Convergence en loi. Théorème de la limite centrale.

1. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée. Montrer qu'on a

$$\lim_{n\to+\infty} e^{-n} \sum_{k=0}^{\infty} f\left(\frac{k-n}{\sqrt{n}}\right) \frac{n^k}{k!} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-\frac{t^2}{2}}\, dt.$$

Solution de l'exercice 1. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de loi de Poisson de paramètre 1. On a $\mathbb{E}[X_1] = 1$ et $\mathrm{Var}(X_1) = 1$.

D'une part, le théorème central limite assure que la suite $\left(\frac{X_1+\ldots+X_n-n}{\sqrt{n}}\right)_{n\geq 1}$ converge en loi vers la loi normale centrée réduite. Ainsi, puisque f est continue et bornée,

$$\lim_{n\to\infty} \mathbb{E}\left[f\left(\frac{X_1+\ldots+X_n-n}{\sqrt{n}}\right)\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-\frac{t^2}{2}}\, dt.$$

D'autre part, pour tout $n \geq 1$, la variable aléatoire $X_1 + \ldots + X_n$ suit la loi de Poisson de paramètre n. Ainsi,

$$\mathbb{E}\left[f\left(\frac{X_1+\ldots+X_n-n}{\sqrt{n}}\right)\right] = e^{-n} \sum_{k=0}^{\infty} f\left(\frac{k-n}{\sqrt{n}}\right) \frac{n^k}{k!}.$$

En comparant ces deux égalités, on a la convergence voulue.

2. *(Examen 2010)* Durant l'année qui vient de s'écouler, 10000 enfants sont nés dans une certaine ville. Parmi ces enfants, 5242 sont des filles et 4758 sont des garçons. Vous paraît-il raisonnable de soutenir l'hypothèse qu'un enfant qui naît dans cette ville a exactement autant de chance d'être un garçon que d'être unee fille ?

Solution de l'exercice 2. voir correction de l'examen 2010, 2eme session

3. Un marchand d'accessoires de magie vend des dés de deux sortes : des dés équitables et des dés pipés où le chiffre 6 sort avec probabilité 1/5. Dans le dernier lot que son fabricant lui a envoyé, le fabricant a oublié d'étiqueter les dés. Le marchand ne sait donc pas quel dé est de quelle sorte.

Pour les départager, il les prend un à un, les lance mille fois chacun et note le nombre de fois où il a obtenu 6. Il décide d'étiqueter "équitable" tout dé qui donne moins de

183 fois 6 et “pipé” un dé qui donne plus de 184 fois 6. En effet, raisonne-t-il, $183 \simeq \frac{1}{2}(1000 * \frac{1}{6} + 1000 * \frac{1}{5})$.

a. Quelle proportion des dés équitables étiquette-t-il comme pipés ? Et quelle proportion des dés pipés comme équitables ?

b. Le marchand sait que dans le lot qu'il a reçu, il y a à peu près autant de dés pipés que de dés équitables. Le seuil de 183 est-il celui qui lui fait commettre au total le moins d'erreurs ?

On pourra utiliser les valeurs suivantes de la fonction de répartition de la loi $\mathcal{N}(0, 1)$.

x	1	1.05	1.1	1.15	1.2	1.25	1.3	1.35	1.4	1.45	1.5
$\Phi(x)$	0.841	0.853	0.864	0.875	0.885	0.894	0.903	0.911	0.919	0.926	0.933

Solution de l'exercice 3. a. Soit $(X_n)_{n\geq 1}$ une suite de v.a. indépendantes de Bernouilli de paramètre $p = 1/6$ et soit $S_n = X_1 + \cdots + X_n$ le nombre de 6 obtenus sur n lancers pour avec un dé équitable. On a $\mathbb{E}[X_1] = p$ et $\mathrm{Var}(X_1) = p(1-p)$. D'où, par le théorème de la limite centrale :

$$\frac{S_n - pn}{\sqrt{p(1-p)n}} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0, 1).$$

Soit Z une v.a. suivant la loi normale centrée réduite. Ici, $n = 1000$, donc on peut approcher la probabilité que le marchand déclare à tort un dé pipé à l'aide de la limite ci-dessus :

$$\begin{aligned} \mathbb{P}(S_{1000} \geq 184) &= \mathbb{P}\left(\frac{S_{1000} - 1000/6}{\sqrt{1000 \times 5/36}} \geq \frac{184 - 1000/6}{\sqrt{1000 \times 5/36}}\right) \\ &\approx \mathbb{P}\left(Z \geq \frac{184 - 1000/6}{\sqrt{1000 \times 5/36}}\right) \approx \mathbb{P}(Z \geq 1,47) \approx 0,071. \end{aligned}$$

Par la loi des grands nombres, c'est à peu près la proportion de dés équitables qui seront étiquetés pipés.

