

Examen final

Durée : 2h
Correction

Tout matériel qui ne ressemblerait pas à un simple stylo ou à une feuille blanche (calculatrice, portable, note de cours...) est interdit. Si vous vous demandez si vous avez le droit d'avoir un objet sur la table, c'est qu'il est certainement interdit.

Les parties sont indépendantes. Si un résultat n'est pas démontré, il pourra être admis dans la question suivante.

1 Loi exponentielle (7.5 points)

Dans cet exercice, nous étudions la variable X suivant la loi exponentielle de paramètre inconnu θ^* dont nous rappelons la densité définie pour tout $\theta > 0$ par :

$$f_\theta(x) = \theta e^{-\theta x} \mathbf{1}_{\mathbb{R}^+}(x).$$

1.1. Calculer pour tout $m \in \mathbb{N}$ le moment d'ordre m de la loi. En déduire la valeur de la variance.

Nous notons $M_m = \mathbb{E}[X^m]$ le moment d'ordre m . Lorsque $m = 0$, nous avons directement $M_0 = \mathbb{E}[X^0] = 1$. Lorsque $m > 0$, nous avons :

$$\begin{aligned} M_m &= \int_{\mathbb{R}} x^m f_\theta(x) dx \\ &= \int_0^{+\infty} x^m \theta e^{-\theta x} dx \\ u(x) &= x^m, & u'(x) &= mx^{m-1}, \\ v'(x) &= \theta e^{-\theta x} & \text{et } v(x) &= -e^{-\theta x}. \\ &= \underbrace{[-x^m e^{-\theta x}]_0^{+\infty}}_{=0} + \int_0^{+\infty} mx^{m-1} e^{-\theta x} dx \\ &= \frac{m}{\theta} \underbrace{\int_0^{+\infty} x^{m-1} \theta e^{-\theta x} dx}_{=M_{m-1}}. \end{aligned}$$

Par récurrence, nous obtenons :

$$M_m = \frac{m}{\theta} M_{m-1}$$

$$\begin{aligned}
&= \frac{m(m-1)}{\theta^2} M_{m-2} \\
&\quad \vdots \\
&= \frac{m!}{\theta^m} M_0 \\
&= \frac{m!}{\theta^m}.
\end{aligned}$$

Enfin, nous calculons la variance en prenant :

$$\begin{aligned}
\mathbb{V}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= M_2 - M_1^2 \\
&= \frac{2!}{\theta^2} - \left(\frac{1!}{\theta^1}\right)^2 \\
&= \frac{1}{\theta^2}.
\end{aligned}$$

1.2. Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ et montrer qu'il est égal à :

$$\hat{\theta}_n = \frac{1}{\bar{X}_n}.$$

Nous commençons par calculer la log-vraisemblance :

$$\begin{aligned}
\log V_{\mathbf{X}}(\theta) &= \log \left(\prod_{i=1}^n f_{\theta}(X_i) \right) \\
&= \sum_{i=1}^n [\log(\theta e^{-\theta X_i})] \\
&= \sum_{i=1}^n [\log \theta - \theta X_i] \\
&= n \log \theta - \theta \sum_{i=1}^n X_i.
\end{aligned}$$

En dérivant, nous avons :

$$\begin{aligned}
\frac{\partial}{\partial \theta} \log V_{\mathbf{X}}(\hat{\theta}_n) = 0 &\Leftrightarrow \frac{n}{\hat{\theta}_n} - \sum_{i=1}^n X_i = 0 \\
&\Leftrightarrow \frac{1}{\hat{\theta}_n} = \frac{1}{n} \sum_{i=1}^n X_i
\end{aligned}$$

$$\Leftrightarrow \hat{\theta}_n = \frac{1}{\bar{X}_n}.$$

1.3. Montrer que $\hat{\theta}_n$ est asymptotiquement normal dont vous préciserez la variance.

D'après le théorème de la limite centrale, nous avons :

$$\sqrt{n} \left(\bar{X}_n - \frac{1}{\theta^*} \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{1}{\theta^{*2}} \right).$$

