Correction, Statistique et Probabilités, L2 MIASHS, CC2
décembre 2017, Durée : deux heures

$\boxed{\text{AM}}$

Exercice 1. Loi normale
La définition de X_i peut s'écrire comme suit

$$X_i = f_i(\varepsilon_i), \quad \text{où} \quad f_i(x) = \theta i^2 + x. \qquad (*)$$

1. D'après (*) X_i ne dépend que de ε_i. Puisque les ε_i sont indépendantes, les X_i sont indépendantes. Puisque $\mathbb{E}(X_i) = \theta i^2$, les espérances de X_i sont differentes. Donc les X_i ne sont pas identiquement distribuées.

2. Puisque la fonction $f_i(x)$ dans (*) est linéaire et ε_i est gaussienne, la loi de X_i est gaussienne. On a $\operatorname{Var} X_i = \operatorname{Var} \varepsilon_i = 1$ donc $X_i \sim \mathcal{N}(\theta i^2, 1)$. On a donc la vraisemblance donnée dans la question.

3. On calcule d'abord la log-vraisemblance $H(\theta)$

$$H(\theta) = \ln L(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i - \theta i^2)^2.$$

On a ensuite

$$H'(\theta) = \sum_{i=1}^{n} i^2(x_i - \theta i^2) = 0 \Leftrightarrow \theta = \frac{\sum i^2 x_i}{\sum i^4},$$

et $H''(\theta) = -\sum i^4 < 0$. Donc $\hat\theta = \frac{\sum i^2 x_i}{\sum i^4}$.

4.

$$\mathbb{E}(\hat\theta) = \frac{\sum i^2 \mathbb{E} X_i}{\sum i^4} = \theta,$$

$$\operatorname{Var}(\hat\theta) = \frac{\sum i^4 \operatorname{Var} X_i}{(\sum i^4)^2} = \frac{1}{\sum i^4}.$$

5. On a $\frac{\hat\theta - \theta}{\sqrt{\frac{1}{\sum i^4}}} \sim \mathcal{N}(0,1)$ et $\mathbb{P}(|\mathcal{N}(0,1)| < 1,96) = 0,95$. Donc l'intervalle de confiance à 95% pour θ et $[\hat\theta - \frac{1,96}{\sqrt{\sum i^4}}, \hat\theta + \frac{1,96}{\sqrt{\sum i^4}}]$.

Exercice 2. Loi uniforme

1. Puisque $\mathbb{P}(Y \le y) = \mathbb{P}(2 - X \le y) = \mathbb{P}(X \ge 2 - y) = 1 - (2 - y) = y - 1$, la fonction de répartition de Y est

$$G(y) = \begin{cases} 0, & y < 1, \\ y - 1, & 1 \le y \le 2, \\ 1, & y > 2. \end{cases}$$

2.

$$g(y) = G'(y) = \begin{cases} 0, & y < 1, \\ 1, & 1 \le y \le 2, \\ 0, & y > 2. \end{cases}$$

3. $\mathbb{E}Y = \int_1^2 y dy = 3/2$, $\mathbb{E}Y^2 = \int_1^2 y^2 dy = 7/3$, $\operatorname{Var} Y = 7/3 - 9/4 = 1/12$.

Exercice 3. Moments de la loi gaussienne standard

La densité de X est $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

1. La fonction $f(x)$ est paire, i.e. $f(-x) = f(x), \forall x \in \mathbb{R}$. Puisque l'intégrande $x^{2n+1} f(x)$ est impair, $\mathbb{E}X^{2n+1} = \int_{-\infty}^{\infty} x^{2n+1} f(x) dx = 0$. (Par changement de variable $x = -t$ on peut prouver $\int_{-\infty}^{0} x^{2n+1} f(x) dx = -\int_0^\infty t^{2n+1} f(t) dt$.)

2. En intégrant par partie on obtient

$$\int_{-\infty}^{\infty} \frac{x^{2n+2}}{\sqrt{2\pi}} e^{-x^2/2} dx = -\frac{1}{\sqrt{2\pi}} ([x^{2n+1} e^{-x^2/2}]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{-x^2/2} (2n+1) x^{2n} dx) = (2n+1) \int_{-\infty}^{\infty} \frac{x^{2n}}{\sqrt{2\pi}} e^{-x^2/2} dx,$$

d'où vient

$$c_{n+1} = \mathbb{E}X^{2n+2} = (2n+1)\mathbb{E}X^{2n} = (2n+1)c_n.$$

D'après la relation précédente on a $\mathbb{E}X^4 = 3\mathbb{E}X^2 = 3, \mathbb{E}X^6 = 5\mathbb{E}X^4 = 5 \cdot 3$,... ainsi

$$\mathbb{E}X^{2n} = (2n-1)(2n-3) \cdots 3 \cdot 1 = (2n-1)!!.$$

$\boxed{\text{BM}}$

Exercice 1. Loi uniforme

1. Puisque $\mathbb{P}(Y \leq y) = \mathbb{P}(3 - X \leq y) = \mathbb{P}(X \geq 3 - y) = 1 - (3 - y) = y - 2$, la fonction de répartition de Y est

$$G(y) = \begin{cases} 0, & y < 2, \\ y - 2, & 2 \leq y \leq 3, \\ 1, & y > 3. \end{cases}$$

