Convergences et approximations en probabilités

Toutes les variables aléatoires de ce chapitre sont définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.

1 Suites de variables aléatoires

On a déjà rencontré, notamment dans l'étude des chaînes de Markov, des suites de variables aléatoires. Le but de ce préliminaire est de donner quelques exemples où cette notion est utile, et de définir proprement l'indépendance mutuelle dans le cas d'une infinité de variables.

Exemple 1.

- Une chaîne de Markov donne un exemple d'une suite de variables aléatoires : pour tout $n \in \mathbb{N}$, X_n est la variable aléatoire donnant la position d'un « individu » sur le graphe au bout de n étapes.
- Soit une variable aléatoire X mesurant le résultat d'une expérience aléatoire quelconque.
Pour modéliser la répétition de cette expérience un grand nombre de fois (en vue de produire des statistiques, d'estimer des paramètres...), on introduit une suite de variables aléatoires indépendantes, de même loi que X.
Par exemple, on modélisera une succession infinie de lancers d'un dé équilibré par une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ indépendantes suivant toutes la loi $\mathscr{U}(\llbracket 1,6 \rrbracket)$.
- On considère une *marche au hasard à deux dimensions* : un objet est initialement en $(0,0)$ (origine du repère) et peut, à chaque étape de temps, sauter d'une unité vers le haut, vers le bas, vers la gauche, ou vers la droite. Si on note X_n la variable aléatoire égale à son abscisse au temps n, et Y_n son ordonnée au temps n, alors $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ sont deux suites de variables aléatoires à valeurs dans $\mathbb{Z}$.

On peut alors étendre la définition de l'indépendance à une infinité de variables :

Définition 1. *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires. On dit que les variables X_n sont indépendantes si et seulement si, pour toute partie $I \subset \mathbb{N}$, les variables $(X_i)_{i\in I}$ sont mutuellement indépendantes.*

En pratique, si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables indépendantes, tous les résultats d'indépendance peuvent s'appliquer : propriétés de l'espérance du produit, de la variance de la somme, lemme des coalitions, stabilités des lois, etc.

Exemple 2.

- Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables indépendantes suivant $\mathscr{B}(1/2)$. On pose, pour $n \in \mathbb{N}^*$, $Y_n = \sum_{k=1}^{n} X_k$.
Alors, pour tout $n \in \mathbb{N}^*$, $Y_n \hookrightarrow \mathscr{B}(n, 1/2)$.
- Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables indépendantes. Alors, pour tout $n \in \mathbb{N}^*$, $V\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} V(X_k)$.
- Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables indépendantes, et $S_n = \sum_{k=1}^{n} X_k$. Alors S_n et X_{n+1} sont indépendantes.

Dans ce chapitre et celui qui suit, nous serons donc souvent en présence d'une suite de variables aléatoires suivant toutes la même loi.

Définition 2. *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires. On dit que les X_n sont indépendantes et identiquement distribuées (en abrégé : i.i.d ou iid) ssi elles sont indépendantes, et suivent toutes la même loi.*

2 Majorations : Markov et Bienaymé-Tchebychev

Les deux théorèmes de cette section sont valables pour des variables discrètes, et pour des variables à densité.

2.1 Inégalité de Markov

L'inégalité de Markov permet de quantifier, pour des variables aléatoires admettant une espérance, la probabilité que la variable prenne de « grandes » valeurs.

Théorème 1 (Inégalité de Markov). *Soit X une variable aléatoire* ***positive****, admettant une espérance. On a :*

$$\forall\, t > 0,\ \mathbb{P}(X \geqslant t) \leqslant \frac{\mathbb{E}(X)}{t}$$

Démonstration. En annexe. □

L'inégalité de Markov permet donc de donner une majoration de la « queue » de la distribution de probabilité ; notamment dans le cas d'une variable à densité, cela montre que l'aire sous la courbe de la densité, entre t et $+\infty$, décroît au plus comme $\frac{1}{t}$.

