Convergence

1. (a) Soit $\left(X_n\right)_{n\geqslant 1}$ une suite de variables aléatoires définies sur un même espace probabilisé, mutuellement indépendantes, de même loi uniforme sur $[0,1]$.
   Pour $n \in \mathbb{N}^*$, on définit les variables $M_n = \max(X_1, \ldots, X_n)$ et $Y_n = n\,(1 - M_n)$.

   (b) Déterminer la fonction de répartition de M_n puis celle de Y_n

2. Une entreprise compte 300 employés. Chacun d'eux téléphone en moyenne 6 minutes par heure. Quel nombre minimal de lignes installer pour que la probabilité qu'elles soient toutes occupées à un moment donné soit inférieure à $2,5\%$?

3. Un dé normal est lancé 9000 fois ; déterminer la probabilité que le 6 soit obtenu entre 1400 et 1600 fois. Comparer les résultats obtenus par l'inégalité de Bienaymé-Tchebychev et par approximation à l'aide d'une loi normale.

4. Un étudiant fait une faute d'orthographe en moyenne tous les 100 mots. Quelle est la probabilité qu'il ne fasse pas plus de 4 fautes dans un devoir de 200 mots ? *(on pourra utiliser une approximation par une loi de Poisson).*

5. Dans un élevage de poules pondeuses, aux heures de ponte maximale, le nombre X d'œufs pondus en 15 minutes dans un enclos suit une loi de Poisson de paramètre 10.

   (a) Quelles sont les valeurs de X ayant la probabilité maximale ?

   (b) Quelle est la probabilité qu'il y ait entre 7 et 12 œufs (au sens large) pondus en 15 minutes ?

   (c) Certains œufs sont déclassés (cassés ou poids insuffisant). La probabilité qu'un œuf soit déclassé est de $0,03$. Sur un lot de N œufs pondus, quelle est la loi Y_N suivie par le nombre d'œufs déclassés ?

   (d) On suppose à présent $N = 4000$.

      i. Par quelle loi peut-on approcher la loi de Y_N ? Déterminer f la densité continue et F la fonction de répartition associées à cette approximation.

      ii. Pour tout réel $h > 0$, on définit une fonction φ de la variable x par $\varphi(x) = F(x+h) - F(x-h)$. Étudier les variations de φ ; où faut-il placer le segment $[a,b]$ pour que, l'amplitude $(b-a)$ étant fixée, $P\big(a \leqslant Y_N \leqslant b\big)$ soit maximale ?

      iii. Trouver un intervalle $[a,b]$ centré en 120, de longueur minimale, tel que $P\big(a \leqslant Y_N \leqslant b\big) \geqslant 0,95$.

6. Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles mutuellement indépendantes, de même loi, définies sur un même espace probabilisé. On note F la fonction de répartition de cette loi. Pour tout réel x et pour tout entier naturel non nul, on note $Z_{n,x}$ le nombre de variables aléatoires X_i, $(i \in [\![1,n]\!])$ qui ont une valeur inférieure ou égale à x, et on pose $T_{n,x} = \frac{Z_{n,x}}{n}$.

   (a) Donner la loi de $Z_{n,x}$, son espérance et sa variance. Déterminer un majorant de $V(Z_{n,x})$ indépendant de x.

   (b) Montrer que $\forall \varepsilon > 0,\ \lim\limits_{n\to+\infty} P\Big(\big|T_{n,x} - F(x)\big| \geqslant \varepsilon\Big) = 0$

7. On considère une suite de variables aléatoires réelles $(X_n)_{n\geqslant 1}$, mutuellement indépendantes, de même loi de Poisson de paramètre 1, et on pose $S_n = X_1 + \cdots + X_n$.

   (a) Quelle est la loi de S_n ?

   (b) Déterminer $P\Big(S_n \leqslant n\Big)$.

   (c) En utilisant le théorème de la limite centrée, montrer que $\lim\limits_{n\to+\infty} \mathrm{e}^{-n} \sum\limits_{k=0}^{n} \dfrac{n^k}{k!} = \dfrac{1}{2}$.