2.
$$g(y) = G'(y) = \begin{cases} 0, & y < 2, \\ 1, & 2 \leq y \leq 3, \\ 0, & y > 3. \end{cases}$$

3. $\mathbb{E}Y = \int_2^3 y dy = 5/2$, $\mathbb{E}Y^2 = \int_2^3 y^2 dy = 19/3$, $\operatorname{Var} Y = 19/3 - 25/4 = 1/12$.

Exercice 2. Moments de la loi gaussienne standard

La densité de X est $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}$.

1. La fonction $f(x)$ est paire, i.e. $f(-x) = f(x), \forall x \in \mathbb{R}$. Puisque l'intégrande $x^{2n+1} f(x)$ est impair, $\mathbb{E}X^{2n+1} = \int_{-\infty}^{\infty} x^{2n+1} f(x) dx = 0$. (Par changement de variable $x = -t$ on peut prouver $\int_{-\infty}^{0} x^{2n+1} f(x) dx = -\int_0^{\infty} t^{2n+1} f(t) dt$.)

2. En intégrant par partie on obtient

$$\int_{-\infty}^{\infty} \frac{x^{2n+2}}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx = -\frac{\sigma}{\sqrt{2\pi}} ([x^{2n+1} e^{-\frac{x^2}{2\sigma^2}}]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} (2n+1) x^{2n} dx) = \sigma^2 (2n+1) \int_{-\infty}^{\infty} \frac{x^{2n}}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx,$$

d'où vient

$$a_{n+1} = \mathbb{E}X^{2n+2} = \sigma^2 (2n+1) \mathbb{E}X^{2n} = \sigma^2 (2n+1) a_n.$$

D'après la relation précédente on a $\mathbb{E}X^4 = 3\sigma^2 \mathbb{E}X^2 = \sigma^4 3, \mathbb{E}X^6 = 5\sigma^2 \mathbb{E}X^4 = \sigma^6 5 \cdot 3$,... ainsi

$$\mathbb{E}X^{2n} = \sigma^{2n} (2n-1)(2n-3) \cdots 3 \cdot 1 = \sigma^{2n} (2n-1)!!.$$

(Comme il y a une erreur dans l'énoncé, on donne les points complets pour les réponses à σ^2 près, c.a.d. celles comme dans AM.)

Exercice 3. Loi normale

La définition de X_i peut s'écrire comme suit

$$X_i = f_i(\varepsilon_i), \quad \text{où} \quad f_i(x) = \theta/i + x. \qquad (*)$$

1. D'après (*) X_i ne dépend que de ε_i. Puisque les ε_i sont indépendantes, les X_i sont indépendantes. Puisque $\mathbb{E}(X_i) = \theta/i$, les espérances de X_i sont différentes. Donc les X_i ne sont pas identiquement distribuées.

2. Puisque la fonction $f_i(x)$ dans (*) est linéaire et ε_i est gaussienne, la loi de X_i est gaussienne. On a $\operatorname{Var} X_i = \operatorname{Var} \varepsilon_i = 1$ donc $X_i \sim \mathcal{N}(\theta/i, 1)$. On a donc la vraisemblance donnée dans la question.

3. On calcule d'abord la log-vraisemblance $H(\theta)$

$$H(\theta) = \ln L(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^{n} (x_i - \theta/i)^2.$$

On a ensuite

$$H'(\theta) = \sum_{i=1}^{n} (x_i - \theta/i)/i = 0 \Leftrightarrow \theta = \frac{\sum x_i/i}{\sum 1/i^2},$$

et $H''(\theta) = -\sum 1/i^2 < 0$. Donc $\hat{\theta} = \frac{\sum x_i/i}{\sum 1/i^2}$.

