Lois limites

Exercice 1

A une entrée d'autoroute arrivent, en temps normal (hors période de pointe), 600 véhicules par heure en moyenne, répartis au hasard dans le temps. Soit N le nombre de véhicules qui entrent sur l'autoroute en 30 secondes.

1. Quelle est la loi de probabilité de N ?
2. Quelle est la probabilité d'avoir 10 entrées en 30 secondes ?
3. Quelle est la probabilité qu'il entre sur l'autoroute entre 10 et 20 véhicules en une minute (bornes comprises) ?

Exercice 2

Un livre de 350 pages contient 450 erreurs d'impression réparties au hasard.

1. Quelle est la probabilité qu'il y ait au moins trois erreurs dans une page donnée ?
2. Déterminer le nombre moyen de pages contenant n erreurs pour tout $0 \leqslant n \leqslant 450$.

Exercice 3

L'opératrice d'un standard téléphonique arcadien reçoit en moyenne deux appels par minute. Les appels sont répartis au hasard dans le temps.

1. Quelle est la loi de probabilité régissant le nombre d'appels reçus en 5 minutes ? On pourra supposer qu'il ne peut pas y avoir plus d'un appel par seconde.
2. Quelle est la probabilité que ce nombre dépasse 15 ?

Exercice 4 (L'ultimatum de France Telecom)

Le téléphone mobile dévastant le marché actuel, France Telecom évoque lors d'une allocution comminatoire la suppression de toute cabine téléphonique qui serait utilisée par moins de n personnes par mois. Le nombre de personnes utilisant une certaine cabine bucolique un jour donné suit une loi de Poisson de paramètre λ.

Quelle est la probabilité de survie du vestige abandonnique ? On considérera un mois de 30 jours.

Exemple numérique : $\lambda = 10$, $n = 290, 295, 299, 301, 305, 310$.

Exercice 5

Combien de fois au moins faut-il lancer une pièce pour que la fréquence des piles ne diffère de $1/2$ que de 0.05 au plus avec une probabilité de 99 chances sur 100 ?

Exercice 6

Une entreprise emploie 100 personnes. Chacune utilise son téléphone en moyenne 10 minutes par heure. Quel nombre minimum de lignes l'entreprise doit-elle posséder pour que, à un instant donné, la probabilité de saturation des lignes soit inférieure à 5% ?

Exercice 7 (Quelques convergences)

1. Montrer que la loi hypergéométrique de paramètres (N, n, p) converge vers la loi binomiale de paramètres (n, p) lorsque $N \longrightarrow +\infty$. En pratique, on considère la qualité de l'approximation bonne dès que $N \geqslant 10n$.
2. Soit pour tout $n \geqslant 1$ X_n une v.a. suivant la loi Gamma de paramètres $(n, 1)$. Montrer que la suite $\left(\frac{X_n - n}{\sqrt{n}}\right)_{n \geqslant 1}$ converge en loi vers une v.a. normalement distribuée.
3. Soit pour tout $n \geqslant 1$ X_n une v.a. suivant la loi du Khi-Deux à n d.d.l. Montrer que la suite $(\sqrt{2X_n} - \sqrt{2n-1}\,)_{n \geqslant 1}$ converge en loi vers une v.a. normalement distribuée.

On rappelle la formule de Stirling : $n! \sim_{n \to +\infty} \sqrt{2\pi}\, n^{n+\frac{1}{2}} e^{-n}$ et $\Gamma(x) \sim_{x \to +\infty} \sqrt{2\pi}\, x^{x-\frac{1}{2}} e^{-x}$.

Exercice 8

Rappeler la valeur de l'espérance, de la variance ainsi que la fonction caractéristique d'une loi de Poisson de paramètre λ. Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. indépendantes suivant une loi de Poisson de paramètre $\lambda = 1$.