1. (a) $M_n(\Omega) \subset [0,1]$ donc F_{M_n} est constante égale à 0 sur $]-\infty, 0[$, et constante égale à 1 sur $]1, +\infty[$.
$\big[M_n \leqslant t\big] = \big[X_1 \leqslant t\big] \cap \cdots \cap \big[X_n \leqslant t\big]$, donc par indépendance de $X_1, \ldots, X_n$,
$P\big(M_n \leqslant t\big) = P\big(X_1 \leqslant t\big) \times \cdots \times P\big(X_n \leqslant t\big)$ donc si $t \in [0,1]$, $F_{M_n}(t) = t^n$.

$$F_{M_n}(t) = \begin{cases} 0 & \text{si } t < 0 \\ t^n & \text{si } 0 \leqslant t \leqslant 1 \\ 1 & \text{si } t > 1 \end{cases}$$

$Y_n(\Omega) \subset [0,n]$ donc F_{Y_n} est constante égale à 0 sur $]-\infty, 0[$, et constante égale à 1 sur $]n, +\infty[$.
$\big[Y_n \leqslant t\big] = \big[M_n \geqslant 1 - \frac{t}{n}\big]$, donc si $t \in [0,n]$, $P\big(Y_n \leqslant t\big) = 1 - \big(1 - \frac{t}{n}\big)^n$

$$F_{Y_n}(t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 - \left(1 - \dfrac{t}{n}\right)^n & \text{si } 0 \leqslant t \leqslant n \\ 1 & \text{si } t > n \end{cases}$$

(b) Soit $t \geqslant 0$: pour $n \geqslant t$ on a $P\big(Y_n \leqslant t\big) = 1 - \left(1 - \dfrac{t}{n}\right)^n$; or $\ln\left(1 - \dfrac{t}{n}\right)^n = n \ln\left(1 - \dfrac{t}{n}\right) \underset{n\infty}{\sim} -t$. Donc $\lim\limits_{n\to+\infty} P\big(Y_n \leqslant t\big) = 1 - \mathrm{e}^{-t}$.
(Y_n) converge en loi vers une variable aléatoire qui suit une loi exponentielle de paramètre 1.

2. La probabilité qu'une ligne soit occupée à un instant donné est de $\frac{1}{10}$, donc le nombre X de lignes occupées suit une loi binômiale $\mathscr{B}\left(300, \frac{1}{10}\right)$. $E(X) = 30$ et $V(X) = 27$.

• On peut appliquer l'inégalité de Bienaymé-Tchebychev : $P\left(\big|X - E(X)\big| \geqslant \varepsilon\right) \leqslant \dfrac{V(X)}{\varepsilon^2}$.
En considérant que $P\left(X - E(X) \geqslant \varepsilon\right)$ et $P\left(X - E(X) \leqslant -\varepsilon\right)$ sont quasiment égaux (ce qui est une approximation), il suffit de prendre $\dfrac{V(X)}{\varepsilon^2} \leqslant 2 \times 0,025$ c'est à dire $\varepsilon \geqslant \sqrt{20 \times 27} = 6\sqrt{15}$.
$P\left(\big|X - 30\big| \geqslant 6\sqrt{15}\right) \geqslant P\left(X \geqslant 30 + 6\sqrt{15}\right)$, donc n est supérieur ou égal à 54.

• On peut aussi utiliser une approximation par une loi normale $\mathscr{N}(30, 27)$:

On a $n = 300 \geqslant 30$, $np = 30 \geqslant 5$, donc on peut approcher la loi de X par une loi normale $\mathscr{N}(30, 27)$, c'est à dire considérer que $X^* = \dfrac{X - 30}{\sqrt{27}} \hookrightarrow \mathscr{N}(0,1)$.

Le nombre de lignes N cherché est tel que $P(X \geqslant N) = P\left(X^* \geqslant \dfrac{N - 30}{\sqrt{27}}\right) \leqslant 0,025$, c'est à dire $\Phi\left(\dfrac{N-30}{\sqrt{27}}\right) \geqslant 1 - 0,025 = 0,975$. À l'aide de la table de Φ, on trouve $\dfrac{N - 30}{\sqrt{27}} \geqslant 1,96$ soit $N \geqslant 40,19$.
Il faut donc installer 41 lignes au minimum.

