Statistique Mathématique

Corrigé du partiel — 28 février 2022

En italiques des remarques qui ne font pas à proprement parler de la correction.

Estimation des paramètres d'une loi bêta

I.1. (1 point) Soit $\alpha, \beta > 1$. Puisque $t^{\alpha-1}(1-t)^{\beta-1}$ est positif pour tout $t \in [01]$, $B(\alpha, \beta) > 0$ et on peut étudier les variations de $g : t \mapsto t^{\alpha-1}(1-t)^{\beta-1}$ au lieu de celles de $t \mapsto f_{\alpha,\beta}(t)$. On remarque tout d'abord que $g(0) = g(1) = 0$, puisque α et β sont strictement plus grands que 1. De plus, g est différentiable sur $[0, 1]$, avec

$$g'(t) = [(\alpha - 1) - t(\alpha + \beta - 2)] \cdot t^{\alpha-2}(1-t)^{\beta-2} .$$

Notons $t_0 := (\alpha - 1)/(\alpha + \beta - 2)$. Il est clair que $t_0 \in [0, 1]$. De plus, g' est positive sur $[0, t_0]$ et négative sur $[t_0, 1]$, donc t_0 est le mode de la distribution.

Beaucoup d'erreurs dans le calcul de la dérivée.

I.2. (1.5 point) Voir Figure 1.

J'ai donné un demi point par cas.

I.3. (2 points) Le jacobien du changement de variable donné par l'énoncé est

$$\begin{vmatrix} t & 1-t \\ z & -z \end{vmatrix} = t \cdot (-z) - z \cdot (1-t) = -z .$$

Remarquant que $u + v = z$, on en déduit que

$$\begin{aligned} \int_0^{+\infty} \int_0^{+\infty} \mathrm{e}^{-u-v} u^{\alpha-1} v^{\beta-1} \mathrm{d}u\mathrm{d}v &= \int_0^{+\infty} \int_0^1 \mathrm{e}^{-z} (zt)^{\alpha-1} (z(1-t))^{\beta-1} \cdot |-z| \, \mathrm{d}t\mathrm{d}z \\ &= \int_0^{+\infty} \int_0^1 \mathrm{e}^{-z} z^{\alpha+\beta-1} t^{\alpha-1} (1-t)^{\beta-1} \mathrm{d}t\mathrm{d}z \\ &= \int_0^{+\infty} \mathrm{e}^{-z} z^{\alpha+\beta-1} \mathrm{d}z \cdot \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} \mathrm{d}t . \end{aligned}$$

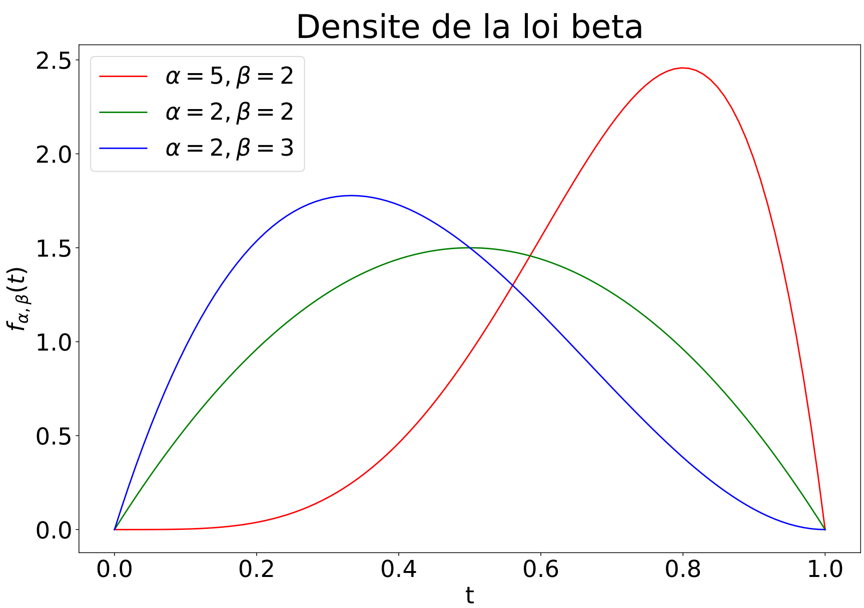

Figure 1: Densité de la loi bêta pour $\alpha, \beta > 1$ en fonction de leur taille relative.

On reconnaît la définition de $\Gamma(\alpha+\beta)$ et on retrouve le résultat de l'énoncé.

Question banalisée puisque certains ne savaient pas faire un changement de variables multivarié.

