EXAMEN STATISTIQUE - 2ÈME ANNÉE APPRENTISSAGE

Lundi 18 Novembre 2024 (14h-15h30)

Partiel sans documents autorisés, avec feuille A4 recto-verso de notes personnelles.

Exercice 1 : Estimation (10 points)

On considère n observations $x_1, ..., x_n$ issues d'un vecteur de n variables aléatoires X_i indépendantes de lois Beta de paramètres $B\left(\frac{1}{\theta}, 1\right)$ de densités

$$p(x_i;\theta) = \begin{cases} \frac{1}{\theta} x_i^{\frac{1}{\theta}-1} \text{ si } x_i \in]0,1[\\ 0 \text{ sinon} \end{cases}$$

avec $\theta > 0$ un paramètre inconnu.

1. Montrer que l'estimateur du maximum de vraisemblance du paramètre θ noté $\widehat{\theta}_{\text{MV}}$ est défini par
$$\widehat{\theta}_{\text{MV}} = -\frac{1}{n}\sum_{i=1}^{n} \ln(X_i).$$

2. Montrer que la variable aléatoire $Y_i = -\ln(X_i)$ suit une loi gamma $\Gamma\left(1, \frac{1}{\theta}\right)$. En déduire la moyenne et la variance de la variable Y_i notées $E[Y_i]$ et var$[Y_i]$.

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

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. Montrer que l'estimateur des moments de θ défini à partir de $E[X_i]$ est
$$\widehat{\theta}_{\text{Mo}} = \frac{n}{\sum_{i=1}^{n} X_i} - 1.$$

6. Lequel des deux estimateurs $\widehat{\theta}_{\text{Mo}}$ et $\widehat{\theta}_{\text{MV}}$ choisiriez vous (justifier votre réponse) ?

Exercice 2 : 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 lois de Poisson de paramètre $i\lambda$, c'est-à-dire, telles que

$$P[X_i = x_i; \lambda] = \frac{(i\lambda)^{x_i}}{x_i!} e^{-i\lambda},\ x_i \in \mathbb{N}.$$

avec $\lambda > 0$. On notera que le paramètre de la loi de Poisson pour la variable aléatoire X_i dépend de l'indice i. On désire utiliser les observations $x_1, ..., x_n$ pour déterminer si $\lambda = \lambda_0 > 0$ ou si $\lambda = \lambda_1 \in]0, \lambda_0[$. On considère donc le test d'hypothèses

$$H_0 : \lambda = \lambda_0, \quad H_1 : \lambda = \lambda_1 \qquad \text{avec } 0 < \lambda_1 < \lambda_0.$$

1. Montrer que la statistique du test de Neyman Pearson est $T_n = \sum_{i=1}^n X_i$ et que la région critique associée est $\{(x_1, ..., x_n) \in \mathbb{N}^n | T_n < S_\alpha\}$.

2. On donne la relation $\sum_{i=1}^n i = \frac{n(n+1)}{2}$. Montrer que la loi approchée de T_n issue du théorème central limite est la loi normale $\mathcal{N}\left(\frac{n(n+1)}{2}\lambda, \frac{n(n+1)}{2}\lambda\right)$. On supposera dans la suite de cet exercice qu'on peut approcher la loi de T_n par cette loi normale.

3. On note F 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 $F(\alpha)$, n et de λ_0. En déduire la valeur de S_α en fonction de $F^{-1}(\alpha)$, n et λ_0.

4. Déterminer les caractéristiques opérationnelles du récepteur (courbes COR) pour ce test en fonction de n, $F^{-1}(\alpha)$, λ_0 et λ_1. Représenter l'allure de ces courbes COR pour diverses valeurs de n. Les performances du test seront-elles meilleures pour $(\lambda_0, \lambda_1) = (10, 1)$ ou pour $(\lambda_0, \lambda_1) = (1000, 100)$?

LOIS DE PROBABILITÉ CONTINUES

$\mathbf{m}$: 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 $\mathcal{G}(\nu, \theta)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\theta x} x^{\nu-1}$ $\theta > 0,\ \nu > 0$ $x \geq 0$ 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 $\mathcal{N}(m, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Normale multivariée $\mathcal{N}_p(\boldsymbol{m}, \boldsymbol{\Sigma})$	$f(x) = K e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{m})}$ $K = \frac{1}{\sqrt{(2\pi)^p \det(\Sigma)}}$ $x \in \mathbb{R}^p$	$\boldsymbol{m}$	$\boldsymbol{\Sigma}$	$e^{i\boldsymbol{u}^T\boldsymbol{m} - \frac{1}{2}\boldsymbol{u}^T\boldsymbol{\Sigma}\boldsymbol{u}}$
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 $B(a, b)$	$f(x) = k x^{a-1} (1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0$ $x \in]0, 1[$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)

LOIS DE PROBABILITÉ DISCRÈTES

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

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

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