Remarque : L'approximation par la loi normale est bien plus précise que l'inégalité de Bienaymé-Tchebychev et la valeur trouvée pour n est nettement inférieure ; ce n'est pas surprenant car cette inégalité est valable quelle que soit la loi (à condition que la variance existe toutefois).

3. Le nombre X de 6 obtenus suit une loi binômiale $\mathscr{B}(9000, \frac{1}{6})$, d'espérance 1500 et de variance 1250.
— $P(1400 \leqslant X \leqslant 1600) = P\big(|X - E(X)| \leqslant 100\big)$; or $P\big(|X - E(X)| \geqslant 100\big) \leqslant \dfrac{V(X)}{100^2} = \dfrac{1}{8}$.
La probabilité que le nombre de 6 obtenus soit dans la fourchette donnée est supérieure à $0,875$.
— $P(1400 \leqslant X \leqslant 1600) \simeq P\left(\dfrac{1400 - 1500}{\sqrt{1250}} \leqslant X^* \leqslant \dfrac{1600 - 1500}{\sqrt{1250}}\right) = P\big(|X^*| \leqslant 2\sqrt{2}\big)$
$P\big(|X^*| \leqslant 2\sqrt{2}\big) = \Phi(2\sqrt{2}) - \Phi(-2\sqrt{2}) = 2\,\Phi(2\sqrt{2}) - 1 \simeq 2 \times 0,9977 - 1 = 0,9954$
En conclusion, l'inégalité de Bienaymé-Tchebychev fournit une approximation assez grossière.

4. Le nombre X de fautes suit une loi binômiale $\mathscr{B}(200, \frac{1}{100})$; $E(X) = 2$, $V(X) = 1,98$.
$n = 200 \geqslant 30$ et $p = \frac{1}{100} \leqslant 0,1$ donc l'approximation par une loi de Poisson $\mathscr{P}(2)$ est licite.

Donc $P(X \leqslant 4) \simeq \mathrm{e}^{-2} \displaystyle\sum_{k=0}^{4} \frac{2^k}{k!} = 7\,\mathrm{e}^{-2} \simeq 0,95$

5. (a) La probabilité de l'événement $[X = k]$ est $p_k = \mathrm{e}^{-10} \times \dfrac{10^k}{k!}$.

$\dfrac{p_{k+1}}{p_k} = \dfrac{10}{k+1}$ donc la suite (p_k) croît jusqu'à $k = 10$ (en fait $p_9 = p_{10}$) puis décroît ensuite.

Donc les valeurs les plus probables de X sont 9 et 10, avec $p_9 = p_{10} = \mathrm{e}^{-10} \times \dfrac{10^9}{9!} \simeq 0,125$

(b) $P\Big(7 \leqslant X \leqslant 12\Big) = \displaystyle\sum_{k=7}^{12} \mathrm{e}^{-10} \times \frac{10^k}{k!} = \frac{\mathrm{e}^{-10}}{7!}\left(1 + \frac{10}{8} + \frac{2 \times 10^2}{8 \times 9} + \frac{10^4}{8 \times 9 \times 11} + \frac{10^5}{8 \times 9 \times 11 \times 12}\right) \simeq 0,66$

(c) En faisant l'hypothèse que le déclassement d'un œuf est indépendant de celui des autres, Y_N suit une loi binômiale $\mathscr{B}(N, \frac{3}{100})$.

(d) i. $E(Y_{4000}) = 120$ et $V(Y_{4000}) = 120 \times 0,97 = 116,4$ donc la loi de Y_N peut être approchée par une loi normale de paramètres 120 et 116, 4.

Donc une densité de cette loi approchée est pour $x \in \mathbb{R}$, $f(x) = \dfrac{1}{\sqrt{232,8\,\pi}} \exp\left(-\frac{(x-120)^2}{232,8}\right)$, et la fonction de répartition $F(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t$.

