Fiche 10 d'exercices : Estimation

1 Exercices d'assimilation du cours

Exercice 1 (Espérance et variance d'une somme de variables aléatoires indépendantes)

On considère une suite $(X_k)_{k\in\mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes, dont la loi commune est donnée pour tout $k \in \mathbb{N}^*$ par :

$$P(X_k = 1) = \frac{1}{3} \quad \text{et} \quad P(X_k = 2) = \frac{2}{3}.$$

Pour tout $n \in \mathbb{N}^*$, on pose $S_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$.

1) Calculer pour tout $n \in \mathbb{N}^*$, $E(S_n)$.
2) Calculer pour tout $n \in \mathbb{N}^*$, $V(S_n)$.

Exercice 2 (Estimation du paramètre d'une loi de Bernoulli)

On considère un dé non équilibré que l'on lance 500 fois et on compte le nombre de fois où l'on obtient les différentes faces du dé.

Face du dé	1	2	3	4	5	6
Nombre d'apparitions	110	87	42	69	91	101

On considère la variable aléatoire X qui vaut 1 si le dé donne 6 et 0 sinon.
1) Reconnaître la loi suivie par X.
2) Donner une estimation à $0,01$ du paramètre de cette loi.

Exercice 3 (Estimation du paramètre d'une loi de Poisson)

Le nombre de fautes de frappe dans un livre de Mathématiques de 300 pages peut être modélisé par une variable aléatoire X suivant une loi de poisson de paramètre λ inconnu, que l'on cherche à estimer.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
6	1	5	5	7	5	10	8	5	6

x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}
6	5	3	6	8	9	2	6	7	7

On suppose que l'on dispose d'un échantillon de 20 réalisations $x_1, \ldots, x_{20}$ de cette variable X. Donner une estimation à $0,01$ près du paramètre λ.

Exercice 4 (Intervalle de confiance)

On suppose que le paramètre p d'une loi de Bernoulli est inconnu et on cherche à l'estimer. Pour ce faire, on considère une suite de variables aléatoires $(X_k)_{k\in\mathbb{N}^*}$ indépendantes, suivant toutes la loi de Bernoulli de paramètre p. Pour tout $n \in \mathbb{N}^*$, on pose $S_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$.

1) a) Montrer que $E(S_n) = p$.
1) b) Calculer la variance de S_n.

2) a) On admet que pour tout $p \in [0,1], p(1-p) \leq \dfrac{1}{4}$. A l'aide de l'inégalité de Bienaymé-Tchebychev, montrer que pour tout $\varepsilon > 0$,

$$P(|S_n - p| \leq \varepsilon) \geq 1 - \frac{1}{4n\varepsilon^2}.$$

2) b) Montrer que $|S_n - p| \leq \varepsilon \iff p \in [S_n - \varepsilon, S_n + \varepsilon]$.

2) c) En déduire que l'intervalle $\left[S_n - \sqrt{\frac{5}{n}}, S_n + \sqrt{\frac{5}{n}}\right]$ est un intervalle de confiance de p au niveau de confiance $0,95$.

2 Exercices d'entraînement

Exercice 5 (Espérance et variance d'une somme de variables aléatoires indépendantes)

On considère une suite $(X_k)_{k\in\mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes, dont la loi commune est donnée pour tout $k \in \mathbb{N}^*$ par le tableau suivant.

x	0	1	2
$P(X_k = x)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

Pour tout $n \in \mathbb{N}^*$, on pose $S_n = \dfrac{3}{4n}\displaystyle\sum_{k=1}^{n} X_k$.

1) Calculer pour tout $n \in \mathbb{N}^*$, $E(S_n)$.
2) Calculer pour tout $n \in \mathbb{N}^*$, $V(S_n)$.

Exercice 6 (Application de l'inégalité de Bienaymé-Tchebychev)

Soient p un réel appartenant à $]0,1[$ et $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes qui suivent toutes une loi de Bernoulli de paramètre p. On pose pour tout $n \in \mathbb{N}^*$, $S_n = \displaystyle\sum_{k=1}^{n} X_k$.

Montrer que pour tout entier $n \geq 1$, $\forall \varepsilon > 0, P\left(\left|\dfrac{S_n}{n} - p\right| \geq \varepsilon\right) \leq \dfrac{p(1-p)}{n\varepsilon^2}$.

Exercice 7 (Application de la loi faible des grands nombres)

On considère une suite de lancers indépendants d'une même pièce équilibrée. Pour tout $i \in \mathbb{N}^*$, on note X_i la variable aléatoire égale à 1 si l'événement « obtenir pile » est réalisé au $i^{\text{ième}}$ lancer et 0 sinon. Pour tout $n \in \mathbb{N}^*$, on pose $\overline{X}_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$.

1) Que représente $\overline{X}_n$?

2) Montrer que : $\forall \varepsilon > 0$, $\displaystyle\lim_{n\to+\infty} P\left(\left|\overline{X}_n - \frac{1}{2}\right| < \varepsilon\right) = 1$. Le résultat précédent est-il conforme à l'intuition ?

Exercice 8 (Intervalle de confiance)

Soit a un réel strictement positif.