Ensuite, nous posons la fonction $\ell(x) = \frac{1}{x}$ qui est dérivable sur \mathbb{R}_*^+ donc en $1/\theta^*$ et nous avons :

$$\begin{aligned} \ell' \left(\frac{1}{\theta^*} \right) &= -\frac{1}{\left(\frac{1}{\theta^*} \right)^2} \\ &= -\theta^{*2}. \end{aligned}$$

Donc, nous avons :

- $\sqrt{n} \rightarrow +\infty$,
- ℓ est dérivable en $1/\theta^{*2}$,
- nous avons la loi normale asymptotique.

Donc, d'après la méthode Delta, nous avons :

$$\begin{aligned} \sqrt{n} \left(\hat{\theta}_n - \theta^* \right) &\xrightarrow[n \rightarrow +\infty]{\mathcal{L}} -\theta^{*2} \mathcal{N} \left(0, \frac{1}{\theta^{*2}} \right) \\ &\xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \theta^{*2} \right). \end{aligned}$$

L'estimateur converge en loi vers une gaussienne de moyenne θ^* et de variance θ^{*2} .

1.4. Montrer que, sous certaines conditions que vous préciserez, l'intervalle

$$IC_{1-\alpha}(\theta^*) = \left[\frac{\hat{\theta}_n}{1 + \frac{q_{1-\alpha/2}}{\sqrt{n}}}; \frac{\hat{\theta}_n}{1 - \frac{q_{1-\alpha/2}}{\sqrt{n}}} \right]$$

possède un niveau asymptotique $1 - \alpha$.

D'après la question précédente, nous avons :

$$\frac{\sqrt{n}}{\theta^*} \left(\hat{\theta}_n - \theta^* \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} (0, 1).$$

Étant donnée Z une variable normale centrée réduite et q_α le quantile d'ordre α d'une loi gaussienne centrée réduite donc nous avons :

$$1 - \alpha = \mathbb{P} \left(q_{\alpha/2} \leq Z \leq q_{1-\alpha/2} \right)$$

$$\begin{aligned}
&= \mathbb{P}(-q_{1-\alpha/2} \leq Z \leq q_{1-\alpha/2}) \\
&= \lim_{n \rightarrow +\infty} \mathbb{P}\left(-q_{1-\alpha/2} \leq \frac{\sqrt{n}}{\theta^*} (\hat{\theta}_n - \theta^*) \leq q_{1-\alpha/2}\right) \\
&= \lim_{n \rightarrow +\infty} \mathbb{P}\left(-\frac{q_{1-\alpha/2}}{\sqrt{n}} \leq \frac{\hat{\theta}_n}{\theta^*} - 1 \leq \frac{q_{1-\alpha/2}}{\sqrt{n}}\right) \\
&= \lim_{n \rightarrow +\infty} \mathbb{P}\left(1 - \frac{q_{1-\alpha/2}}{\sqrt{n}} \leq \frac{\hat{\theta}_n}{\theta^*} \leq 1 + \frac{q_{1-\alpha/2}}{\sqrt{n}}\right) \\
&= \lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{1}{1 - \frac{q_{1-\alpha/2}}{\sqrt{n}}} \geq \frac{\theta^*}{\hat{\theta}_n} \geq \frac{1}{1 + \frac{q_{1-\alpha/2}}{\sqrt{n}}}\right) \text{ si } \sqrt{n} > q_{1-\alpha/2}, \\
&= \lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{\hat{\theta}_n}{1 + \frac{q_{1-\alpha/2}}{\sqrt{n}}} \leq \theta^* \leq \frac{\hat{\theta}_n}{1 - \frac{q_{1-\alpha/2}}{\sqrt{n}}}\right).
\end{aligned}$$

Donc, pour n assez grand, nous avons l'intervalle de niveau asymptotique suivant :

$$IC_{1-\alpha}(\theta^*) = \left[\frac{\hat{\theta}_n}{1 + \frac{q_{1-\alpha/2}}{\sqrt{n}}}; \frac{\hat{\theta}_n}{1 - \frac{q_{1-\alpha/2}}{\sqrt{n}}} \right].$$

2 Loi binomiale (7.5 points)

Dans cet exercice, nous étudions un n -échantillon de loi binomiale $\mathcal{Bin}(m, \theta^*)$ de paramètres m (supposé connu) et θ^* que nous cherchons à estimer.

