EXAMEN STATISTIQUE - 1TR

Mardi 8 Décembre 2015

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

Exercice 1 : Estimation (11 points)

On considère n variables aléatoires $X_1, ..., X_n$ indépendantes suivant la même loi continue de densité

$$p(x_i; r, \alpha) = \begin{cases} \frac{r\alpha^r}{(\alpha+x_i)^{r+1}} & \text{si } x_i > 0 \\ 0 \text{ sinon} \end{cases}$$

avec $\alpha > 0$ et $r > 2$. On admettra qu'une telle loi est de moyenne $E[X_i] = \frac{\alpha}{r-1}$ et de variance $\text{Var}[X_i] = \frac{2\alpha^2}{(r-1)(r-2)}$. On suppose que le paramètre α est connu et on cherche à étudier des estimateurs du paramètre r.

1. Montrer que la vraisemblance associée à l'échantillon $(X_1, ..., X_n)$ admet un unique maximum global pour une valeur de r que l'on déterminera. En déduire l'estimateur du maximum de vraisemblance de r noté $\hat{r}_{\text{MV}}$ en fonction de n, α et de l'échantillon $(X_1, ..., X_n)$.

2. En utilisant l'expression de $E[X_i]$, déterminer un estimateur des moments du paramètre r noté $\hat{r}_{\text{Mo}}$.

3. Montrer que si X_i possède la densité $p(x_i; r, \alpha)$ définie ci-dessus, alors $V_i = \ln\left(\frac{\alpha+X_i}{\alpha}\right)$ suit une loi exponentielle de densité

   $$g(v_i; r) = \begin{cases} re^{-rv_i} & \text{si } v_i > 0 \\ 0 \text{ sinon} \end{cases}$$

   En utilisant les tables de lois, déterminer la fonction caractéristique de V_i et montrer que $V = \sum_{i=1}^n V_i$ suit une loi gamma $\Gamma(r, n)$. En déduire que $\hat{r}_{\text{MV}}$ suit une loi inverse gamma $\text{IG}(\theta, \nu)$ de paramètres $\theta = nr$ et $\nu = n$.

4. En utilisant le fait que $\hat{r}_{\text{MV}}$ suit une loi inverse gamma $\text{IG}(\theta, \nu)$ de paramètres $\theta = nr$ et $\nu = n$, montrer que $\hat{r}_{\text{MV}}$ est un estimateur biaisé du paramètre r et en déduire un estimateur non biaisé de r noté $\hat{r}^*$. Déterminer la variance de $\hat{r}^*$ et en déduire que cet estimateur est convergent.

5. Déterminer la borne de Cramer-Rao d'un estimateur non biaisé de r. L'estimateur $\hat{r}^*$ est-il l'estimateur efficace du paramètre r ?

6. On suppose qu'on dispose d'une information a priori sur le paramètre r résumée dans une loi gamma $\Gamma(a, b)$. Montrer que la loi a posteriori de $r|x_1, ..., x_n$ est également une loi gamma dont on précisera les paramètres. Quel est l'estimateur du maximum a posteriori du paramètre r noté $\hat{r}_{\text{MAP}}$?

Exercice 2: Tests Statistiques (9 points)

Comme dans l'exercice précédent, on considère n variables aléatoires $X_1, ..., X_n$ indépendantes suivant la même loi continue de densité

$$p(x_i; r, \alpha) = \begin{cases} \frac{r\alpha^r}{(\alpha+x_i)^{r+1}} & \text{si } x_i > 0 \\ 0 \text{ sinon} \end{cases}$$

avec $\alpha > 0$ et $r > 2$. Dans un premier temps, on cherche à effectuer le test d'hypothèses simples

$$\begin{aligned} H_0 &: r = r_0 \\ H_1 &: r = r_1 \end{aligned}$$

avec $r_1 > r_0 > 2$.

1. Montrer que le test de Neyman Pearson conduit à la statistique de test

$$T_n = \sum_{i=1}^{n} \ln\left(\frac{\alpha + X_i}{\alpha}\right)$$

   et indiquer la région critique de ce test.

2. On admet que T_n suit une loi gamma $\Gamma(r_j, n)$ sous l'hypothèse H_j avec $j \in \{0, 1\}$, et on suppose qu'on sait calculer la fonction Q_n définie par

$$Q_n(x) = \frac{1}{\Gamma(n)} \int_0^x u^{n-1} e^{-u}\, du$$

   ainsi que son inverse $Q_n^{-1}(x)$ pour toutes les valeurs de n et de x. Déterminer la valeur du seuil K_α du test de Neyman Pearson en fonction de r_0, α et de Q_n^{-1}.

3. Déterminer la puissance du test en fonction de K_α, r_1 et Q_n. En déduire les courbes COR du test étudié dans cet exercice et tracer la forme de ces courbes pour différentes valeurs du couple (r_0, r_1).

4. Lorsqu'on ne peut pas calculer les fonctions Q_n et Q_n^{-1}, on peut approcher la loi de T_n par une loi normale d'après le théorème de la limite centrale. Calculer le seuil K_α et la puissance du test lorsqu'on utilise cette loi approchée en fonction de n, r_0, r_1 et de la fonction de répartition de la loi normale $\mathcal{N}(0, 1)$. En déduire l'expression des courbes COR résultant de ces approximations.

LOIS DE PROBABILITÉ CONTINUES

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

LOI	**Densité de probabilité**	$\mathbf{m}$	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	$f(x) = \frac{1}{2} e^{-\lvert x \rvert}$	0	2	$\dfrac{1}{1+t^2}$
Normale $\mathcal{N}\left(m, \sigma^2\right)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$	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 $c_{\lambda, \alpha}$	$f(x) = \dfrac{1}{\pi\lambda\left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0,\ \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda \lvert t \rvert}$
Beta	$f(x) = k x^{a-1} (1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0,\ x \in]0, 1[$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)