On peut en fait raffiner l'estimation si la variable X admet des moments d'ordre supérieur. Si X positive admet un moment d'ordre $r \in \mathbb{N}^*$, on applique Markov à la variable X^r (qui admet donc une espérance !) et on trouve :

$$\forall\, t > 0,\ \mathbb{P}(X \geqslant t) = \mathbb{P}(X^r \geqslant t^r) \leqslant \frac{\mathbb{E}(X^r)}{t^r}$$

Ceci fournit cette fois une décroissante de la queue de distribution en $\frac{1}{t^r}$.
Si une variable X n'est pas positive, la condition de convergence absolue montrer que X admet une espérance ssi $|X|$ en admet une. On peut alors appliquer Markov à $|X|$. On obtient :

$$\forall\, t > 0,\ \mathbb{P}(|X| \geqslant t) \leqslant \frac{\mathbb{E}(|X|)}{t}$$

et on estime cette fois la probabilité $\mathbb{P}(X \leqslant -t) + \mathbb{P}(X \geqslant t)$ (**les** queues de distribution donc, en $+\infty$ et $-\infty$).
En combinant les deux extensions précédentes, on a, si X admet un moment d'ordre r :

$$\forall\, t > 0,\ \mathbb{P}(X \leqslant -t) + \mathbb{P}(X \geqslant t) \leqslant \frac{\mathbb{E}(|X|^r)}{t^r}$$

2.2 Inégalité de Bienaymé-Tchebychev

L'inégalité de Bienaymé-Tchebychev permet, pour une variable admettant une variance, de quantifier la probabilité que la variable s'écarte de sa valeur moyenne.

Théorème 2 (Inégalité de Bienaymé-Tchebychev). *Soit* X *une variable admettant une variance. Alors elle admet aussi une espérance, et on a :*

$$\mathbb{P}\big(|X - \mathbb{E}(X)| \geqslant t\big) \leqslant \frac{V(X)}{t^2}$$

Démonstration. En annexe.

□

On examine donc cette fois les probabilités d'écart à la moyenne ; il est normal que leur ordre de grandeur fasse intervenir la variance.

3 Loi faible des grands nombres

Commençons par l'énoncé du théorème :

> **Théorème 3** (Loi faible des grands nombres)**.** *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires admettant des variances ; on suppose que toutes les* X_n *ont la même espérance notée* m*, et la même variance notée* v*.*
> *Pour* $n \in \mathbb{N}^*$*, on note* $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$*. On a alors :*
>
> $$\forall\, \varepsilon > 0,\ \lim_{n\to+\infty} \mathbb{P}\left(\left|\overline{X_n} - m\right| \geqslant \varepsilon\right) = 0$$

Démonstration. On commence par montrer deux propriétés qu'il faudra savoir redémontrer :

- $\mathbb{E}\left(\overline{X_n}\right) = m$.
  En effet la linéarité de l'espérance donne :
  $$\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}(X_k) = \frac{1}{n}\sum_{k=1}^{n} m = m$$
- $V\left(\overline{X_n}\right) = \frac{v}{n}$.
  En effet l'indépendance mutuelle des (X_i) et le caractère quadratique de la variance donnent :
  $$V\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} V(X_k) = \frac{1}{n^2}\sum_{k=1}^{n} v = \frac{v}{n}$$

Bienaymé-Tchebychev donne alors :

$$\forall\, n \in \mathbb{N}^*,\ \mathbb{P}\left(\left|\overline{X_n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{v/n}{\varepsilon^2} = \frac{v}{n\varepsilon^2}$$

et on en déduit bien $\lim_{n\to+\infty} \mathbb{P}\left(\left|\overline{X_n} - m\right| \geqslant \varepsilon\right) = 0$. □

Interprétons maintenant ce résultat, dans le contexte où on l'appliquera la plupart du temps : celui de variables iid.