2.5. Calculer l'estimateur $\tilde{\theta}_n$ basé sur le moment d'ordre 1. Montrer qu'il n'est pas possible de proposer un estimateur basé sur la variance sauf si nous savons que θ^* appartient $[0; 1/2]$ ou à $[1/2; 1]$.

Nous savons que si X suit une loi binomiale de paramètre m et θ^* alors il existe Y_1, \dots, Y_m m variables de loi de Bernoulli indépendantes de paramètre θ^* telles que :

$$X = \sum_{j=1}^m Y_j.$$

Donc, nous avons :

$$\mathbb{E}[X] = m\theta^* \text{ et } \mathbb{V}[X] = m\theta^*(1 - \theta^*).$$

À partir de l'espérance, nous obtenons :

$$\bar{X}_n = m\tilde{\theta}_n^{(1)} \Leftrightarrow \tilde{\theta}_n^{(1)} = \frac{1}{m}\bar{X}_n.$$

Pour la variance, nous voyons que nous devons résoudre l'équation :

$$x^2 - x + \frac{\mathbb{V}[X]}{m} = 0.$$

Pour cela, nous calculons $\Delta = 1 - 4\frac{\mathbb{V}[X]}{m}$ et nous obtenons deux solutions :

$$x_1 = \frac{1 - \sqrt{1 - 4\frac{\mathbb{V}[X]}{m}}}{2} \text{ et } x_2 = \frac{1 + \sqrt{1 - 4\frac{\mathbb{V}[X]}{m}}}{2}$$

et les deux estimateurs sont possibles puisque $0 \leq \sqrt{1 - 4\frac{\mathbb{V}[X]}{m}} \leq 1$ car $0 \leq \mathbb{V}[X] \leq m/4$: le premier si nous savons que θ^* appartient à $[1/2; 1]$ et le second si nous savons que θ^* appartient à $[0; 1/2]$.

2.6. Montrer que $\tilde{\theta}_n$ est sans biais. Est-il asymptotiquement sans biais ?

Nous avons :

$$\begin{aligned} \mathbb{E}[\tilde{\theta}_n] &= \mathbb{E}\left[\frac{1}{m}\bar{X}_n\right] \\ &= \frac{1}{m}\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] \\ &= \frac{1}{mn}\sum_{i=1}^n \mathbb{E}[X_i] \\ &= \frac{1}{mn}nm\theta^* \\ &= \theta^*. \end{aligned}$$

Donc l'estimateur est sans biais donc il est asymptotiquement sans biais.

2.7. Donner la loi exacte de $\tilde{\theta}_n$ et montrer qu'il est asymptotiquement normal.

Nous savons qu'il existe Y_1, \dots, Y_{nm} nm variables de Bernoulli indépendantes de paramètres θ^* telles que :

$$nm\tilde{\theta}_n = \sum_{j=1}^{nm} Y_j$$

donc $nm\tilde{\theta}_n$ suit une loi binomiale de paramètres nm et θ^* .

