
EXAMEN STATISTIQUE - 2ÈME ANNÉE APPRENTISSAGE

Mardi 21 Novembre 2023 (14h-15h30)

Partiel avec documents autorisés

Exercice 1 : Estimation (10 points)

On considère n observations x_1, \dots, x_n issues d'un vecteur de n variables aléatoires X_i indépendantes de lois de Poisson de paramètre $i\lambda$, c'est-à-dire, telles que

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

avec $\lambda > 0$ un paramètre inconnu. On notera que le paramètre de la loi de Poisson pour la variable aléatoire X_i dépend de l'indice i .

1. On cherche à estimer le paramètre λ à l'aide de l'estimateur

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n \frac{X_i}{i}.$$

Déterminer le biais de cet estimateur. En utilisant la propriété $\sum_{i=1}^n \frac{1}{i} \leq n$, donner un majorant de $\text{var}(\hat{\lambda})$. En déduire que $\hat{\lambda}$ est un estimateur convergent de λ .

2. En utilisant la relation $\sum_{i=1}^n i = \frac{n(n+1)}{2}$, montrer que l'estimateur du maximum de vraisemblance du paramètre λ noté $\hat{\lambda}_{\text{MV}}$ est défini par

$$\hat{\lambda}_{\text{MV}} = \frac{2}{n(n+1)} \sum_{i=1}^n X_i.$$

Cet estimateur est-il sans biais et convergent ?

3. Déterminer la borne de Cramer-Rao pour un estimateur non biaisé du paramètre λ . L'estimateur $\hat{\lambda}_{\text{MV}}$ est-il l'estimateur efficace du paramètre λ ?
4. Lequel des deux estimateurs $\hat{\lambda}$ et $\hat{\lambda}_{\text{MV}}$ choisiriez vous (justifier votre réponse) ?

Exercice 2 : Tests Statistiques (10 points)

On considère n observations x_1, \dots, x_n issues d'un vecteur (X_1, \dots, X_n) de n variables aléatoires indépendantes de même loi de densités

$$f(x_i; a) = \begin{cases} ax_i^{a-1} & \text{si } x_i \in]0, 1[\\ 0 & \text{sinon} \end{cases}$$

avec $a > 0$. On désire utiliser les observations x_1, \dots, x_n pour déterminer si $a = a_0 > 0$ ou si $a = a_1 > a_0$. On considère donc le test d'hypothèses

$$H_0 : a = a_0, \quad H_1 : a = a_1 \quad \text{avec } a_1 > a_0 > 0.$$

1. Montrer que la variable aléatoire $U_i = -\ln(X_i)$ possède la loi gamma $\mathcal{G}(1, a)$ et en déduire à l'aide des tables sa moyenne $E[U_i]$ et sa variance $\text{var}[U_i]$.
Remarque : on prendra soin de déterminer le domaine de définition de U_i .
2. Montrer que la statistique du test de Neyman Pearson est $T_n = \sum_{i=1}^n U_i$ et déterminer la région critique associée.
3. Montrer que la loi approchée de T_n issue du théorème central limite est la loi normale $\mathcal{N}\left(\frac{n}{a}, \frac{n}{a^2}\right)$. On supposera dans la suite de cet exercice qu'on peut approcher la loi de T_n par cette loi normale.
4. 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 a_0 . En déduire la valeur du seuil S_α en fonction de $F^{-1}(\alpha)$ et de n et a_0 .
5. Déterminer les caractéristiques opérationnelles du récepteur (courbes COR) pour ce test et montrer qu'elles ne dépendent que de n , $F^{-1}(\alpha)$ et de $\frac{a_1}{a_0}$. Analyser les performances du test en fonction de $\frac{a_1}{a_0}$ et de n et représenter l'allure de ces courbes pour diverses valeurs de $\frac{a_1}{a_0}$ et de n .

LOIS DE PROBABILITÉ CONTINUES

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

LOI	Densité de probabilité	m	σ^2	F. C.
Uniforme	$f(x) = \frac{1}{b-a}$ $x \in]a, b[$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{itb} - e^{ita}}{it(b-a)}$
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 $\Gamma(n+1) = n! \forall n \in \mathbb{N}$	$\frac{\nu}{\theta}$	$\frac{\nu}{\theta^2}$	$\frac{1}{(1 - i\frac{t}{\theta})^\nu}$
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	$f(x) = \frac{1}{2} e^{- x }, \quad x \in \mathbb{R}$	0	2	$\frac{1}{1+t^2}$
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 $\mathcal{N}_p(\mathbf{m}, \Sigma)$	$f(x) = K e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \Sigma^{-1}(\mathbf{x}-\mathbf{m})}$ $K = \frac{1}{\sqrt{(2\pi)^p \det(\Sigma)}}$ $x \in \mathbb{R}^p$	\mathbf{m}	Σ	$e^{i\mathbf{u}^T \mathbf{m} - \frac{1}{2} \mathbf{u}^T \Sigma \mathbf{u}}$
Khi ₂ χ_ν^2 $\Gamma(\frac{1}{2}, \frac{\nu}{2})$	$f(x) = k e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}$ $k = \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})}$ $\nu \in \mathbb{N}^*, x \geq 0$	ν	2ν	$\frac{1}{(1-2it)^{\frac{\nu}{2}}}$
Cauchy $c_{\lambda, \alpha}$	$f(x) = \frac{1}{\pi\lambda \left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0, \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda t }$
Beta $B(a, b)$	$f(x) = k x^{a-1} (1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0, b > 0$ $x \in]0, 1[$ 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

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

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

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