> **Définition 3.** *Soit X une varible aléatoire, et* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires iid, de même loi que X. On appelle, pour tout* $n \in \mathbb{N}^*$, $(X_1,\dots,X_n)$ *un* n*-échantillon de la loi de X.*
> *On appelle* moyenne empirique des X_i *la variable aléatoire :*
>
> $$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

Remarque 1. Si X admet une variance (et donc une espérance), les propriétés démontrées précédemment montrent que $\mathbb{E}\left(\overline{X_n}\right) = \mathbb{E}(X)$, et $V\left(\overline{X_n}\right) = \frac{V(X)}{n}$.

La loi faible des grands nombres s'applique donc aux (X_n) et donne une première traduction formelle du lien entre probabilités et statistiques, qui peut s'énoncer comme :

> Soit une variable X d'espérance $\mathbb{E}(X)$; on effectue n tirages indépendants de X.
> Alors la moyenne de ces tirages se rapproche de $\mathbb{E}(X)$ lorsque $n \to +\infty$.

C'est en fait cela qui justifie les programmes informatiques habituels qui approximent la valeur d'une variable aléatoire en effectuant la moyenne d'un grand nombre de tirages. En effet la loi faible des grands nombres montre que la probabilité qu'une valeur de $\overline{X_n}$ (qui est donc la moyenne de n résultats expérimentaux) s'écarte de $\mathbb{E}(X)$ à une précision quelconque tend vers 0 lorsque $n \to +\infty$.

4 Convergence en loi

On peut considérer un autre type de convergence, mettant en jeu la loi de probabilité des X_n (donc la fonction de répartition).

Définition 4. *Soit X une variable aléatoire,* F_X *sa fonction de répartition ; soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires. On dit que* (X_n) converge en loi *vers X ssi :*

$$\textbf{\textit{en tout point }} x \textbf{\textit{ où }} F_X \textbf{\textit{ est continue}}, \quad \lim_{n\to+\infty} F_{X_n}(x) = F_X(x)$$

Si (X_n) *converge en loi vers X, on note* $X_n \xrightarrow{\mathscr{L}} X$.

Exemple 3. Soient X_n des variables aléatoires tq : $\forall\, n \in \mathbb{N}^*, X_n \hookrightarrow \mathscr{E}(n)$. Montrer que (X_n) converge en loi vers une variable aléatoire constante égale à 0.

Remarque 2.

- Il s'agit donc d'une convergence de $F_{X_n}(x)$ *à x fixé*, et pour $n \to +\infty$.
- Remarquer qu'on ne considère que les points où F_X est continue. Dans le cas où la loi limite est une loi discrète, on n'a donc pas besoin des limites aux points de $X(\Omega)$: par exemple si on veut montrer que (X_n) converge en loi vers une va constante égale à 0, il suffit de montrer :

  $$\forall\, x < 0, \ \lim_{n\to+\infty} F_{X_n}(x) = 0 \quad \text{et} \quad \forall\, x > 0, \ \lim_{n\to+\infty} F_{X_n}(x) = 1$$

  et on n'a pas à examiner $\lim_{n\to+\infty} F_{X_n}(0)$.

  Si par contre on tend vers une variable à densité (donc de fonction de répartition continue !), il faut examiner $\lim_{n\to+\infty} F_{X_n}(x)$ pour tout $x \in \mathbb{R}$.

Dans le cas où les variables formant la suite et la variable limite sont toutes à valeurs dans $\mathbb{Z}$, on peut se ramener à un critère faisant intervenir les $\mathbb{P}(X = k)$ pour $k \in \mathbb{Z}$.

