

- EXERCICE 1 -

Dans cet exercice R désigne un réel strictement positif fixé.

On considère une variable aléatoire réelle X de densité f définie par : $f(t) = \begin{cases} 0 & \text{si } t \notin [0; R] \\ \frac{2t}{R^2} & \text{si } t \in [0; R] \end{cases}$.

Soit $n \in \mathbb{N}^*$. On considère $(X_1, X_2, \dots, X_n)$ un n -échantillon de X . On cherche à estimer le réel R .

1. Déterminer la fonction de répartition de X , notée F_X .
2. Montrer que X admet une espérance et une variance et que $E(X) = \frac{2R}{3}$ et $V(X) = \frac{R^2}{18}$.
3. On note $\overline{X}_n$ la moyenne empirique de X et on considère la variable aléatoire $T_n = \frac{3}{2}\overline{X}_n$.
 - (a) Montrer que T_n est un estimateur de R .
 - (b) Calculer $E(T_n)$ et $V(T_n)$.
 - (c) Montrer à l'aide de l'inégalité de Bienaymé-Tchebychev que T_n est un estimateur convergent de R .

- EXERCICE 2 -

Pour tout entier $m \geq 1$, on considère un échantillon $(Y_1, \dots, Y_m)$ de la loi de Bernoulli de paramètre p .

On rappelle que $\overline{Y}_m = \frac{1}{m} \sum_{i=1}^m Y_i$.

1. A l'aide de l'inégalité de Bienaymé-Tchebychev, montrer que : $P(\overline{Y}_m - \epsilon \leq p \leq \overline{Y}_m + \epsilon) \geq 1 - \frac{1}{4m\epsilon^2}$.
2. En déduire que $\left[\overline{Y}_m - \sqrt{\frac{5}{m}}; \overline{Y}_m + \sqrt{\frac{5}{m}} \right]$ est un intervalle de confiance de p au niveau de confiance 0,95.

- EXERCICE 3 -

Soit a un réel strictement positif et X une variable aléatoire vérifiant $E(X^2) = 2a$ et $V(X^2) = 4a^2$.

On considère n variables aléatoires $(X_1, X_2, \dots, X_n)$ indépendantes et suivant toutes la même loi que X .

On note S_n la variable aléatoire définie par : $S_n = \frac{1}{2n} \sum_{k=1}^n X_k^2$

1. Justifier que S_n est un estimateur de a .
2. A l'aide de l'inégalité de Bienaymé-Tchebychev, montrer que S_n est un estimateur convergent de a .
3. On suppose que a est inférieur ou égal à 1.
 - (a) Montrer que : $\forall \epsilon > 0, P(|S_n - a| \leq \epsilon) \geq 1 - \frac{1}{n\epsilon^2}$.
 - (b) Déterminer une valeur de n pour laquelle $\left[S_n - \frac{1}{10}; S_n + \frac{1}{10} \right]$ est un intervalle de confiance de a au niveau de confiance au moins égal à 95%.

- EXERCICE 4 -

Soit a un réel strictement positif et f la fonction définie par : $f : x \mapsto \begin{cases} 0 & \text{si } x < 1 \\ \frac{a}{x^{a+1}} & \text{si } x \geq 1 \end{cases}$.

On admet que f est une densité de probabilité et on note X une variable aléatoire admettant f pour densité. Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes, toutes de même loi que X .

1. Déterminer la fonction de répartition F_X de X .
2. Soit $n \in \mathbb{N}^*$. On pose : $W_n = \ln(X_n)$.
Montrer que la variable aléatoire W_n suit une loi exponentielle de paramètre a .
En déduire l'espérance et la variance de W_n .
3. On définit, pour tout n de $\mathbb{N}^*$: $M_n = \frac{\ln(X_1) + \dots + \ln(X_n)}{n}$ et $T_n = \sqrt{n}(aM_n - 1)$.
 - (a) Montrer que la variable aléatoire M_n admet une espérance et une variance et les calculer.
 - (b) Justifier que la suite de variables aléatoires $(T_n)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.
 - (c) En déduire que l'intervalle $\left[\frac{\sqrt{n}-2}{\sqrt{n}M_n}; \frac{\sqrt{n}+2}{\sqrt{n}M_n} \right]$ est un intervalle de confiance asymptotique pour a au niveau de confiance 95%.
On admettra que $\Phi(2) \geq 0,975$, où Φ désigne la fonction de répartition d'une variable aléatoire suivant la loi normale centrée réduite.

- EXERCICE 5 -

Soit $(X_1, \dots, X_n)$ un n -échantillon i.i.d de loi parente la loi normale $\mathcal{N}(m, \sigma^2)$ de variance σ^2 connue. i.i.d signifie indépendantes et identiquement distribuées.

On cherche un intervalle de confiance pour l'espérance m et on considère pour cela l'estimateur $T_n = \overline{X}_n$.

