Compléments de probabilités
Exercices

Dans toute cette feuille, on note Φ la fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0,1)$. On en donne un tableau de valeurs à la fin de la feuille.

Convergences en loi

Exercice 1. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires iid, de loi commune $\mathcal{B}(p)$ (avec $p\in]0,1[$). On pose : $\forall\, n\in\mathbb{N}^*,\ T_n=\prod_{i=1}^{n}X_i$.

1. Donner la loi de T_n.
2. Montrer que $(T_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable constante égale à 0.

Exercice 2. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires iid, de loi $\mathcal{U}([0,1])$. On pose :

$$\forall\, n\in\mathbb{N}^*,\ I_n=\min(X_1,\dots,X_n)\quad\text{et}\quad S_n=\max(X_1,\dots,X_n)$$

On a vu dans le TD précédent que les fonctions de répartition de I_n et S_n sont :

$$F_{I_n}(x)=\begin{cases}0 \text{ si } x<0\\ 1-(1-x)^n \text{ si } 0\leqslant x\leqslant 1\\ 1 \text{ si } x>1\end{cases}\qquad\text{et}\qquad F_{S_n}(x)=\begin{cases}0 \text{ si } x<0\\ x^n \text{ si } 0\leqslant x\leqslant 1\\ 1 \text{ si } x>1\end{cases}$$

1. Montrer que (I_n) converge en loi vers une variable constante égale à 0.
2. Montrer que (S_n) converge en loi vers une variable constante égale à 1.
3. Reprendre l'exercice avec les $(X_n)_{n\in\mathbb{N}^*}$ iid, de loi $\mathcal{U}([a,b])$

Exercice 3. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telle que : $\forall\, n\in\mathbb{N}^*,\ X_n\hookrightarrow\mathcal{N}\left(0,\frac{1}{n}\right)$.

1. Exprimer, pour tout réel t, $F_{X_n}(t)$ en fonction de Φ.
2. Montrer que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire constante égale à 0.

Exercice 4. Soit (X_n) une suite de va indépendantes, telles que : $\forall\, n\in\mathbb{N}^*,\ X_n\hookrightarrow\mathcal{E}\left(\frac{1}{n}\right)$.
On note $Z_n=X_n-\lfloor X_n\rfloor$ (partie fractionnaire de X_n).
On a déjà vu, dans la feuille sur les variables à densité (exercice 9), que :

$$\forall\, x\in[0,1],\ P(Z_n\leqslant x)=\frac{1-e^{-x/n}}{1-e^{-1/n}}$$

Montrer que (Z_n) converge en loi vers une variable aléatoire Z suivant $\mathcal{U}([0,1])$.

Exercice 5. On considère une urne constituée de $2n$ boules ($n \geqslant 2$), dont la moitié sont noires, et l'autre moitié sont blanches. On pioche dans cette urne deux boules successivement et sans remise, et on note X_n le nombre de boules blanches obtenues.

1. Déterminer la loi de X_n.
2. Montrer que la suite $(X_n)_{n\geqslant 2}$ converge en loi vers une variable aléatoire Z suivant la loi binomiale $\mathscr{B}\left(2, \frac{1}{2}\right)$.
3. Interpréter ce résultat.

Exercice 6. Soit $(U_n)_{n\in\mathbb{N}^*}$ une suite de variables iid suivant la loi $\mathscr{U}([0,\theta])$ avec $\theta > 0$. Pour $n \in \mathbb{N}^*$, on note

$$X_n = \max(U_1, \dots, U_n)$$

1. Déterminer la fonction de répartition de X_n.
2. (a) Soit $\varepsilon > 0$. Montrer que
$$P\big(|X_n - \theta| > \varepsilon\big) = \begin{cases} 0 & \text{si } \varepsilon \geqslant \theta \\ \left(1 - \dfrac{\varepsilon}{\theta}\right)^n & \text{si } \varepsilon < \theta \end{cases}$$
   (b) En déduire, pour tout $\varepsilon > 0$,
$$\lim_{n\to+\infty} P\big(|X_n - \theta| > \varepsilon\big) = 0$$
3. On note, pour tout $n \in \mathbb{N}^*$, $Y_n = n(\theta - X_n)$.
   (a) Déterminer $Y_n(\Omega)$.
   (b) Montrer que, si $t \in [0, \theta n]$:
$$F_{Y_n}(t) = 1 - \left(1 - \frac{t}{\theta n}\right)^n$$
   (c) Montrer que, pour tout $t \in \mathbb{R}_+$, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geqslant n_0$, $t \in [0, \theta n]$.
   (d) En déduire que Y_n converge en loi vers une variable Y de loi $\mathscr{E}\left(\frac{1}{\theta}\right)$.

Estimations

Exercice 7. Soit une population constituée à 40% de cadres. On considère un échantillon en tirant au hasard 200 individus dans cette population. On considère que la population à étudier est assez grande et homogène, de sorte que les tirages des différents individus soient des expériences indépendantes.
On cherche à déterminer s'il est légitime de supposer, avec une proba $\geqslant 99\%$, que la fraction de cadres dans l'échantillon considéré sera comprise entre 30% et 50%.