Théorème 4. *Soient* $(X_n)_{n\in\mathbb{N}}$ *et* X *des variables aléatoires discrètes, à valeurs dans* $\mathbb{Z}$. *Alors* $X_n \xrightarrow{\mathscr{L}} X$ *si et seulement si :*

$$\forall\, k \in \mathbb{Z}, \ \lim_{n\to+\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$$

Démonstration. En annexe. □

Remarque 3. Attention, en cas de convergence en loi on n'a pas forcément convergence des espérances : $X_n \xrightarrow{\mathscr{L}} X$ n'implique pas $\lim_{n\to+\infty} \mathbb{E}(X_n) = \mathbb{E}(X)$.
Contre exemple : soient $(X_n)_{n\in\mathbb{N}^*}$ telles que

$$\mathbb{P}(X_n = 0) = 1 - \frac{1}{n} \ , \ \mathbb{P}(X_n = n) = \frac{1}{n}$$

On vérifie facilement que (X_n) converge en loi vers une variable constante égale à 0 ; et que par ailleurs : $\forall\, n \in \mathbb{N}^*$, $\mathbb{E}(X_n) = 1$.

Un exemple fameux de convergence en loi est donné par **l'approximation binomiale d'une loi de Poisson** :

Proposition 5. *Soient* $\lambda > 0$, *et* X_n *des variables aléatoires suivant* $\mathscr{B}(n, \lambda/n)$.
Alors $X_n \xrightarrow{\mathscr{L}} X$, *avec* $X \hookrightarrow \mathscr{P}(\lambda)$.

Démonstration. En annexe. $\square$

En pratique, on pourra donc, pour n grand, approximer la loi $\mathscr{B}(n,p)$ par la loi $\mathscr{P}(np)$. Usuellement, on considère cette approximation correcte sous les hypothèses : $\begin{cases} p \leqslant 0,1 \\ n \geqslant 30 \\ np \leqslant 15 \end{cases}$.

Remarque 4. Ces critères ne sont pas à connaître ; si vous avez à utiliser une telle approximation dans un exercice, l'énoncé vous indiquera qu'il est légitime de le faire.

5 Théorème central limite et conséquences

Jusqu'à la fin du cours, on note Φ la fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0,1)$.

5.1 Le théorème

Ce théorème montre, en un sens, que la loi normale est une distribution « universelle » dans le cas d'un grand nombre de mesures.

La loi faible des grands nombres montre que la moyenne empirique de n variables aléatoires iid, d'espérance commune m, se comportera, pour n grand, comme une variable prenant des valeurs proches de m. Le théorème central limite donne cette fois une convergence *en loi*, et précise la répartition autour de cette valeur m : il s'agit d'une loi normale.

Pour qu'une convergence en loi ait lieu, il est raisonnable de se ramener à des espérances et des variances constantes ; et donc de considérer des variables centrées réduites.
Soient (X_n) des variables aléatoires iid, d'espérance m et de variance σ^2. Alors $\overline{X_n} = \dfrac{1}{n}\sum_{k=1}^{n} X_i$ a pour espérance m et pour variance $\dfrac{\sigma^2}{n}$; de sorte que $\sqrt{n}\left(\dfrac{\overline{X_n}-m}{\sigma}\right)$ est centrée réduite.
Ceci motive la définition suivante :

Définition 5 (Moyenne centrée réduite)**.** *Soit* $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires iid, admettant une espérance m et une variance* σ^2 *non nulle. On pose :*

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad et \quad \overline{X_n}^* = \sqrt{n}\left(\frac{\overline{X_n}-m}{\sigma}\right)$$

On a vu que $\left(\overline{X_n}\right)_n$ *est la suite des moyennes empiriques.*
On appelle $\left(\overline{X_n}^*\right)_n$ *la suite des* ***moyennes centrées réduites*** *associées à la suite* (X_n).

On a alors :

Théorème 6 (Théorème central limite)**.** *Soit* $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires iid, admettant une espérance m et une variance* σ^2 *non nulle. On note* $\overline{X_n}^*$ *les moyennes centrées réduites associées à la suite* (X_n).
Alors

$$\overline{X_n}^* \xrightarrow{\mathcal{L}} X \text{ , où } X \hookrightarrow \mathcal{N}(0,1)$$

Démonstration. Admis (c'est pas facile !!). □

Remarque 5. Évidemment, on ne peut pas repasser les variables à l'échelle et conclure que $\overline{X_n}$ converge en loi vers une variable suivant $\mathcal{N}\left(m,\sigma^2/n\right)$: la variable n ne peut pas apparaître dans une limite $n\to+\infty$. Mais on peut quand même interpréter le théorème central limite en disant que $\overline{X_n}$ suit approximativement la loi $\mathcal{N}\left(m,\sigma^2/n\right)$ pour n grand. Ceci permet de raffiner les considérations faites avant l'énoncé de la loi faible des grands nombres.