1. Calculer l'espérance et la variance de T_n .
2. Utilisation de l'inégalité de BT.
 - (a) Montrer, à l'aide de l'inégalité de BT, que pour tout $t > 0$, on a : $P(T_n - t \leq m \leq T_n + t) \geq 1 - \frac{\sigma^2}{nt^2}$.
 - (b) En déduire que l'intervalle $I_n(\alpha) = \left[T_n - \frac{\sigma}{\sqrt{n\alpha}}; T_n + \frac{\sigma}{\sqrt{n\alpha}} \right]$ est un intervalle de confiance de m au niveau de confiance $1 - \alpha$.
3. Utilisation de la loi normale centrée réduite.
 - (a) Montrer que T_n suit la loi normale $\mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$.
 - (b) En déduire que pour tout $t > 0$: $P\left(\left|\sqrt{\frac{n}{\sigma^2}}(T_n - m)\right| \leq t\right) = 2\Phi(t) - 1$.
 - (c) Justifier que pour tout $\alpha \in]0, 1[$, il existe un réel t_α tel que : $2\Phi(t) - 1 = 1 - \alpha$.
 - (d) En déduire que l'intervalle $J_n(\alpha) = \left[T_n - \frac{t_\alpha \sigma}{\sqrt{n}}; T_n + \frac{t_\alpha \sigma}{\sqrt{n}} \right]$ est un intervalle de confiance de m au niveau de confiance $1 - \alpha$.
4.
 - (a) Déterminer l'étendue des intervalles $I_n(\alpha)$ et $J_n(\alpha)$ (i.e leur longueur).
 - (b) On donne les valeurs numériques de t_α et $\frac{1}{\sqrt{\alpha}}$ suivantes pour différentes valeurs de α .

α	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
$1/\sqrt{\alpha}$	10.	7.071	5.773	5.	4.472	4.082	3.779	3.535	3.333	3.162
t_α	2.575	2.326	2.170	2.053	1.959	1.880	1.811	1.750	1.695	1.644

Au vu de ces résultats, déterminer quel est l'intervalle le plus précis entre $I_n(\alpha)$ et $J_n(\alpha)$?

5. On suppose que le temps de travail hebdomadaire d'un étudiant en classe prépa est une variable aléatoire suivant la loi normale $\mathcal{N}(m, 100)$ de moyenne m inconnue que l'on cherche à estimer.
Une enquête réalisée auprès de 100 étudiants donne un temps de travail hebdomadaire moyen de 29.32h.
 - (a) Déterminer l'intervalle de confiance $J_n(0.05)$ au niveau de confiance $\alpha = 0.05$ pour la moyenne m .
 - (b) Quelle valeur de n faut-il choisir pour que la longueur de l'intervalle de confiance au niveau de confiance $\alpha = 0.05$ soit inférieure à 2?

- EXERCICE 6 -

Soit $(X_1, \dots, X_n)$ un n -échantillon d'une variable aléatoire X admettant pour espérance m et pour variance σ^2 .

1. On suppose que m est connu.

Montrer que $T_n = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$ est un estimateur de σ^2 et calculer $E(T_n)$.

2. On suppose que m n'est pas connu.

On note $\overline{X}_n$ la moyenne empirique de l'échantillon et $U_n = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2$.

(a) Montrer que, pour tout $i \in \llbracket 1; n \rrbracket$: $E\left((X_i - \overline{X}_n)^2\right) = V(X_i - \overline{X}_n)$.

(b) Montrer que, pour tout $i \in \llbracket 1; n \rrbracket$: $V(X_i - \overline{X}_n) = \left(1 - \frac{1}{n}\right)^2 V(X_i) + \frac{1}{n^2} \sum_{k \neq i} V(X_k)$.

(c) En déduire que : $V(X_i - \overline{X}_n) = \frac{n-1}{n} \sigma^2$.

(d) Montrer alors que U_n est un estimateur de σ^2 et calculer $E(U_n)$.

(e) En déduire un estimateur V_n de σ^2 qui vérifie $E(V_n) = \sigma^2$.

- EXERCICE 7 -

Résultat préliminaire.

1. Soit N un entier supérieur à 2 et Y une variable aléatoire à valeurs dans $\llbracket 1; N \rrbracket$.

Pour $k \in \llbracket 1; N \rrbracket$, montrer que $P(Y = k) = P(Y \geq k) - P(Y \geq k+1)$.

En déduire que pour toute variable aléatoire Y à valeurs dans $\llbracket 1; N \rrbracket$ on a : $E(Y) = \sum_{k=1}^N P(Y \geq k)$.

Une urne contient N boules numérotées de 1 à N . On sait que $N \geq 2$ mais on ne connaît pas sa valeur exacte et on cherche à l'estimer. Pour cela, on effectue n tirages avec remise et on note, pour tout $k \in \llbracket 1; n \rrbracket$, Z_k le numéro de la boule obtenue au k -ième tirage. Ainsi, les variables $(Z_1, \dots, Z_n)$ sont mutuellement indépendantes.

2. Pour tout $k \in \llbracket 1; n \rrbracket$, déterminer la loi de Z_k . Donner son espérance et sa variance.

3. On pose $S_n = \max(Z_1, \dots, Z_n)$.

(a) Calculer pour tout $k \in \llbracket 1; N \rrbracket$, $P(S_n \leq k)$.

(b) Montrer à l'aide du résultat préliminaire que : $E(S_n) = N - \sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^n$.

(c) Montrer que pour tout $k \in \llbracket 0; N-1 \rrbracket$, on a : $\int_{k/N}^{(k+1)/N} x^n dx \geq \left(\frac{k}{N}\right)^n$.

(d) En déduire que : $\frac{1}{N} \sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^n \leq \int_0^1 x^n dx$.

(e) Montrer que : $E(S_n) \geq N - \frac{N}{n+1}$ puis en déduire que $\lim_{n \rightarrow +\infty} E(S_n) = 0$.
On dit que S_n est un estimateur asymptotiquement sans biais de N .