LSTAT2040: Examen

Durée 3 heures

Les feuilles blanches sont vos feuilles de réponses. Inscrivez vos nom et prénom sur chacune d'entre elles et numérotez-les dans l'ordre de lecture. Seules ces feuilles seront corrigées. Les feuilles de couleur vous serviront de brouillon.
À la fin de l'examen, vous devez rendre toutes les feuilles que vous avez reçues, y compris les feuilles de brouillon et les feuilles vierges que vous n'avez pas utilisées.
Tout échange d'informations, sous quelque forme que ce soit, est interdit et sera considéré comme une tricherie.
Toute réponse non justifiée sera considérée comme étant erronée.
Vous pouvez consulter le syllabus du cours.
L'examen est sur 22 et compte 2 points de bonus.

Excercice 1 (15 pts)

X est une variable aléatoire dont la fonction de densité est donnée par

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} I(0 < x < 1), \ \alpha, \beta > 0,$$

où $B(\alpha, \beta) := \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, $\forall \alpha, \beta$.
Γ est la fonction gamma $\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t)\, dt$. Cette fonction possède un certain nombre de propriétés, notamment le fait que $\Gamma(x+1) = x\Gamma(x)$, $\forall x > 0$ et $\Gamma(x+1) = x!$, $\forall x \in \mathbb{N}$.
Dans ce qui suit, nous aurons également besoin de la fonction ψ (connue sous le nom `digamma`), définie comme suit

$$\psi(x) := \frac{d \log(\Gamma(x))}{dx} = \frac{\Gamma'(x)}{\Gamma(x)}.$$

Dans cet exercice, `digamma` est à considérer comme étant une fonction connue (au même titre que $\log(\cdot)$, ou $\exp(\cdot)$, par exemple), même si il n'est pas possible de la définir explicitement. L'une des caractéristiques intéressantes de cette fonction est le fait que $\psi(x+1) = \frac{1}{x} + \psi(x)$.
Sauf indication contraire, nous supposons que α et β sont inconnus et que nous observons X_i, $i = 1, \ldots, n$, un échantillon *iid* de X.

1 (2 pts)

Montrez que

$$E(X) = \frac{\alpha}{\alpha+\beta}, \text{ et } E(X^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}.$$

2 (2 pts)

Montrez que f fait partie des familles exponentielles. En suivant les notations du cours pour les familles exponentielles, précisez ses différentes composantes (h, B, ...) et donnez sa forme canonique. Servez-vous en pour prouver que

$$E(\log(X)) = \psi(\alpha) - \psi(\alpha+\beta), \text{ et que } Var(\log(X)) = \psi'(\alpha) - \psi'(\alpha+\beta).$$

3 (2 pts)

Supposons que β soit connu. Trouver un estimateur de α par la méthode de moments (MM). Quelle est sa distribution asymptotique ? Donnez sa variance asymptotique sous forme d'une expression explicite en fonction de α et β uniquement.

4 (3 pts)

Montrez que l'estimateur MM de $\boldsymbol{\theta} := (\alpha, \beta)$ est donné par $\hat{\boldsymbol{\theta}}_{MM} = (\hat{\alpha}_{MM}, \hat{\beta}_{MM})$, où

$$\hat{\alpha}_{MM} = \overline{X}\frac{\overline{X} - \overline{X^2}}{\overline{X^2} - \overline{X}^2}, \text{ et } \hat{\beta}_{MM} = (1-\overline{X})\frac{\overline{X} - \overline{X^2}}{\overline{X^2} - \overline{X}^2}.$$

$\hat{\boldsymbol{\theta}}_{MM}$ est-il consistant ? Converge-t-il en distribution vers $\boldsymbol{\theta}$?

5 (2 pts)

Donnez, en fonction de $\psi(\cdot)$, les équations de vraisemblances qui permettent de calculer l'estimateur du maximum de vraisemblance (MV) de $\boldsymbol{\theta}$. Expliquez, le plus précisément possible, la procédure à suivre pour résoudre ces équations numériquement.

6 (3 pts)

Supposons que $\beta = 1$. Trouver l'expression explicite de l'estimateur MV de α. Cet estimateur est-il efficace ? Prouvez sa consistance et trouvez sa variance asymptotique.

7 (1 pt)

Lequel des estimateurs MM ou MV choisirez-vous si vous savez que $\beta = 1$ et que votre taille d'échantillon est "grande" ?

Excercice 2 (7 pts)

Soit X_i, $i = 1, \ldots, n$, un échantillon *iid* de X dont la densité est donnée par

$$f(x) = \frac{1}{2\theta} I(0 \leq x \leq 2\theta),$$

où $\theta > 0$ est inconnu.

1 (2 pts)

Calculez la moyenne, la variance et la fonction de répartition (cdf) de X.

2 (2 pts)

Soit T l'estimateur du maximum de vraisemblance de θ. Trouvez T et montrez que sa densité est donnée par

$$t \mapsto \frac{n}{\theta^n} t^{n-1} I(0 \leq t \leq \theta).$$

On en déduit que T est biaisé. Effectuez une correction pour le transformer en un estimateur non-biaisé.

3 (2 pts)

Soit $\hat{\theta}_1 = \frac{n+1}{2n} \max_{1 \leq i \leq n} X_i$ un estimateur de θ. Montrez que la variance (exact) de $\hat{\theta}_1$ est donnée par

$$\frac{1}{n(n+2)} \theta^2.$$

Montez que $\hat{\theta}_1$ est consistant.

4 (1 pt)

Un autre estimateur possible de θ est $\hat{\theta}_2 = \overline{X}$. De ces deux estimateurs ($\hat{\theta}_1$ and $\hat{\theta}_2$), lequel est le meilleur ?