La convergence en loi étant définie par la convergence point par point de la fonction de répartition (ici en tout point car la fonction de répartition de $\mathcal{N}(0,1)$ est continue sur $\mathbb{R}$), le théorème central limite implique le résultat suivant :

Proposition 7. *Soient $a < b$ deux réels (avec éventuellement $a = -\infty$ ou $b = +\infty$) Avec les notations du théorème précédent, on a*

$$\lim_{n\to+\infty} \mathbb{P}\big(a \leqslant \overline{X_n}^* \leqslant b\big) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\,\mathrm{d}t = \Phi(b) - \Phi(a)$$

Ceci permet d'obtenir l'expression de probabilités de la forme $\lim\limits_{n\to+\infty} \mathbb{P}(a_n \leqslant \overline{X_n} \leqslant b_n)$, où a_n et b_n auront des expressions permettant aux calculs de bien s'arranger.

Exemple 4. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires iid de même loi $\mathscr{B}(1/3)$.
Pour tout $n \in \mathbb{N}^*$, on pose $\overline{X_n}$ la moyenne empirique de $X_1, \dots, X_n$. On a alors classiquement

$$\mathbb{E}(X_n) = \mathbb{E}(X_1) = \frac{1}{3} \quad \text{et} \quad V(X_n) = \frac{V(X_1)}{n} = \frac{2}{9n}$$

Montrer que :

$$\lim_{n\to+\infty} \mathbb{P}\left(\frac{1}{3} - \sqrt{\frac{2}{n}} \leqslant \overline{X_n} \leqslant \frac{1}{3} + \sqrt{\frac{2}{n}}\right) = \Phi(3) - \Phi(-3)$$

Remarque 6. Par propriété de Φ, on trouve aussi $\lim\limits_{n\to+\infty} \mathbb{P}\left(\frac{1}{3} - \sqrt{\frac{2}{n}} \leqslant \overline{X_n} \leqslant \frac{1}{3} + \sqrt{\frac{2}{n}}\right) = 2\Phi(3) - 1 \simeq 0,9974$ (voir tables de la loi normale en annexe). À n grand, il y a 99.74% de chances que la moyenne empirique donne la probabilité de succès avec une précision $\pm\sqrt{2/n}$.
Nous reverrons ces considérations dans le chapitre sur les estimations.

En pratique on n'examine pas forcément la limite, mais une valeur approchée à n « grand » . Nous verrons dans la suite des critères rendant valides ces considérations.

5.2 Approximation d'une loi binomiale par une loi normale

Une application de ce théorème consiste à considérer des variables de Bernoulli indépendantes de paramètre p. Soit une suite $(X_i)_{i\in\mathbb{N}^*}$ une suite de telles variables. En reprenant les notations précédentes, la stabilité de la loi binomiale donne $S_n = n\overline{X_n} \hookrightarrow \mathscr{B}(n,p)$.

Toujours avec les notations précédentes, on a $m = p$ (espérance de $\mathscr{B}(p)$) et $\sigma^2 = pq$ (variance de $\mathscr{B}(p)$). Le théorème central limite montre alors que la suite de variables $\sqrt{n}\left(\frac{\overline{X_n} - p}{\sqrt{pq}}\right)$ converge en loi vers une variable suivant $\mathscr{N}(0,1)$.