D'un autre côté, par le théorème de la limite centrale, nous avons :

$$\sqrt{nm} \left(\tilde{\theta}_n - \mathbb{E}[Y_i] \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \mathbb{V}[Y_i])$$

donc nous avons :

$$\sqrt{n} \left(\tilde{\theta}_n - \theta^* \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{\theta^*(1 - \theta^*)}{m}\right).$$

2.8. Montrer que l'estimateur $\tilde{\theta}_n$ est un *UMVU*.

D'après la question 1.6, l'estimateur est non-biaisé donc il reste à calculer la borne de Cramer-Rao. Comme nous avons un n -échantillon, nous avons :

$$I_n(\theta^*) = nI_1(\theta^*).$$

Et, nous savons que :

$$\begin{aligned} I_1(\theta) &= -\mathbb{E} \left[\frac{\partial^2}{d\theta^2} \log(f_\theta(X)) \right] \\ &= -\mathbb{E} \left[\frac{\partial^2}{d\theta^2} \log \left(\binom{m}{X} \theta^X (1-\theta)^{m-X} \right) \right] \\ &= -\mathbb{E} \left[\frac{\partial^2}{d\theta^2} \left(\log \binom{m}{X} + X \log \theta + (m-X) \log(1-\theta) \right) \right] \\ &= -\mathbb{E} \left[\frac{\partial}{d\theta} \left(X \frac{1}{\theta} - (m-X) \frac{1}{1-\theta} \right) \right] \\ &= -\mathbb{E} \left[-X \frac{1}{\theta^2} - (m-X) \frac{1}{(1-\theta)^2} \right] \\ &= \mathbb{E} \left[X \frac{1}{\theta^2} + m \frac{1}{(1-\theta)^2} - X \frac{1}{(1-\theta)^2} \right] \\ &= \mathbb{E} \left[X \frac{(1-\theta)^2 - \theta^2}{[\theta(1-\theta)]^2} \right] + m \frac{1}{(1-\theta)^2} \\ &= \mathbb{E}[X] \frac{1-2\theta+\theta^2-\theta^2}{[\theta(1-\theta)]^2} + m \frac{1}{(1-\theta)^2} \\ &= m \left[\frac{1-2\theta}{\theta(1-\theta)^2} + \frac{1}{(1-\theta)^2} \right] \\ &= \frac{m}{\theta(1-\theta)^2} [1-2\theta+\theta] \\ &= \frac{m}{\theta(1-\theta)}. \end{aligned}$$

Donc nous avons :

$$I_n(\theta^*) = \frac{mn}{\theta^*(1-\theta^*)}.$$

D'un autre côté, nous avons :

$$\begin{aligned} \mathbb{V}[\tilde{\theta}_n] &= \mathbb{V} \left[\frac{1}{nm} \sum_{j=1}^{nm} Y_j \right] \\ &= \frac{1}{n^2 m^2} \sum_{j=1}^{nm} \mathbb{V}[Y_j] \\ &= \frac{1}{n^2 m^2} \sum_{j=1}^{nm} [\theta^*(1-\theta^*)] \end{aligned}$$

$$= \frac{\theta^* (1 - \theta^*)}{nm}$$

Donc, nous avons :

$$\mathbb{V} [\tilde{\theta}_n] = \frac{\theta^* (1 - \theta^*)}{nm} = \frac{1}{I_n(\theta^*)}.$$

Donc la borne de Cramer-Rao est atteinte et l'estimateur est bien un UMVU.

3 Loi géométrique (5 points)

Dans cet exercice, nous étudions la loi géométrique de paramètre $\theta^* \in]0; 1[$ dont nous rappelons la densité définie pour tout $k \in \mathbb{N}^*$ par :

$$f_{\theta^*}(k) = (1 - \theta^*)^{k-1} \theta^*.$$

Nous cherchons à tester les hypothèses :

$$\begin{cases} \mathcal{H}_0 : \theta^* = 1/4, \\ \mathcal{H}_1 : \theta^* = 3/4. \end{cases}$$

3.9. Donner les ensembles Θ_0 et Θ_1 associés au test.

Nous commençons par remarquer que $\Theta_0 = \{1/4\}$ et $\Theta_1 = \{3/4\}$ qui sont des singletons.