4.

$$\mathbb{E}(\hat{\theta}) = \frac{\sum 1/i \mathbb{E} X_i}{\sum 1/i^2} = \theta,$$

$$\mathrm{Var}\,(\hat{\theta}) = \frac{\sum 1/i^2 \mathrm{Var}\, X_i}{(\sum 1/i^2)^2} = \frac{1}{\sum 1/i^2}.$$

5. On a $\frac{\hat{\theta}-\theta}{\sqrt{\frac{1}{\sum 1/i^2}}} \sim \mathcal{N}(0,1)$ et $\mathbb{P}(|\mathcal{N}(0,1)| < 1,64) = 0,90$. Donc l'intervalle de confiance à 90% pour θ et $[\hat{\theta} - \frac{1,64}{\sqrt{\sum 1/i^2}}, \hat{\theta} + \frac{1,64}{\sqrt{\sum 1/i^2}}]$.

$\boxed{\textbf{AJ}}$

Exercice 1. Loi uniforme

1. La vraisemblance est

$$L(\theta) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{1}{\theta} 1\!\!1_{[0,\theta]}(x_i) = \frac{1}{\theta^n} \prod_{i=1}^{n} 1\!\!1_{[0,\theta]}(x_i).$$

Si les x_i sont tous inférieurs à θ, i.e. $\max(x_1, \ldots, x_n) \leq \theta$, alors $L(\theta) = \frac{1}{\theta^n}$ et $L(\theta)$ est maximisée, sinon $L(\theta) = 0$. Donc l'estimateur du maximum de vraisemblance de θ est $\max(x_1, \ldots, x_n)$.

2. La fonction de répartition de X_i est

$$F(x) = \begin{cases} 0, & x < 0, \\ \int_0^x \frac{1}{\theta} dt = \frac{x}{\theta}, & 0 \leq x \leq \theta, \\ 1, & x > \theta. \end{cases}$$

La fonction de répartition de Z_n est

$$F_{Z_n}(z) = F(z)^n = \begin{cases} 0, & z < 0, \\ \frac{z^n}{\theta^n}, & 0 \leq z \leq \theta, \\ 1, & z > \theta. \end{cases}$$

3. La densité de Z_n est

$$f_{Z_n}(z) = F'_{Z_n}(z) = \begin{cases} 0, & z < 0, \\ \frac{n}{\theta^n} z^{n-1}, & 0 \leq z \leq \theta, \\ 0, & z > \theta. \end{cases}$$

4. $\mathbb{E}Z_n = \int_0^\theta \frac{n}{\theta^n} z^n dz = \frac{n}{n+1}\theta$. $\mathbb{E}Z_n^2 = \int_0^\theta \frac{n}{\theta^n} z^{n+1} dz = \frac{n}{n+2}\theta^2$. $\mathrm{Var}\, Z_n = \frac{n}{n+2}\theta^2 - (\frac{n}{n+1}\theta)^2 = \frac{n}{(n+1)^2(n+2)}\theta^2$.

5.
$$\lim_{n\to\infty} F_{Z_n}(z) = \begin{cases} 0, & z < \theta, \\ 1, & z \geq \theta. \end{cases}$$

On en déduit que $\mathbb{P}(Z = \theta) = 1$, Z est la constante θ.

6. $\mathbb{P}(|Z_n - \theta| > \varepsilon) = \mathbb{P}(Z_n < \theta - \varepsilon) = (\frac{\theta-\varepsilon}{\theta})^n$. On en déduit que $\sum_{n=1}^\infty \mathbb{P}(|Z_n - \theta| > \varepsilon) = \sum_{n=1}^\infty (\frac{\theta-\varepsilon}{\theta})^n < \infty$ car $0 < \frac{\theta-\varepsilon}{\theta} < 1$.

Exercice 2. Loi exponentielle

1. $\mathbb{P}(Y = 1) = \mathbb{P}(1 \leq X < 2) = F(2) - F(1) = e^{-3} - e^{-6}$.

2. $Y \in \{0, 1, 2, \ldots\} = \mathbb{N}$. $\mathbb{P}(Y = k) = \mathbb{P}(k \leq X < k+1) = F(k+1) - F(k) = e^{-3k} - e^{-3(k+1)}$.

3. On a d'abord $\mathbb{P}(Z \leq z) = \sum_{y=0}^\infty \mathbb{P}(y \leq X \leq y + z)$, ensuite $\sum_{y=0}^\infty \mathbb{P}(y \leq X \leq y + z) = \sum_{y=0}^\infty e^{-3y} - e^{-3(y+z)} = \frac{1-\exp(-3z)}{1-\exp(-3)}$.

4. La densité de Z est $f(z) = (\frac{1-\exp(-3z)}{1-\exp(-3)})' = \frac{3\exp(-3z)}{1-\exp(-3)}$.

5. On a d'abord $\mathbb{E}(Z) = \int_0^1 z f(z) dz = \frac{3}{1-e^{-3}} \int_0^1 z e^{-3z} dz$. En intégrant par parties on a $\int_0^1 z e^{-3z} dz = \frac{e^3-4}{9e^3}$. Donc $\mathbb{E}(Z) = \frac{e^3-4}{3(e^3-1)} = \frac{1}{3} - \frac{1}{e^3-1}$.