On réécrit alors : $\sqrt{n}\left(\frac{\overline{X_n} - p}{\sqrt{pq}}\right) = \left(\frac{n\overline{X_n} - np}{\sqrt{npq}}\right)$; avec $n\overline{X_n} \hookrightarrow \mathscr{B}(n,p)$.
Ceci justifie donc que :

Théorème 8. *Soit (S_n) une suite de variables aléatoires telles que $S_n \hookrightarrow \mathscr{B}(n,p)$ (avec $p \in]0,1[$, et $q = 1-p$). Alors la suite de variables*

$$\left(\frac{S_n - np}{\sqrt{npq}}\right)_{n\in\mathbb{N}^*}$$

converge en loi vers une variable suivant $\mathscr{N}(0,1)$.

Remarque 7. On note que $\frac{S_n - np}{\sqrt{npq}}$ est la « version centrée-réduite » de S_n (on retranche l'espérance, puis on divise par l'écart-type).

Remarque 8. Ici aussi, en repassant à des variables gaussiennes d'espérance non nulle et de variance quelconque, on pourra dire qu'à n grand, la loi $\mathscr{B}(n,p)$ peut être approchée par la loi $\mathscr{N}(np,npq)$. En pratique, on considère valide cette approximation sous les conditions suivantes : $\begin{cases} n \geqslant 30 \\ np \geqslant 15 \\ npq \geqslant 5 \end{cases}$.

Remarque 9. Notons que si on se souvient qu'il existe une approximation valide de la loi binomiale par une loi normale, les paramètres de celle-ci sont tels qu'on obtient les bonnes valeurs de l'espérance et de la variance : donc $\mathscr{N}(np,npq)$.

5.3 Approximation d'une loi de Poisson par une loi normale

On peut effectuer un raisonnement similaire sur une loi de Poisson, pour laquelle on dispose aussi de résultats de stabilité. Soit (T_n) une suite de variables aléatoires telles que $\mathrm{T}_n \hookrightarrow \mathscr{P}(n\alpha)$.Le théorème central limite permet cette fois d'obtenir :

Théorème 9. *Soit* (T_n) *une suite de variables aléatoires telles que* $\mathrm{T}_n \hookrightarrow \mathscr{P}(n\alpha)$. *Alors la suite de variables*

$$\left(\frac{\mathrm{T}_n - n\alpha}{\sqrt{n\alpha}}\right)_{n\in\mathbb{N}^*}$$

converge en loi vers une variable suivant $\mathscr{N}(0,1)$.

Remarque 10. Ici aussi, $\dfrac{\mathrm{T}_n - n\alpha}{\sqrt{n\alpha}}$ est la « version centrée-réduite » de T_n.

Corollaire 10. *Ce dernier résultat peut être utilisé pour approximer une loi de Poisson* $\mathscr{P}(\lambda)$ *pour* λ *grand (en identifiant* λ *à* $n\alpha$. *On obtient donc que, à* λ *grand, on peut approximer la loi* $\mathscr{P}(\lambda)$ *par la loi* $\mathscr{N}(\lambda,\lambda)$.
En pratique, on considère cette approximation valable pour $\lambda \geqslant 15$.

Remarque 11. Et ici encore, si on se souvient qu'il existe une approximation valide de la loi de Poisson $\mathscr{P}(\lambda)$ par une loi normale, les paramètres de celle-ci sont tels qu'on obtient les bonnes valeurs de l'espérance et de la variance : donc $\mathscr{N}(\lambda,\lambda)$.

5.4 Loi discrète ≃ loi à densité ??

On voit dans ces deux derniers exemple qu'on approxime une loi discrète (binomiale / de Poisson) par une loi à densité (normale).
Or si $\mathrm{S}_n \hookrightarrow \mathscr{B}(n,p)$ est approximée par la loi de $\mathrm{N} \hookrightarrow \mathscr{N}(np,npq)$ il est clair qu'on ne peut pas approximer, pour $k \in \mathbb{N}$ quelconque $\mathbb{P}(\mathrm{S}_n = k)$ par $\mathbb{P}(\mathrm{N} = k)$ (cette dernière proba est nulle car N est à densité).
La bonne approximation est $\mathbb{P}(\mathrm{S}_n = k) \simeq \mathbb{P}(k - 0.5 \leqslant \mathrm{N} \leqslant k + 0.5)$.