3.10. Montrer que la statistique du rapport de vraisemblance $h(\mathbf{X})$ est égale $g_n(\bar{X}_n)$ avec :

$$\begin{aligned} g_n : \mathbb{R}_*^+ &\rightarrow \mathbb{R}^+ \\ x &\mapsto 9^n 3^{-nx}. \end{aligned}$$

Nous calculons la vraisemblance :

$$\begin{aligned} V_{\mathbf{X}}(\theta) &= \prod_{i=1}^n f_{\theta}(X_i) \\ &= \prod_{i=1}^n [(1 - \theta)^{X_i-1} \theta] \\ &= \left[\frac{\theta}{1 - \theta} \right]^n (1 - \theta)^{\sum_{i=1}^n X_i} \\ &= \left[\frac{\theta}{1 - \theta} \right]^n (1 - \theta)^{n\bar{X}_n} \end{aligned}$$

Ainsi, la fonction de la statistique du rapport de vraisemblance est de la forme :

$$\begin{aligned}
 h(\mathbf{X}, 1/4, 3/4) &= \frac{V_{\mathbf{X}}(3/4)}{V_{\mathbf{X}}(1/4)} \\
 &= \frac{\left[\frac{3/4}{1-3/4}\right]^n (1-3/4)^{n\bar{X}_n}}{\left[\frac{1/4}{1-1/4}\right]^n (1-1/4)^{n\bar{X}_n}} \\
 &= \frac{\left[\frac{3/4}{1/4}\right]^n (1/4)^{n\bar{X}_n}}{\left[\frac{1/4}{3/4}\right]^n (3/4)^{n\bar{X}_n}} \\
 &= 9^n 3^{-n\bar{X}_n}.
 \end{aligned}$$

3.11. En admettant que l'espérance d'une loi géométrique de paramètre p est $1/p$ et sa variance $(1-p)/p^2$, proposer un test uniformément plus puissant de niveau asymptotique $1-\alpha$.

Nous savons qu'il existe $k_\alpha > 0$ tel que :

$$\begin{aligned}
 \Phi(\mathbf{X}) &= \mathbb{1}_{\{g_n(\bar{X}_n) > k_\alpha\}} \\
 &= \mathbb{1}_{\{9^n 3^{-n\bar{X}_n} > k_\alpha\}} \\
 &= \mathbb{1}_{\{3^{-n\bar{X}_n} > k_\alpha 9^{-n}\}} \\
 &= \mathbb{1}_{\{-n\bar{X}_n \ln(3) > \ln(k_\alpha) - n \ln(9)\}} \\
 &= \mathbb{1}_{\{\bar{X}_n \ln(3) < -\frac{1}{n} \ln(k_\alpha) + 2 \ln(3)\}} \\
 &= \mathbb{1}_{\{\bar{X}_n < -\frac{1}{n \ln(3)} \ln(k_\alpha) + 2\}} \\
 &= \mathbb{1}_{\{\bar{X}_n < \ell_\alpha\}}
 \end{aligned}$$

où $\ell_\alpha = -\frac{1}{n \ln(3)} \ln(k_\alpha) + 2$. D'après le théorème de la limite centrale, nous avons sous l'hypothèse \mathcal{H}_0 la loi asymptotique suivante :

$$\begin{aligned}
 \sqrt{n} \left(\bar{X}_n - \frac{1}{1/4} \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{1-1/4}{[1/4]^2} \right) &\Leftrightarrow \sqrt{n} (\bar{X}_n - 4) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, 4^2 \times \frac{3}{4} \right) \\
 &\Leftrightarrow \sqrt{n} (\bar{X}_n - 4) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} (0, 12) \\
 &\Leftrightarrow \frac{\sqrt{n}}{2\sqrt{3}} (\bar{X}_n - 4) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} (0, 1).
 \end{aligned}$$