ii. $\varphi'(x) = f(x+h) - f(x-h)$ s'annule pour $x = 120$ ($\exp\left(-\frac{(x+h-120)^2}{232,8}\right)$ et $\exp\left(-\frac{(x-h-120)^2}{232,8}\right)$ sont égaux) et φ est maximale pour $x = 120$, donc $b-a$ étant fixé, la probabilité cherchée (plus exactement une valeur approchée de cette probabilité, obtenue en utilisant la loi normale) sera maximale si le segment $[a, b]$ est centré en 120, c'est à dire $a + b = 240$ *(on peut poser $a = 120 - h$ et $b = 120 + h$, avec $h > 0$ égal à la moitié de l'amplitude de l'intervalle).*

iii. Notons Z la variable aléatoire qui approxime Y_N, ($Z \hookrightarrow \mathscr{N}(120\,;\,116,4)$; la probabilité d'avoir $a \leqslant Z \leqslant b$ est donnée par : $P\big([a \leqslant Z \leqslant b]\big) = P\big([\frac{a-120}{\sqrt{116,4}} \leqslant \frac{Z-120}{\sqrt{116,4}} \leqslant \frac{b-120}{\sqrt{116,4}}]\big)$.

La probabilité est maximale si $b-a$ est donné, revient à dire que la probabilité étant fixée (à 0, 95 ici) l'amplitude est minimale ; ainsi l'intervalle $[a, b]$ cherché est tel que $a+b = 240$, donc avec $a = 120-h$ et $b = 120 + h$, on obtient : $P\Big([\frac{-h}{\sqrt{116,4}} \leqslant \frac{Z-120}{\sqrt{116,4}} \leqslant \frac{h}{\sqrt{116,4}}]\Big) \geqslant 0,95$.

Or $\frac{Z-120}{\sqrt{116,4}}$ suit une loi normale centrée réduite, d'où

$$P\Big([\frac{-h}{\sqrt{116,4}} \leqslant \frac{Z-120}{\sqrt{116,4}} \leqslant \frac{h}{\sqrt{116,4}}]\Big) = \phi\left(\frac{-h}{\sqrt{116,4}}\right) - \phi\left(\frac{h}{\sqrt{116,4}}\right) = 2\,\phi\left(\frac{h}{\sqrt{116,4}}\right) - 1 \geqslant 0,95.$$

$\phi\left(\dfrac{h}{\sqrt{116,4}}\right) \geqslant 0,975$ ce qui, en consultant la table, donne $\dfrac{h}{\sqrt{116,4}} \geqslant 1,96$ donc $h \geqslant 21,4$.

L'amplitude minimale de l'intervalle est égale à 42, 8. Ainsi $[98,6\,;\,141,4]$ est l'intervalle cherché.

6. (a) $Z_{n,x} \hookrightarrow \mathscr{B}\big(n, F(x)\big)$, donc $E\big(Z_{n,x}\big) = n\,F(x)$ et $V(Z_{n,x}) = n\,F(x)\,\big(1 - F(x)\big)$.

La fonction $u \mapsto u\,(1-u)$ définie sur $[0, 1]$ admet un maximum égal à $\frac{1}{4}$, atteint pour $u = \frac{1}{2}$, donc $V(Z_{n,x}) \leqslant \dfrac{n}{4}$.

(b) L'inégalité de Bienaymé-Tchebychev donne $P\Big(\big|T_{n,x} - F(x)\big| \geqslant \varepsilon\Big) \leqslant \dfrac{V(T_{n,x})}{\varepsilon^2}$.

Or $V(T_{n,x}) = \dfrac{V(Z_{n,x})}{n^2} = \dfrac{1}{4\,n}$; ainsi $P\Big(\big|T_{n,x} - F(x)\big| \geqslant \varepsilon\Big) \leqslant \dfrac{1}{4\,n\,\varepsilon^2}$ tend vers 0 à ε fixé lorsque n tend vers l'infini.

7. (a) S_n suit une loi de Poisson de paramètre n (cours).

(b) $P\Big(S_n \leqslant n\Big) = \mathrm{e}^{-n} \displaystyle\sum_{k=0}^{n} \frac{n^k}{k!}$.

(c) $E(S_n) = V(S_n) = n$. D'après le théorème de la limite centrée, $\dfrac{S_n - n}{\sqrt{n}}$ converge en loi vers une loi normale centrée réduite. $\displaystyle\lim_{n\to+\infty} P\Big(S_n \leqslant n\Big) = \lim_{n\to+\infty} P\Big(\frac{S_n - n}{\sqrt{n}} \leqslant 0\Big) = \Phi(0) = \frac{1}{2}$. Donc $\displaystyle\lim_{n\to+\infty} \mathrm{e}^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}$.