Statistique Mathématique

Partiel — 28 février 2022

Durée: 2 heures. Documents interdits. Justifiez bien vos réponses et soignez votre présentation. La partie II utilise des résultats de la partie I, que l'on pourra admettre si besoin est. Bon courage !

Estimation des paramètres d'une loi bêta

La *loi bêta* est une loi continue, à valeurs dans $[0,1]$. Elle est par exemple utilisée en génétique dans le modèle de Baldind-Nichols pour modéliser la fréquence d'apparition d'allèles dans des sous-populations. Dans ce sujet, nous nous proposons d'étudier succinctement la loi bêta et de proposer un estimateur pour les paramètres de cette loi.

Généralement, la loi bêta est paramétrisée par deux paramètres strictement positifs α et β. Sa densité par rapport à la mesure de Lebesgue est donnée par

$$\forall t \in \mathbb{R}, \qquad f_{\alpha,\beta}(t) = \frac{1}{B(\alpha,\beta)} t^{\alpha-1}(1-t)^{\beta-1}\mathbb{1}_{t\in[0,1]}\,,$$

où $B(\alpha,\beta)$ est une constante de normalisation.

Partie I : Préliminaires

Dans cette partie, on suppose que X est une variable aléatoire suivant la loi bêta de paramètres (α,β) (ce qu'on notera par la suite $X \sim B(\alpha,\beta)$). On note Γ la *fonction Gamma*, définie par

$$\forall z \in \mathbb{R}_+, \qquad \Gamma(z) = \int_0^{+\infty} x^{z-1}\mathrm{e}^{-x}\mathrm{d}x\,.$$

On pourra utiliser, sans les démontrer, les faits suivants : Γ est bien définie, $\Gamma(1)=1$, et vérifie la propriété remarquable $\Gamma(z+1) = z\Gamma(z)$ pour tout $z>0$.

I.1. On suppose dans cette question que α et β sont strictement supérieurs à 1. Quel est le mode (l'élément maximisant la densité, s'il existe) de $f_{\alpha,\beta}$?

I.2. Toujours en supposant que $\alpha,\beta>1$ et en utilisant la question I.1., tracez $f_{\alpha,\beta}(t)$ en fonction de t dans les trois cas suivants: $\alpha \gg \beta$, $\alpha=\beta$, $\alpha \ll \beta$.

I.3. Pour tout $\alpha,\beta>0$, montrer que

$$\int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-u-v}u^{\alpha-1}v^{\beta-1}\mathrm{d}u\mathrm{d}v = \Gamma(\alpha+\beta)\cdot\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\mathrm{d}t\,.$$

Indication: on pourra utiliser le changement de variables $u = zt$ et $v = z(1-t)$.

I.4. Déduire de la question I.3. la valeur de la constante de normalisation $B(\alpha,\beta)$.

I.5. Montrez que $\mathbb{E}_{\alpha,\beta}[X] = \frac{\alpha}{\alpha+\beta}$.

I.6. Calculez le moment d'ordre deux de X. En utilisant la question I.5., en déduire que

$$\mathrm{Var}_{\alpha,\beta}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}\,.$$

Partie II : Estimation des paramètres

Nous nous intéressons maintenant à l'estimation des paramètres (inconnus) α et β. On considère dans cette partie un n-échantillon i.i.d. $X_1,\ldots,X_n$ de loi $B(\alpha,\beta)$.

II.1. Rappelez brièvement le principe de la méthode des moments.

II.2. En utilisant les résultats de la partie I, proposez un estimateur des paramètres α et β. On pourra utiliser les notations suivantes :

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \text{et} \qquad \hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu}_n)^2 .$$

II.3. Étudiez la consistance de l'estimateur proposé à la question II.2.

II.4. On suppose dans cette question que $\alpha > 1$. Calculez $\mathbb{E}[1/X]$. Pouvez-vous proposer un autre estimateur de (α, β) dans ce cas?