Statistique (1), 2016/2017

Si $\{X_1, \cdots, X_n\}$ est un échantillon, on note $X_{i:n}$ la i-ème statistique d'ordre. Si $X \sim \mathcal{N}(\mu, \sigma^2)$, le second terme désigne la *variance*; si $X \sim \mathcal{S}t(k)$, X suit une loi de Student à k degrés de liberté. Enfin, X suit une loi Gamma $\mathcal{G}(\alpha, \beta)$ si sa densité s'écrit

$$f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x\beta} \mathbf{1}(x \geq 0),$$

et X suit une loi Beta $\mathcal{B}(\alpha, \beta)$ si sa densité s'écrit

$$f_{\alpha,\beta}(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbf{1}(x \in [0,1]),$$

où $\Gamma(\cdot)$ désigne la fonction Gamma,

$$\Gamma : z \mapsto \int_0^{+\infty} t^{z-1} e^{-t} \, \mathrm{d}t.$$

$\boxed{1}$ On dispose d'un échantillon $\{x_1, \cdots, x_n\}$ tiré suivant un mélange

$$X = \begin{cases} U_1 \sim \mathcal{U}([0,a]) \text{ avec probabilité } \theta \\ \\ U_2 \sim \mathcal{U}([0,b]) \text{ avec probabilité } 1-\theta \end{cases}$$

avec $\theta \in (0,1)$ et $a < b$. Soit N le nombre d'observations entre 0 et a dans l'échantillon. Quelle est la loi de N ? Quel est l'estimateur du maximum de vraisemblance de θ.

$\boxed{2}$ Considérons un échantillon tiré suivant une loi de densité

$$f_\theta(x) = \frac{2x}{\theta^2} \text{ pour } x \in [0, \theta].$$

Trouver l'estimateur du maximum de vraisemblance de la médiane de la distribution.

$\boxed{3}$ Considérons un échantillon $\{x_1, \cdots, x_n\}$ tiré suivant une loi $\mathcal{N}(\theta, 1)$. Malheureusement, seules des indicatrices, indiquant que les observations étaient positives, ont été gardées, i.e. $\{y_1, \cdots, y_n\}$, avec $y_i = \mathbf{1}(x_i > 0)$. Quelle est l'estimateur du maximum de vraisemblance de θ?

$\boxed{4}$ Soit $X_1, \cdots, X_n$ un échantillon i.i.d. de loi de Weibull, de densité

$$f(x) = \beta\alpha^{-\beta} x^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right] \text{ pour } x \geq 0.$$

Montrer que $Y = \min\{X_1, \cdots, X_n\}$ admet pour densité

$$g(y) = n\beta y^{\beta-1}\alpha^{-\beta} \exp\left[-n\left(\frac{y}{\alpha}\right)^\beta\right] \text{ pour } y \geq 0.$$

$\boxed{5}$ Soit $X_1, \cdots, X_5$ un échantillon i.i.d. de loi $\mathcal{U}([0,1])$. Quelle est la probabilité que les 5 observations soient dans l'intervalle $[0, 1/2]$?

$\boxed{6}$ Supposons F continue et strictement croissante. Soit $\alpha \in (0,1)$ tel quel $k/n \to \alpha$ lorsque $n \to \infty$. Montrer que

$$X_{k:n} \overset{p.s.}{\to} Q(\alpha) \text{ lorsque } n \to \infty.$$

$\boxed{7}$ Soient X_1 et X_2 deux variables indépendantes de densité $f(x) = 2(1-x)$ pour $x \in [0,1]$. Montrer que

$$\mathbb{P}(\max\{X_1, X_2\} \geq 2\min\{X_1, X_2\}) = \frac{7}{12}.$$

$\boxed{8}$ Soient X_1 et X_2 deux variables indépendantes $\mathcal{E}(\lambda)$. Donner la loi de $Z = \max\{X_1, X_2\} - \min\{X_1, X_2\}$.

$\boxed{9}$ Soit $X_1, \cdots, X_n$ un échantillon i.i.d. de loi $\mathcal{U}([0,1])$. On pose $E = X_{n:n} - X_{1:n}$. Montrer que la densité de E est

$$g(x) = n(n-1)(1-x)x^{n-2} \text{ pour } x \in [0,1].$$

Montrer que

$$\mathbb{E}(E) = \frac{n-1}{n+1}$$

$\boxed{10}$ Soit $X_1, \cdots, X_n$ un échantillon i.i.d. de loi $\mathcal{E}(1)$. On pose $D_{i,j} = X_{i:n} - X_{j:n}$. Montrer que $D_{k+1,k}$ suit une loi exponentielle de paramètre $n-k$.

$\boxed{11}$ Soit $X_1, \cdots, X_n$ un échantillon i.i.d. de loi $\mathcal{U}([0,1])$. On pose $D_{i,j} = X_{i:n} - X_{j:n}$. Montrer que si $i > j$, alors $D_{i,j}$ suit une loi Beta, $\mathcal{B}(i-j, n+1-[i-j])$.
On suppose disposer d'une nouvelle observation X_{n+1} (indépendante des autres, et de loi $\mathcal{U}([0,1])$).
Calculer

$$\begin{cases} \mathbb{P}(X_{n+1} > \max\{X_1, \cdots, X_n\}) \\ \\ \mathbb{P}(X_{n+1} \in [\min\{X_1, \cdots, X_n\}, \max\{X_1, \cdots, X_n\}]) \end{cases}$$

$\boxed{12}$ Soient X_1 et X_2 deux variables indépendantes de loi $\mathcal{N}(0, \sigma^2)$. Montrer que

$$\mathbb{E}(\max\{X_1, X_2\}) = \frac{\sigma}{\sqrt{\pi}}.$$

Montrer que

$$\mathbb{P}(\min\{X_1, X_2\} \leq 0 \leq \max\{X_1, X_2\} = \frac{1}{2}.$$

$\boxed{13}$ Soit $X_1, \cdots, X_4$ un échantillon i.i.d. de densité $f(x) = 2x$ pour $x \in [0,1]$ (et 0 ailleurs). On note $Y_i = X_{i:n}$

(a) donner la loi du couple (Y_3, Y_4)

(b) donner la loi conditionnelle de Y_3 sachant $Y_4 = y_4$

(c) montrer que $\mathbb{E}(Y_3|Y_4 = y_4) = \frac{6}{7}y_4$.