1. Justifier que la variable aléatoire donnant la fraction de cadres dans l'échantillon est la moyenne empirique de 200 variables iid de loi $\mathscr{B}(p)$, avec $p = 0,4$. On note cette moyenne $\overline{X}$; donner les valeurs de E(X) et V(X).
2. On cherche donc à estimer la proba que cette moyenne empirique soit entre 0,3 et 0,5. À l'aide de Bienaymé-Tchebychev, montrer que $P(0,3 \leqslant \overline{X} \leqslant 0,5) \geqslant 0,8$. Peut-on conclure ?
3. On cherche maintenant à estimer cette probabilité en utilisant le TCL. Former la moyenne centrée réduite des X_i. Montrer que $P(0,3 \leqslant \overline{X} \leqslant 0,5) \simeq 2\Phi\left(\frac{10}{\sqrt{12}}\right) - 1$. Conclure, en utilisant $\frac{10}{\sqrt{12}} \simeq 2,89$.

Exercice 8 (Sondage)**.** Dans une population, une proportion $p \in]0,1[$ de personnes souhaitent voter pour le candidat A aux prochaines élections. On choisit un échantillon de n individus ; on pose les variables aléatoires $X_1, \ldots, X_n$ telles que $X_i = 1$ ssi l'individu i vote pour le candidat A. On considère alors que les X_i sont iid, de loi $\mathscr{B}(p)$.

1. Donner la variable aléatoire qui mesure la proportion d'individus votant pour le candidat A dans l'échantillon considéré. On note M_n cette variable.

2. Donner $E(M_n)$ et $V(M_n)$.

On cherche à mesurer la taille de l'échantillon à considérer pour avoir une bonne estimation de la proportion p.

3. Soit $\varepsilon > 0$. Montrer que $P\left(|M_n - p| > \varepsilon\right) \leqslant \dfrac{p(1-p)}{n\varepsilon^2}$; puis que $P\left(|M_n - p| > \varepsilon\right) \leqslant \dfrac{1}{4n\varepsilon^2}$.

4. Comment choisir n pour que M_n soit voisin de p à 0,01 près, avec une probabilité supérieure à $0,95$?

Exercice 9. Un vendeur de journaux propose au choix deux quotidiens A et B. Il dispose d'un stock de 40 exemplaires de A et 40 exemplaires de B. Chaque client demande soit A, soit B avec la probabilité 0,5. Les demandes des clients sont indépendantes les unes des autres.
60 clients se présentent dans une journée. On note x la probabilité de l'événement : «le vendeur ne satisfait pas toutes les demandes dans la journée».

1. Déterminer la loi de Y , nombre de clients demandant A dans la journée.

2. Exprimer x à l'aide de Y .

3. Déduire de l'inégalité de Bienaymé-Tchebychev un majorant de x.

4. En approchant la loi de Y par une loi normale, exprimer x à l'aide de la fonction Φ.

Approximations de lois

Exercice 10. Soit $n \geqslant 2$, et $Z_1, \ldots, Z_n$ iid suivant $\mathscr{G}(p)$ (avec $p \in]0,1[$). On pose $M_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} Z_i$.

1. Donner l'espérance m et l'écart-type σ_n de M_n.

2. Justifier l'existence de $\displaystyle\lim_{n\to+\infty} P\big(-\sigma_n \leqslant M_n - m \leqslant \sigma_n\big)$, et donner sa valeur en fonction de $\displaystyle\int_{-1}^{1} e^{-x^2/2}\,\mathrm{d}x$.

Exercice 11. Soit $X_1, \ldots, X_n$ iid suivant $\mathscr{E}(1)$. On pose $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$, et $X_n^* = \sqrt{n}\left(\overline{X_n} - 1\right)$.

1. Montrer que, pour tout $\varepsilon > 0$, $\displaystyle\lim_{n\to+\infty} P\left(|\overline{X_n} - 1| \geqslant \varepsilon\right) = 0$.

2. On note F_n la fonction de répartition de X_n^*.

   (a) Donner, pour $x \in \mathbb{R}$, $\displaystyle\lim_{n\to+\infty} F_n(x)$.

   (b) À l'aide de la table donnée en annexe, donner une valeur approchée de $\Phi(2)$. Calculer, pour n assez grand, une valeur approchée de $P\left(1 - \dfrac{2}{\sqrt{n}} \leqslant \overline{X_n} \leqslant 1 + \dfrac{2}{\sqrt{n}}\right)$.

Exercice 12. On modélise le nombre de clients par jour dans un magasin par une variable suivant une loi de Poisson de paramètre 10. On suppose que les variables associées aux différents jours sont mutuellement indépendantes. On cherche à estimer la probabilité qu'en 20 jours, le magasin ait plus de 180 clients.

1. Soient $X_1, \dots, X_{20}$ les va donnant le nombre de clients aux différents jours. Rappeler les valeurs de $E(X_i)$ et $V(X_i)$.

2. Soit $\overline{X} = \dfrac{1}{20} \displaystyle\sum_{i=1}^{20} X_i$ la moyenne empirique des X_i. Donner l'espérance et la variance de $\overline{X}$. En déduire l'expression de la moyenne centrée réduite, notée $\overline{X}^*$.

3. Exprimer l'événement $\left(\displaystyle\sum_{i=1}^{20} X_i \geqslant 180 \right)$ comme un événement portant sur $\overline{X}^*$.

4. En déduire qu'une valeur approchée de la probabilité recherchée est $\Phi\left(\sqrt{2}\right)$. Donner cette valeur approchée.