Annexe : Table de valeurs de Φ

Fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0; 1)$.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\,\mathrm{d}t \quad \text{et} \quad \Phi(-x) = 1 - \Phi(x).$$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	**0.9236**	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Exemple de lecture : $\Phi(1,43) \simeq 0,9236$ (en gras).

Démonstrations

Théorème 1 (Inégalité de Markov)

Soit X une variable aléatoire **positive**, admettant une espérance. On a :

$$\forall\, t > 0,\ \mathbb{P}(X \geqslant t) \leqslant \frac{\mathbb{E}(X)}{t}$$

Démonstration. On commence par le cas d'une variable discrète. Soit $t > 0$. Comme $\mathbb{E}(X)$ existe, on a

$$\begin{aligned}
\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\,\mathbb{P}(X = x) = \sum_{\substack{x \in X(\Omega)\\ x<t}} x\,\mathbb{P}(X = x) + \sum_{\substack{x \in X(\Omega)\\ x \geqslant t}} x\,\mathbb{P}(X = x) & \\
\geqslant \sum_{\substack{x \in X(\Omega)\\ x \geqslant t}} x\,\mathbb{P}(X = x) & \qquad \text{(par positivité de X)}\\
\geqslant t \sum_{\substack{x \in X(\Omega)\\ x \geqslant t}} \mathbb{P}(X = x) & \qquad (x \geqslant t \text{ sur le domaine de sommation})\\
= t\,\mathbb{P}(X \geqslant t) &
\end{aligned}$$

Finalement, on a $\mathbb{E}(X) \geqslant t\,\mathbb{P}(X \geqslant t)$, ce qui permet de conclure en divisant par $t > 0$.
Dans le cas d'une variable à densité c'est un calcul similaire en remplaçant les sommes par des intégrales. Cette fois l'hypothèse se traduit par le fait que la densité de X est nulle sur $\mathbb{R}_-^*$; on a alors

$$\mathbb{E}(X) = \int_0^{+\infty} x f(x)\,\mathrm{d}x = \int_0^t x f(x)\,\mathrm{d}x + \int_t^{+\infty} x f(x)\,\mathrm{d}x \geqslant \int_t^{+\infty} x f(x)\,\mathrm{d}x \geqslant t \int_t^{+\infty} f(x)\,\mathrm{d}x = t\,\mathbb{P}(X \geqslant t)$$

et on conclut de la même manière.

On peut également utiliser des variables indicatrices.
Si $t > 0$ on considère la variable $Y_t = X \times \mathbb{1}_{(X \geqslant t)}$; cette variable est donc telle que :

- Si $X(\omega) < t$ alors $Y_t(\omega) = 0$ (car l'indicatrice vaut 0)
- Si $X(\omega) \geqslant t$ alors $Y_t(\omega) = X(\omega) \geqslant t$ (car l'indicatrice vaut 1)

Observons alors que :

- Comme X est à valeurs positives on a toujours $Y_t(\omega) \leqslant X(\omega)$;
- On a aussi $Y_t \geqslant t\mathbb{1}_{(X \geqslant t)}$: si $X(\omega) < t$, $Y_t(\omega) = 0 = t\mathbb{1}_{(X \geqslant t)}$; et si $X(\omega) \geqslant t$, $Y_t(\omega) = X(\omega) \geqslant t = t\mathbb{1}_{(X \geqslant t)}$.