$\boxed{14}$ Soit $X_1, \cdots, X_n$ un échantillon i.i.d. de loi F absolument continue, de médiane $m = F^{-1}(1/2)$. Trouver la plus petite valeur n telle que

$$\mathbb{P}(m \in [\min\{X_1, \cdots, X_n\}, \max\{X_1, \cdots, X_n\}]) > 99\%.$$

$\boxed{15}$ Considérons une variable aléatoire X suivant une loi de densité $f_\lambda(x) = \lambda|x|\exp(-\lambda x^2)$, pour tout $x \in \mathbb{R}$, pour un paramètre $\lambda \in]0, +\infty[$.

1. Vérifiez que f_λ est une densité.
2. Donnez la loi de $Y = |X|$.
3. Vérifiez que pour $n > 0$, $\mathbb{E}(Y^n) = \frac{n}{2\lambda}\mathbb{E}(Y^{n-2})$
4. (a) Calculez l'espérance, la variance et le moment d'ordre quatre de X
   (b) Donnez la loi limite et la vitesse de $\overline{X^2}_n$, le moment d'ordre deux empirique.
   (c) Déduisez-en un estimateur $\hat{\lambda}_1$ de λ dont vous étudierez la constistance et la loi limite.
5. (a) Calculez l'espérance, la variance et le moment d'ordre quatre de Y
   (b) Donnez la loi limite et la vitesse de $\overline{|X|}_n$.
   (c) Déduisez-en un estimateur $\hat{\lambda}_2$ de λ dont vous étudierez la constistance et la loi limite.
6. Comparez les deux estimateurs obtenus.

$\boxed{16}$ On suppose que les observations $U_1, \cdots, U_n$ sont indṕendantes et uniformément distribuées sur $[-\theta, \theta]$, avec $\theta \in \mathbb{R}_+$.

1. Calculez $\mathbb{E}(U_{n:n}) = \mathbb{E}(\max\{U_i\})$ et $\mathbb{E}(U_{1:n}) = \mathbb{E}(\min\{U_i\})$
2. Quelle est la densité de $|U_i|$?
3. Calculez $\mathbb{E}(|U|_{n:n}) = \mathbb{E}(\max\{|U_i|\})$
4. Quel est l'estimateur du maximum de vraisemblance de θ ?
5. Cet estimateur est-il est un estimateur sans biais de θ ?
6. Suggérez un estimateur pour la méthode des moments de θ ?

$\boxed{17}$ On dispose de $n = 2k$ variables indépendantes, de loi $\mathcal{N}(0,1)$. Pour l'ensemble des variables aléatoires du tableau de la page suivante, trouver leur loi (une seule réponse possible).

	$\mathcal{N}(0,1/n)$	$\mathcal{N}(0,n)$	$\mathcal{S}t(n-1)$	$\mathcal{S}t(n)$	$\chi^2(n-1)$	$\chi^2(n)$	$\mathcal{F}(\star,\star)$	autre
$X_{n:n}$	☐	☐	☐	☐	☐	☐	☐	☐
$\sum_{i=1}^{n} X_i$	☐	☐	☐	☐	☐	☐	☐	☐
$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	☐	☐	☐	☐	☐	☐	☐	☐
$\sum_{i=1}^{n} X_i^2$	☐	☐	☐	☐	☐	☐	☐	☐
$\frac{1}{n}\sum_{i=1}^{n} X_i^2$	☐	☐	☐	☐	☐	☐	☐	☐
$\frac{1}{n^2}\sum_{i=1}^{n} X_i^2$	☐	☐	☐	☐	☐	☐	☐	☐
$\sum_{i=1}^{n}(X_i-\overline{X})^2$	☐	☐	☐	☐	☐	☐	☐	☐
$\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$	☐	☐	☐	☐	☐	☐	☐	☐
$\frac{1}{n^2}\sum_{i=1}^{n}(X_i-\overline{X})^2$	☐	☐	☐	☐	☐	☐	☐	☐
$\sum_{i=1}^{k} X_i - \sum_{i=k+1}^{n} X_i$	☐	☐	☐	☐	☐	☐	☐	☐
$\frac{X_n}{\sqrt{\sum_{i=1}^{n-1} X_i^2}}$	☐	☐	☐	☐	☐	☐	☐	☐
$\sqrt{n-1}\frac{X_n}{\sqrt{\sum_{i=1}^{n-1} X_i^2}}$	☐	☐	☐	☐	☐	☐	☐	☐
$\frac{X_n^2}{\sum_{i=1}^{n-1} X_i^2}$	☐	☐	☐	☐	☐	☐	☐	☐
$(n-1)\frac{X_n^2}{\sum_{i=1}^{n-1} X_i^2}$	☐	☐	☐	☐	☐	☐	☐	☐
$2\frac{\sum_{i=1}^{k} X_i^2}{\sum_{i=1}^{n} X_i^2}$	☐	☐	☐	☐	☐	☐	☐	☐
$\frac{\sum_{i=1}^{k} X_i^2}{\sum_{i=k+1}^{n} X_i^2}$	☐	☐	☐	☐	☐	☐	☐	☐