1. Quelle est la loi suivie par la somme $X_1 + \cdots + X_n$?
2. Calculer la probabilité $\mathbb{P}(X_1 + \cdots + X_n \leqslant n)$.
3. Utiliser le théorème central limite pour montrer que $\lim_{n\to+\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}$.

Exercice 9 (Loi de Cauchy et loi des grands nombres)

La loi de Cauchy de paramètre $a > 0$ est la loi absolument continue de densité $f_a(x) = \frac{1}{\pi} \frac{a}{x^2 + a^2}$ et de fonction caractéristique $\varphi_a(t) = e^{-a|t|}$.

1. Soit X et Y deux v.a. indépendantes suivant les loi de Cauchy de paramètres respectifs a et b, et soit $\lambda \in \mathbb{R}$.
   (a) Montrer que λX suit une loi de Cauchy dont on précisera le paramètre.
   (b) Montrer que $f_a * f_b = f_{a+b}$. En déduire la loi de $X + Y$.
2. Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. i.i.d. suivant la loi de Cauchy de paramètre a. On pose $S_n = X_1 + \ldots + X_n$.
   (a) Quelle est la loi de $\frac{S_n}{n}$? Montrer que $\frac{S_n}{n} \xrightarrow[n\to\infty]{\mathcal{L}} X_1$.
   (b) Peut-on appliquer la loi des grands nombres à la suite $(X_n)_{n\geqslant 1}$? Pourquoi ?
   (c) Calculer la fonction caractéristique de la v.a. $\frac{S_n}{n} - \frac{S_{2n}}{2n}$. Montrer que $\frac{S_n}{n} - \frac{S_{2n}}{2n} \xrightarrow[n\to\infty]{\mathcal{L}} X_1$.
   (d) En déduire que la suite $\left(\frac{S_n}{n}\right)_{n\geqslant 1}$ ne converge pas en probabilité. (On pourra raisonner par l'absurde en supposant qu'il existe une v.a. Y telle que $\frac{S_n}{n} \xrightarrow[n\to\infty]{\mathcal{P}} Y$.)

Exercice 10 (Distributions empiriques (suite de l'exercice 9, fiche TD n°6))

Soit X une v.a. dont la fonction de répartition F est continue et strictement croissante sur $\mathbb{R}$ et M la médiane de X définie par $F(M) = \frac{1}{2}$. Soit $(X_1, \ldots, X_n)$ un échantillon de la v.a. X et $(X_{(1)}, \ldots, X_{(n)})$ l'échantillon ordonné associé. On définit la distribution empirique et la fonction de répartition empirique de cet échantillon respectivement par :

$$D_n(x) = \text{card } \{i \in \{1, \ldots, n\} : X_i \leqslant x\} \quad \text{et} \quad F_n(x) = \frac{1}{n} D_n(x).$$

I — On suppose n impair : $n = 2k + 1$, et l'on introduit la médiane empirique (aléatoire) : $M_n = X_{(k+1)}$.

1. $D_n(x)$ est une v.a. discrète à valeurs dans $\{1, \ldots, n\}$. Exprimer l'événement $\{D_n(x) = i\}$ à l'aide de $X_{(i)}$, $X_{(i+1)}$ et x. En déduire que la v.a. $D_n(x)$ suit une loi binomiale dont on précisera les paramètres.
2. (a) Montrer que pour tout $x \in \mathbb{R}$, $F_n(x) \xrightarrow[n\to\infty]{\mathcal{P}} F(x)$.
   On a le résultat plus précis suivant (théorème de Glivenko-Cantelli) :
   Si la fonction de répartition F est continue sur $\mathbb{R}$, alors $\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \xrightarrow[n\to\infty]{\text{p.s.}} 0$.
   (b) En déduire que $M_n \xrightarrow[n\to\infty]{\text{p.s.}} M$. (On pourra remarquer que $F_n(M_n) \overset{\text{p.s.}}{=} \frac{n+1}{2n} \xrightarrow[n\to\infty]{} F(M) = \frac{1}{2}$.)
3. On suppose maintenant que la v.a. X admet une densité continue $f > 0$. Montrer que la suite de v.a. $(\sqrt{n}(M_n - M))_{n\geqslant 1}$ converge en loi vers une v.a. suivant la loi normale de paramètres $(0, \frac{1}{4f(M)^2})$. (On rappelle la formule des accroissements finis valable pour toute fonction g de classe C^1 : $g(x + h) = g(x) + hg'(x + \theta h)$, $\theta \in]0, 1[$.)