1) Montrer que $\displaystyle\int_a^{+\infty} \frac{1}{t^2}\mathrm{d}t = \frac{1}{a}$, $\displaystyle\int_a^{+\infty} \frac{1}{t^3}\mathrm{d}t = \frac{1}{2a^2}$ et $\displaystyle\int_a^{+\infty} \frac{1}{t^4}\mathrm{d}t = \frac{1}{3a^3}$.

On considère la fonction f définie sur $\mathbb{R}$ par :

$$f(t) = \begin{cases} \frac{3a^3}{t^4} & \text{si } t \geq a, \\ 0 & \text{si } t < a. \end{cases}$$

2) Montrer que f peut être considérée comme une densité de probabilité.

Un fabriquant de téléphones portables veut étudier l'autonomie de ses téléphones. Etant donné un téléphone pris au hasard chargé au maximum, on note X le nombre d'heures écoulées lorsque le téléphone s'éteint. On admet que X est une variable aléatoire de densité f.

3) Montrer que $E(X) = \dfrac{3a}{2}$ et que $V(X) = \dfrac{3a^2}{4}$.

4) On cherche à estimer le paramètre a. Pour cela, on allume en même temps n téléphones pris au hasard que l'on a chargés au maximum, n étant un entier naturel non nul. On note pour tout i compris entre 1 et n, X_i la durée en heures écoulées lorsque le $i^{\text{ième}}$ téléphone s'éteint. On suppose que les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes. On pose $M_n = \dfrac{2}{3n} \displaystyle\sum_{i=1}^{n} X_i$.

a) Calculer $E(M_n)$ et $V(M_n)$.

b) Soit ε un réel strictement positif. Montrer en utilisant l'inégalité de Bienaymé-Tchebychev que :
$P(|M_n - a| > \varepsilon) \leq \dfrac{a^2}{3n\varepsilon^2}$.

c) Soient $\alpha \in]0,1[$ et $\varepsilon = \dfrac{a}{\sqrt{3\alpha n}}$. Montrer que $[M_n - \varepsilon, M_n + \varepsilon]$ est un intervalle de confiance de a au niveau de confiance $1 - \alpha$.

3 Exercices d'approfondissement

Exercice 9 (Intervalle de confiance)

Soit a un réel tel que $-\dfrac{1}{2} \leq a \leq \dfrac{1}{2}$. Soit f la fonction définie sur $\mathbb{R}$ par :

$$f(t) = \begin{cases} \frac{1-a}{2} & \text{si } -1 \leq t \leq 0, \\ \frac{1+a}{2} & \text{si } 0 < t \leq 1, \\ 0 & \text{sinon.} \end{cases}$$

1) Vérifier que f est une densité de probabilité. Dans la suite de l'exercice, on note X une variable aléatoire admettant f comme densité et on suppose que le paramètre a est inconnu.

2) Calculer $E(X)$ et montrer que $V(X) = \dfrac{4 - 3a^2}{12}$.

3) Pour un entier $n \in \mathbb{N}^*$, on considère n variables aléatoires $X_1, X_2, \ldots, X_n$ indépendantes et suivant la même loi que X. Pour tout $n \geq 1$, on définit $\overline{X}_n = \dfrac{1}{n} \displaystyle\sum_{k=1}^{n} X_k$ puis $Y_n = 2\overline{X}_n$.

a) Montrer que $E(Y_n) = a$ et que $V(Y_n) = \dfrac{4 - 3a^2}{3n}$.

b) Soit ε un réel strictement positif. En appliquant l'inégalité de Bienaymé-Tchebychev à Y_n, établir l'inégalité :

$$P(|Y_n - a| \leq \varepsilon) \geq 1 - \frac{4}{3n\varepsilon^2}.$$

Exercice 10 (Intervalle de confiance)

Soit a un réel tel que $0 \leq a \leq 1$. Soit f la fonction définie sur $\mathbb{R}$ par :

$$f(t) = \begin{cases} \frac{t}{a^2} & \text{si } 0 \leq t \leq a, \\ \frac{2a-t}{a^2} & \text{si } a < t \leq 2a, \\ 0 & \text{sinon.} \end{cases}$$

1) Montrer que f est continue sur $\mathbb{R}$.
2) Montrer que f peut être considérée comme une densité de probabilité.
3) a) Montrer que X possède une espérance et que $E(X) = a$.
b) Montrer que X^2 possède une espérance et que $E(X^2) = \dfrac{7a^2}{6}$.
c) En déduire que X possède une variance et la déterminer.
4) Dans cette question, on suppose que le paramètre a est inconnu et on souhaite en faire une estimation. A cet effet, pour tout entier n supérieur ou égal à 2, on considère n variables aléatoires $X_1, \ldots, X_n$ indépendantes, suivant toutes la même loi que X et on pose $\overline{X}_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$.
Déterminer l'espérance et la variance de la variable aléatoire $\overline{X}_n$.
5) On note ε un réel strictement positif.
a) Ecrire l'inégalité de Bienaymé-Tchebychev pour $\overline{X}_n$ puis établir l'inégalité :

$$P(|\overline{X}_n - a| \geq \varepsilon) \leq \frac{1}{6n\varepsilon^2}.$$

b) En déduire l'inégalité :

$$P(\overline{X}_n - \varepsilon \leq a \leq \overline{X}_n + \varepsilon) \geq 1 - \frac{1}{6n\varepsilon^2}.$$

c) Donner un intervalle de confiance de a au niveau de confiance 90%.