$\boxed{\text{BJ}}$

Exercice 1. Loi exponentielle

1. $\mathbb{P}(Y=0) = \mathbb{P}(X<1) = F(1) = 1-e^{-2}$.

2. $Y \in \{0,1,2,\dots\} = \mathbb{N}$. $\mathbb{P}(Y=k) = \mathbb{P}(k \le X < k+1) = F(k+1) - F(k) = e^{-2k} - e^{-2(k+1)}$.

3. On a d'abord $\mathbb{P}(Z \le z) = \sum_{y=0}^{\infty} \mathbb{P}(y \le X \le y+z)$, ensuite $\sum_{y=0}^{\infty} \mathbb{P}(y \le X \le y+z) = \sum_{y=0}^{\infty} e^{-2y} - e^{-2(y+z)} = \frac{1-\exp(-2z)}{1-\exp(-2)}$.

4. La densité de Z est $f(z) = (\frac{1-\exp(-2z)}{1-\exp(-2)})' = \frac{2\exp(-2z)}{1-\exp(-2)}$.

5. On a d'abord $\mathbb{E}(Z) = \int_0^1 zf(z)dz = \frac{2}{1-e^{-2}} \int_0^1 ze^{-2z}dz$. En intégrant par parties on a $\int_0^1 ze^{-2z}dz = \frac{e^2-3}{4e^2}$. Donc $\mathbb{E}(Z) = \frac{e^2-3}{2(e^2-1)} = \frac{1}{2} - \frac{1}{e^2-1}$.

Exercice 2. Loi uniforme

1. La vraisemblance est

$$L(\lambda) = \prod_{i=1}^n f(x_i) = \prod_{i=1}^n \frac{1}{\lambda} 1\!\!1_{[0,\lambda]}(x_i) = \frac{1}{\lambda^n} \prod_{i=1}^n 1\!\!1_{[0,\lambda]}(x_i).$$

Si les x_i sont tous inférieurs à λ, i.e. $\max(x_1,\dots,x_n) \le \lambda$, alors $L(\lambda) = \frac{1}{\lambda^n}$ et $L(\lambda)$ est maximisée, sinon $L(\lambda) = 0$. Donc l'estimateur du maximum de vraisemblance de λ est $\max(x_1,\dots,x_n)$.

2. La fonction de répartition de X_i est

$$F(x) = \begin{cases} 0, & x < 0, \\ \int_0^x \frac{1}{\lambda} dt = \frac{x}{\lambda}, & 0 \le x \le \lambda, \\ 1, & x > \lambda. \end{cases}$$

La fonction de répartition de Z_n est

$$F_{Z_n}(z) = F(z)^n = \begin{cases} 0, & z < 0, \\ \frac{z^n}{\lambda^n}, & 0 \le z \le \lambda, \\ 1, & z > \lambda. \end{cases}$$

3. La densité de Z_n est

$$f_{Z_n}(z) = F'_{Z_n}(z) = \begin{cases} 0, & z < 0, \\ \frac{n}{\lambda^n} z^{n-1}, & 0 \le z \le \lambda, \\ 0, & z > \lambda. \end{cases}$$

4. $\mathbb{E}Z_n = \int_0^\lambda \frac{n}{\lambda^n} z^n dz = \frac{n}{n+1}\lambda$. $\mathbb{E}Z_n^2 = \int_0^\lambda \frac{n}{\lambda^n} z^{n+1} dz = \frac{n}{n+2}\lambda^2$. $\operatorname{Var} Z_n = \frac{n}{n+2}\lambda^2 - (\frac{n}{n+1}\lambda)^2 = \frac{n}{(n+1)^2(n+2)}\lambda^2$.

5.
$$\lim_{n\to\infty} F_{Z_n}(z) = \begin{cases} 0, & z < \lambda, \\ 1, & z \ge \lambda. \end{cases}$$

On en déduit que $\mathbb{P}(Z=\lambda) = 1$, Z est la constante λ.

6. $\mathbb{P}(|Z_n - \lambda| > \varepsilon) = \mathbb{P}(Z_n < \lambda - \varepsilon) = (\frac{\lambda-\varepsilon}{\lambda})^n$. On en déduit que $\sum_{n=1}^\infty \mathbb{P}(|Z_n - \lambda| > \varepsilon) = \sum_{n=1}^\infty (\frac{\lambda-\varepsilon}{\lambda})^n < \infty$ car $0 < \frac{\lambda-\varepsilon}{\lambda} < 1$.