I.4. (2 points) Pour que $f_{\alpha,\beta}$ soit une densité, il faut qu'elle somme à 1. De la question précédente, on déduit que

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

La double intégrale se sépare facilement en un produit,

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

dans lequel on reconnaît une nouvelle fois la définition de Γ. Ainsi,

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\,.$$

La fonction B est parfois appelé fonction Beta *dans la littérature.*

I.5. (2 point) On calcule

$$\begin{aligned}
\mathbb{E}_{\alpha,\beta}[X] &= \int_0^1 t \cdot t^{\alpha-1}(1-t)^{\beta-1} \frac{\mathrm{d}t}{B(\alpha,\beta)} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \int_0^1 t^{\alpha+1-1}(1-t)^{\beta-1}\mathrm{d}t\,.
\end{aligned}$$

D'après les questions précédentes, nous savons que

$$\frac{1}{B(\alpha+1,\beta)} \int_0^1 t^{\alpha+1-1}(1-t)^{\beta-1}\mathrm{d}t = 1\,.$$

On en déduit que

$$\mathbb{E}_{\alpha,\beta}[X] = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)} = \frac{\alpha}{\alpha+\beta}\,,$$

on l'on a utilisé les propriétés de la fonction Gamma données dans l'énoncé.

J'ai donné un demi point pour la méthode (intégrer t contre la densité).

I.6. (2 points) De la même manière que dans la question précédente, on écrit

$$\begin{aligned}
\mathbb{E}_{\alpha,\beta}\left[X^2\right] &= \int_0^1 t^2 \cdot t^{\alpha-1}(1-t)^{\beta-1} \frac{\mathrm{d}t}{B(\alpha,\beta)} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+2+\beta)} \\
\mathbb{E}_{\alpha,\beta}\left[X^2\right] &= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}\,.
\end{aligned}$$

Pour obtenir la variance, nous écrivons

$$\begin{aligned}
\mathrm{Var}_{\alpha,\beta}(X) &= \mathbb{E}_{\alpha,\beta}\left[X^2\right] - \mathbb{E}_{\alpha,\beta}[X]^2 \\
&= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2} \\
&= \frac{(\alpha^2+\alpha)(\alpha+\beta) - \alpha^2(\alpha+\beta-1)}{(\alpha+\beta)^2(\alpha+\beta-1)} \\
\mathrm{Var}_{\alpha,\beta}(X) &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}\,,
\end{aligned}$$

où l'on a utilisé le résultat de la question I.5. pour passer de la première ligne à la deuxième.

II.1. (2 points) Voir le cours.

II.2. (3 points) Notons μ l'espérance de X et σ^2 sa variance. Nous avons à résoudre

$$\begin{cases} \mu &= \frac{\alpha}{\alpha+\beta} \\ \sigma^2 &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \,. \end{cases}$$

Il y a différentes manières de procéder. On peut par exemple écrire que $\beta = \alpha(1-\mu)/\mu$ en partant de la première équation, puis en injectant cette relation dans la second équation obtenir

$$\alpha\sigma^2 = \frac{\mu^2 \cdot \alpha(1-\mu)}{\mu(\alpha/\mu+1)} \,.$$

On obtient

$$\begin{cases} \alpha &= \mu\left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right) \\ \beta &= (1-\mu)\left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right) \,. \end{cases}$$

En utilisant les notations de l'énoncé, on propose les estimateurs

$$\begin{cases} \hat{\alpha} &= \hat{\mu}\left(\frac{\hat{\mu}(1-\hat{\mu})}{\hat{\sigma}^2} - 1\right)_+ \\ \hat{\beta} &= (1-\hat{\mu})\left(\frac{\hat{\mu}(1-\hat{\mu})}{\hat{\sigma}^2} - 1\right)_+ \,. \end{cases}$$

Je n'ai pas pénalisé si la copie ne prenait pas la précaution de rendre le terme entre parenthèse plus grand que 0.

II.3. (2 points) Les X_i sont i.i.d. et ont des moments d'ordre 1 et 2. D'après la loi faible des grands nombres,

$$\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} \mu \qquad \text{et} \qquad \frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{\mathbb{P}} \mathbb{E}_{\alpha,\beta}\left[X^2\right] \,.$$

Nous en déduisons que $\hat{\mu} \xrightarrow{\mathbb{P}} \mu$ et

$$\hat{\sigma} = \left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2 - \frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{\mathbb{P}} \sigma^2 \,.$$

Par le théorème de continuité, nous en déduisons la consistance de $(\hat{\alpha}, \hat{\beta})$.

II.4. (2 points bonus) En raisonnant comme dans la première partie, sous l'hypothèse $\alpha > 1$, on obtenait

$$\mathbb{E}_{\alpha,\beta}\left[\frac{1}{X}\right] = \frac{\alpha+\beta-1}{\alpha-1} \,.$$

Combiné par exemple avec le premier moment, cela donnait lieu à un autre estimateur par la méthode des moments : en appelant $\hat{\nu}$ la moyenne des inverses,

$$\begin{cases} \hat{\alpha} &= \frac{-1}{\frac{\hat{\mu}}{\hat{\mu}-1} + \frac{1}{\hat{\nu}-1}} \\ \hat{\beta} &= \frac{1}{\frac{\hat{\mu}}{\hat{\mu}-1} + \frac{1}{\hat{\nu}-1}} \cdot \frac{\hat{\mu}}{\hat{\mu}-1} \,. \end{cases}$$

Il y avait plusieurs expressions possibles.