II — On suppose n pair : $n = 2k$, et l'on choisit pour médiane empirique le nombre (aléatoire) $M_n = \frac{1}{2}[X_{(k)} + X_{(k+1)}]$.

1. Résultat préliminaire : soit $(Y_n)_{n\geqslant 1}$ et $(Z_n)_{n\geqslant 1}$ deux suites de v.a. convergeant en loi vers une même v.a. U, telles que $\forall n \geqslant 1$, $Y_n \leqslant Z_n$.
   (a) Comparer les fonctions de répartition de deux v.a. Y et Z vérifiant $Y \leqslant Z$.
   (b) En déduire que pour toute suite de v.a. $(U_n)_{n\geqslant 1}$ vérifiant $\forall n \geqslant 1$, $Y_n \leqslant U_n \leqslant Z_n$, on a $U_n \xrightarrow[n\to\infty]{\mathcal{L}} U$.
2. Montrer que les résultats des questions I–2 et I–3 subsistent encore dans ce cas.

Exercice 11 (Propriétés asymptotiques des échantillons d'une v.a.)

Soit X une v.a. telle que $\mathbb{E}(X^4) < +\infty$ et $(X_1, \ldots, X_n)$ un échantillon de X. On pose $m = \mathbb{E}(X)$, $\mu_k = \mathbb{E}(X-m)^k$, $k \in \{2,3,4\}$, et l'on introduit respectivement la moyenne empirique et la variance empirique non corrigée de cet échantillon :

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad V_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}_n^2.$$

1. Montrer que $\overline{X}_n \xrightarrow[n\to\infty]{\mathcal{L}} m$, puis que la suite $\left(\sqrt{n}\,\frac{\overline{X}_n - m}{\sqrt{\mu_2}}\right)_{n\geqslant 1}$ converge en loi vers une v.a. normalement distribuée.
2. Montrer que $V_n \xrightarrow[n\to\infty]{\mathcal{L}} \mu_2$.
3. On admet que $\operatorname{var}(V_n) = \dfrac{\mu_4 - \mu_2^2}{n} - \dfrac{2(\mu_4 - 2\mu_2^2)}{n^2} + \dfrac{\mu_4 - 3\mu_2^2}{n^3}$ (calculs méphistophéliques pour maniaco-dépressif, tendance schizophrénie). Montrer alors que la suite $\left(\frac{V_n - \mathbb{E}(V_n)}{\sigma(V_n)}\right)_{n\geqslant 1}$ converge en loi vers une v.a. normalement distribuée. (On pourra appliquer le théorème central limite à la suite $((X_n - m)^2)_{n\geqslant 1}$ et utiliser le résultat suivant : *soit $Y, Y_1, Y_2, \ldots$ une suite de v.a. et $\alpha, \alpha_1, \alpha_2, \ldots$ une suite de nombres réels telles que $Y_n \xrightarrow[n\to\infty]{\mathcal{L}} Y$ et $\alpha_n \xrightarrow[n\to\infty]{} \alpha$. Alors $\alpha_n Y_n \xrightarrow[n\to\infty]{\mathcal{L}} \alpha Y$.*)
4. Calculer $\operatorname{cov}(\overline{X}_n, V_n)$ puis $\lim_{n\to+\infty} \operatorname{cov}(\overline{X}_n, V_n)$. On pourra supposer que $m = 0$, ce qui n'est pas restrictif puisque V_n reste inchangé lorsqu'on translate les X_i d'une même constante.