Par croissance de l'espérance, $\mathbb{E}(X) \geqslant \mathbb{E}(Y_t)$; et $\mathbb{E}(Y_t) \geqslant \mathbb{E}\left(t\mathbb{1}_{(X \geqslant t)}\right) = t\mathbb{E}\left(\mathbb{1}_{(X \geqslant t)}\right) = t\mathbb{P}(X \geqslant t)$.
La conjonction de ces deux inégalités donne $\mathbb{E}(X) \geqslant t\mathbb{P}(X \geqslant t)$ d'où Markov. □

Théorème 2 (inégalité de Bienaymé-Tchebychev)

Soit X une variable admettant une variance. Alors elle admet aussi une espérance, et on a :

$$\mathbb{P}\big(|X - \mathbb{E}(X)| \geqslant t\big) \leqslant \frac{V(X)}{t^2}$$

Démonstration. C'est une conséquence de Markov. Définissons $Y = (X - \mathbb{E}(X))^2$. Comme X admet une variance, Y admet une espérance et on a par définition $\mathbb{E}(Y) = V(X)$.
On a ensuite : $\mathbb{P}\big(|X - E(X)| \geqslant t\big) = \mathbb{P}\big((X - \mathbb{E}(X))^2 \geqslant t^2\big) = \mathbb{P}\big(Y \geqslant t^2\big)$.
Par Markov (Y est bien à valeurs positives), on a : $\mathbb{P}\big(Y \geqslant t^2\big) \leqslant \dfrac{\mathbb{E}(Y)}{t^2}$.
On a finalement montré : $\mathbb{P}\big(|X - \mathbb{E}(X)| \geqslant t\big) \leqslant \dfrac{V(X)}{t^2}$. □

Théorème 4 (convergence ne loi pour une variable discrète)

Soient $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires discrètes, à valeurs dans $\mathbb{Z}$. Alors $X_n \xrightarrow{\mathscr{L}} X$ si et seulement si :

$$\forall\, k \in \mathbb{Z},\ \lim_{n\to+\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$$

Démonstration. On peut démontrer le sens direct. Supposons que $X_n \xrightarrow{\mathscr{L}} X$. Pour $k \in \mathbb{Z}$, comme $X(\Omega) \subset \mathbb{Z}$, on peut écrire

$$\mathbb{P}(X = k) = F_X(k + 1/2) - F_X(k - 1/2)$$

De même, comme $X_n(\Omega) \subset \mathbb{Z}$, on a aussi :

$$\mathbb{P}(X_n = k) = F_{X_n}(k + 1/2) - F_{X_n}(k - 1/2)$$

Les fonctions F_{X_n} et F_X étant continues en $k \pm 1/2$, on a, par passage à la limite, $\mathbb{P}(X_n = k) \to \mathbb{P}(X = k)$.
Le sens réciproque est admis. □

Proposition 5 (approximation binomiale d'une loi de Poisson)

Soient $\lambda > 0$, et X_n des variables aléatoires suivant $\mathscr{B}(n, \lambda/n)$.
Alors $X_n \xrightarrow{\mathscr{L}} X$, avec $X \hookrightarrow \mathscr{P}(\lambda)$.

Démonstration. Toutes les variables considérées sont à valeurs dans $\mathbb{N}$: il suffit de montrer que :

$$\forall\, k \in \mathbb{N},\ \lim_{n\to+\infty} \mathbb{P}(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

Soit donc $k \in \mathbb{N}$ fixé.
Pour $n \geqslant k$, on a

$$\begin{aligned}
\mathbb{P}(X_n = k) &= \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n(n-1)\dots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!}\left(1 - \frac{\lambda}{n}\right)^{-k}\frac{n(n-1)\dots(n-k+1)}{n^k}\left(1 - \frac{\lambda}{n}\right)^{n}
\end{aligned}$$

Pour $n \to +\infty$, et à k fixé :

- $\left(1 - \frac{\lambda}{n}\right)^{-k} \to 1$
- $n(n-1)\dots(n-k+1) \sim n^k$
- $\left(1 - \frac{\lambda}{n}\right)^{n} \to e^{-\lambda}$ (exercice classique en passant au ln).

ce qui donne finalement

$$\lim_{n\to+\infty} \mathbb{P}(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

On reconnaît bien la loi de Poisson de paramètre λ. □