De même, en notant T_n le nombre de 6 obtenus avec un dé pipé, on trouve que le marchand prendra un dé pipé pour un dé équilibré avec probabilité :

$$\begin{aligned} \mathbb{P}(T_{1000} \leq 183) &= \mathbb{P}\left(\frac{T_{1000} - 1000/5}{\sqrt{1000 \times 4/25}} \leq \frac{183 - 1000/5}{\sqrt{1000 \times 4/25}}\right) \\ &\approx \mathbb{P}\left(Z \leq \frac{183 - 1000/5}{\sqrt{1000 \times 4/25}}\right) \approx \mathbb{P}(Z \leq -1,34) \approx 0,091. \end{aligned}$$

b. Avec autant de dés pipés équilibrés, le marchand fera en moyenne $8,1\%$ d'erreurs avec le seuil de $183-184$. Avec un seuil de $182-183$, on obtiendrait un taux d'erreur pour les dés équitables d'environ $\mathbb{P}(Z \geq 1.39) \approx 0,081$, et pour les dés pipés $\mathbb{P}(Z \leq -1.42) \approx 0,085$. Soit $8,3\%$ d'erreurs en moyenne.

Avec un seuil de $184 - 185$, on obtiendrait un taux d'erreur pour les dés équitables d'environ $\mathbb{P}(Z \geq 1.56) \approx 0,060$, et pour les dés pipés $\mathbb{P}(Z \leq -1.26) \approx 0,104$, soit $8,2\%$ d'erreurs en moyenne.

Le seuil choisi par le marchand minimise donc la proportion d'erreurs qu'il fera. En effet, il s'agit d'un minimum local comme on vient de le voir, et global par convexité (il suffit de voir que la dérivée est croissante) de la fonction

$$x \mapsto \mathbb{P}\left(Z \geq \frac{184 + x - 1000/6}{\sqrt{1000 \times 5/36}}\right) + \mathbb{P}\left(Z \leq \frac{183 + x - 1000/5}{\sqrt{1000 \times 4/25}}\right)$$

sur l'intervalle déterminé par $\frac{184+x-1000/6}{\geq}0$ et $183 + x - 1000/5 \leq 0$. Et en dehors de cet intervalle, l'erreur est d'au moins 1/2 pour un type de dés, donc supérieure à 1/4 en moyenne.

4. Soient $(X_n)_{n\geq 1}$ des variables aléatoires indépendantes identiquement distribuées et de carré intégrable. On suppose que leur loi a la propriété suivante : $X_1 + X_2$ a même loi que $\sqrt{2}X_1$.

a. Exprimer, pour tout $n \geq 0$, la loi de $X_1 + \ldots + X_{2^n}$ en fonction de celle de X_1.

b. Déterminer la loi commune des variables aléatoires $(X_n)_{n\geq 1}$.

Solution de l'exercice 4. a. On va montrer par récurrence que $X_1 + \ldots + X_{2^n}$ a même loi que $\sqrt{2}^n X_1$, pour tout $n \geq 1$. Pour $n = 1$, c'est l'hypothèse. Supposons le résultat démontré jusqu'au rang n et démontrons le au rang $n + 1$. On écrit

$$X_1 + \ldots + X_{2^{n+1}} = (X_1 + \ldots + X_{2^n}) + (X_{2^n+1} + \ldots + X_{2^{n+1}})$$

. La v.a. $X_{2^n+1} + \ldots + X_{2^{n+1}}$ est indépendante de $X_1 + \ldots + X_{2^n}$ (car les X_n sont indépendantes et on utilise des paquets d'indices disjoints) et de même loi. Par l'hypothèse de récurrence, ces deux v.a. ont même loi que $\sqrt{2}^n X_1$. Mais alors leur somme a même loi que $\sqrt{2}^n X_1 + \sqrt{2}^n X_2 = \sqrt{2}^n(X_1 + X_2)$, qui a bien même loi (d'après l'hypothèse pour $n = 1$) que $\sqrt{2}^{n+1} X_1$. Ce qui achève la preuve par récurrence.

b. Comme X_1 est de carré intégrable, X_1 est intégrable. L'hypothèse entraine alors que $\mathbb{E}[X_1] + \mathbb{E}[X_1] = \sqrt{2}\mathbb{E}[X_1]$, et donc $\mathbb{E}[X_1] = 0$. Soit $\sigma^2 = \mathbb{E}[X_1^2]$. On applique le théorème de la limite centrale :

$$\frac{X_1 + \ldots + X_n}{\sqrt{n}} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0, \sigma^2).$$

On a la même convergence pour les sous-suites, et en particulier :

$$\frac{X_1 + \ldots + X_{2^n}}{\sqrt{2^n}} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0, \sigma^2).$$

Or, d'après la question précédente, le terme de gauche a même loi que X_1, qui suit donc la loi $\mathcal{N}(0, \sigma^2)$.