Nous avons donc :

$$\mathbb{P}(\Phi(\mathbf{X}) = 1) = \mathbb{P}(\bar{X}_n < \ell_\alpha)$$

$$\begin{aligned}
&= \mathbb{P}(\bar{X}_n - 4 < \ell_\alpha - 4) \\
&= \mathbb{P}\left(\frac{\sqrt{n}}{2\sqrt{3}}(\bar{X}_n - 4) < \frac{\sqrt{n}}{2\sqrt{3}}(\ell_\alpha - 4)\right)
\end{aligned}$$

donc si nous posons

$$\begin{aligned}
q_\alpha = \frac{\sqrt{n}}{2\sqrt{3}}(\ell_\alpha - 4) &\Leftrightarrow \frac{2\sqrt{3}}{\sqrt{n}}q_\alpha = \ell_\alpha - 4 \\
&\Leftrightarrow \frac{2\sqrt{3}}{\sqrt{n}}q_\alpha + 4 = \ell_\alpha.
\end{aligned}$$

Nous obtenons la statistique

$$\Phi(\mathbf{X}) = \mathbf{1}_{\{\bar{X}_n < \frac{2\sqrt{3}}{\sqrt{n}}q_\alpha + 4 = \ell_\alpha\}}$$

qui est un test uniformément plus puissant de niveau asymptotique $1 - \alpha$ par le lemme de Neyman-Pearson.

3.12. (bonus) Calculer l'espérance et la variance admise à la question précédente.

Nous avons :

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{+\infty} k(1 - \theta^*)^{k-1}\theta^* \\
&= \theta^* \sum_{k=1}^{+\infty} k(1 - \theta^*)^{k-1}.
\end{aligned}$$

Or, nous savons que pour tout $\theta^* \in]0, 1[$:

$$\frac{1}{\theta^*} = \sum_{k=0}^{+\infty} (1 - \theta^*)^k$$

donc nous avons :

$$-\frac{1}{\theta^{*2}} = -\sum_{k=0}^{+\infty} k(1 - \theta^*)^{k-1}$$

donc

$$\frac{1}{\theta^{*2}} = \sum_{k=1}^{+\infty} k(1 - \theta^*)^{k-1}$$

ce qui donne :

$$\mathbb{E}[X] = \theta^* \sum_{k=1}^{+\infty} k(1 - \theta^*)^{k-1}$$

$$\begin{aligned}
&= \theta^* \frac{1}{\theta^{*2}} \\
&= \frac{1}{\theta^*}
\end{aligned}$$

De même, nous avons :

$$\begin{aligned}
\frac{-2}{\theta^{*3}} &= - \sum_{k=2}^{+\infty} k(k-1) (1-\theta^*)^{k-2} \\
\ell - 1 = k - 2, \quad k &= \ell + 1 \\
&= - \sum_{\ell=1}^{+\infty} \ell(\ell+1) (1-\theta^*)^{\ell-1} \\
&= - \sum_{\ell=1}^{+\infty} \ell^2 (1-\theta^*)^{\ell-1} - \sum_{\ell=1}^{+\infty} \ell (1-\theta^*)^{\ell-1} \\
&= - \sum_{\ell=1}^{+\infty} \ell^2 (1-\theta^*)^{\ell-1} - \frac{1}{\theta^{*2}}
\end{aligned}$$

ce qui donne :

$$\sum_{\ell=1}^{+\infty} \ell^2 (1-\theta^*)^{\ell-1} = \frac{2}{\theta^{*3}} - \frac{1}{\theta^{*2}}.$$

Donc, nous avons :

$$\begin{aligned}
\mathbb{V}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \sum_{k=1}^{+\infty} k^2 (1-\theta^*)^{k-1} \theta^* - \left(\frac{1}{\theta^*}\right)^2 \\
&= \theta^* \left[\frac{2}{\theta^{*3}} - \frac{1}{\theta^{*2}} \right] - \left(\frac{1}{\theta^*}\right)^2 \\
&= \frac{2}{\theta^{*2}} - \frac{\theta^*}{\theta^{*2}} - \frac{1}{\theta^{*2}} \\
&= \frac{1-\theta^*}{\theta^{*2}}
\end{aligned}$$