$. On considère donc le test d'hypothèses

$$H_0 : \mu = 0 \text{ contre } H_1 : \mu = \mu_1 \qquad \text{avec } \mu_1 < 0.$$

1. Déterminer la vraisemblance conjointe de $X_1, ..., X_n$ et $Y_1, ..., Y_m$ notée $L(\boldsymbol{x}, \boldsymbol{y}; \mu)$ avec $\boldsymbol{x} = (x_1, ..., x_n)$ et $\boldsymbol{y} = (y_1, ..., y_m)$.

2. En utilisant l'expression de $L(\boldsymbol{x}, \boldsymbol{y}; \mu)$, montrer que la statistique de test du théorème de Neyman-Pearson est
$$T_n = \frac{1}{\sigma_1^2} \sum_{i=1}^{n} X_i + \frac{1}{\sigma_2^2} \sum_{i=1}^{m} Y_i$$
et indiquer la région critique de ce test.

3. Déterminer la loi de T_n sous les deux hypothèses H_0 et H_1.

4. Exprimer les risques de première et seconde espèce α et β en fonction du seuil du test de Neyman-Pearson, de la fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0, 1)$ notée F et des paramètres n, m, σ_1^2 et σ_2^2.

5. Déterminer les caractéristiques opérationnelles du récepteur (courbes COR) pour ce test et analyser les performances du test en fonction de la valeur de μ_1 et des autres paramètres $n, m, \sigma_1^2, \sigma_2^2$.

LOIS DE PROBABILITÉ CONTINUES

m : moyenne σ^2 : variance **F. C. :** fonction caractéristique

LOI	**Densité de probabilité**	**m**	σ^2	**F. C.**
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 $\Gamma(\theta, \nu)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\theta x} x^{\nu-1}$ $\theta > 0,\ \nu > 0$ $x \geq 0$	$\dfrac{\nu}{\theta}$	$\dfrac{\nu}{\theta^2}$	$\dfrac{1}{\left(1 - i\frac{t}{\theta}\right)^\nu}$
Inverse gamma IG(θ, ν)	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\frac{\theta}{x}} \frac{1}{x^{\nu+1}}$ $\theta > 0,\ \nu > 0$ $x \geq 0$	$\frac{\theta}{\nu-1}$ si $\nu > 1$	$\frac{\theta^2}{(\nu-1)^2(\nu-2)}$ si $\nu > 2$	(*)
Première loi de Laplace	$f(x) = \frac{1}{2} e^{-\|x\|}$	0	2	$\dfrac{1}{1+t^2}$
Normale $\mathcal{N}(m, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Khi$_2$ χ^2_ν $\Gamma\left(\frac{1}{2}, \frac{\nu}{2}\right)$	$f(x) = k e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}$ $k = \dfrac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}$ $\nu \in \mathbb{N}^*,\ x \geq 0$	ν	2ν	$\dfrac{1}{(1-2it)^{\frac{\nu}{2}}}$
Cauchy $c_{\lambda,\alpha}$	$f(x) = \dfrac{1}{\pi\lambda\left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0,\ \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda\|t\|}$
Beta	$f(x) = k x^{a-1} (1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0,\ x \in]0, 1[$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)

LOIS DE PROBABILITÉ DISCRÈTES

$\mathbf{m}$: moyenne σ^2 : variance **F. C. :** fonction caractéristique

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

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