

Chapitre 6. Espérance

31/03/2025

Programme

- 1 Combinatoire
- 2 Espaces probabilisés finis
- 3 Lois conjointes
- 4 Espaces probabilisés discrets
- 5 Variable aléatoire discrète
- 6 **Espérance**
- 7 Convergence et approximations
- 8 Statistiques descriptives

6.1. Espérance d'une v.a. r. d.

Définition

Soit X une v.a. r. dénombrable sur un E . $P. (\Omega, P)$. On suppose que le support de X est infini et s'écrit $X(\Omega) = \{x_n, n \in \mathbb{N}\}$. La v. a. X admet une espérance ssi la somme $\sum_{n \in \mathbb{N}} x_n P(X = x_n)$ est absolument convergente. On appelle espérance de X le réel

$$E(X) = \sum_{n \in \mathbb{N}} x_n P(X = x_n).$$

Le résultat suivant montre que la définition ci-dessus a un sens.

Théorème

Soit $(a_n)_{n \in \mathbb{N}}$ une suites réelles. Si la série $\sum_{n=0}^{\infty} a_n$ est absolument convergente, alors pour toute bijection $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ la série $\sum_{n=0}^{\infty} a_{\varphi(n)}$ est convergente. De plus, la somme est indépendante de φ . On la note $\sum_{n=0}^{\infty} a_n$.

Exemples

- Espérance d'une loi géométrique sur $X(\Omega) = \mathbb{N}^*$. Soit $X \sim \mathcal{G}^*(p)$.
Alors

$$E(X) = \sum_{k=1}^{\infty} kpq^{k-1} = p\left(\sum_{k \geq 1} q^k\right)' = \frac{p}{(1-q)^2} = \frac{1}{p}.$$

- Espérance d'une loi géométrique sur $X(\Omega) = \mathbb{N}$. Soit $X \sim \mathcal{G}(p)$.
Alors

$$E(X) = \sum_{k=0}^{\infty} kpq^k = qp \sum_{k \geq 1} kq^{k-1} = \frac{qp}{(1-q)^2} = \frac{q}{p}.$$

- Espérance d'une loi de Poisson. Si $X \sim \mathcal{P}(\lambda)$, alors

$$E(X) = \sum_{n=0}^{\infty} n \frac{\lambda^n}{n!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = \lambda.$$

Un contre exemple

Considérons la v.a. X à valeurs dans $\mathbb{N}^*$ de loi de probabilité donnée par

$$\forall n \in \mathbb{N}^* \quad P(X = n) = \frac{1}{n(n+1)}.$$

Est-ce que X admet une espérance ?

On voit que la série

$$\sum_{n \geq 1} nP(X = n) = \sum_{n \geq 1} \frac{1}{n+1}$$

est divergente. Cette v.a. n'admet donc pas d'espérance.

6.2. Propriétés de l'espérance

Définition

Soit I un ensemble dénombrable d'indices. Soit $(u_i)_{i \in I}$ une famille de complexes indexée par I . On dit que la famille $(u_i)_{i \in I}$ est sommable si la famille $(|u_i|)_{i \in I}$ est sommable.

Théorème (théorème de transfert)

Soit X une variable aléatoire réelle discrète sur (Ω, P) à valeurs dans un ensemble non vide E et f une fonction de E dans $\mathbb{R}$. La variable $f(X)$ admet une espérance ssi la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable et dans ce cas

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x)P(X = x).$$

Dém. Pour $y \in f(X)(\Omega)$, l'évènement $\{f(X) = y\}$ est la réunion disjointe des évènements $\{X = x\}$ où $x \in f^{-1}(y) \subset X(\Omega) \subset E$. Pour $y \in f(X)(\Omega)$, $f^{-1}(y)$ est une partie de $X(\Omega)$ qui est au plus dénombrable et est donc au plus dénombrable puis

$$P(f(X) = y) = \sum_{x \in f^{-1}(y)} P(X = x).$$

De plus, Ω est la réunion disjointe des évènements $\{f(X) = y\}$, $y \in f(X)(\Omega)$.

- Supposons que la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable. Pour $y \in f(X)(\Omega)$ donné, la sous-famille $(f(x)P(X = x))_{x \in f^{-1}(y)}$ est sommable et

$$\sum_{x \in f^{-1}(y)} f(x)P(X = x) = y \sum_{x \in f^{-1}(y)} P(X = x) = yP(f(X) = y) \quad (\star)$$

puis d'après le théorème de sommation par paquets, la famille $(yP(f(X) = y))_{y \in f(X)(\Omega)}$ est sommable ou encore la variable $f(X)$ admet une espérance et de plus (par $(\star)$)

$$\begin{aligned} E(f(X)) &= \sum_{y \in f(X)(\Omega)} yP(f(X) = y) \\ &= \sum_{y \in f(X)(\Omega)} \sum_{x \in f^{-1}(y)} f(x)P(X = x) \\ &= \sum_{x \in X(\Omega)} f(x)P(X = x). \end{aligned}$$

- Supposons que $f(X)$ admette une espérance ou encore supposons la famille $(yP(f(X) = y))_{y \in f(X)(\Omega)}$ sommable. Montrons que la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable.

Puisque la famille $(|f(x)|P(X = x))_{x \in X(\Omega)}$ est réelle positive, on peut regrouper par paquets et on obtient

$$\begin{aligned} \sum_{x \in X(\Omega)} |f(x)|P(X = x) &= \sum_{y \in f(X)(\Omega)} |y| \sum_{x \in f^{-1}(y)} P(X = x) \\ &= \sum_{y \in f(X)(\Omega)} |y|P(f(X) = y) < +\infty. \end{aligned}$$

Mais alors, d'après le paragraphe précédent, la variable $f(X)$ admet une espérance et de plus

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x)P(X = x).$$



Conséquences du théorème de transfert

Théorème (Linéarité de l'espérance)

Soient X et Y deux v.a. réelles discrètes sur (Ω, P) . Soient a et b deux réels. Si X et Y admettent toutes deux une espérance, alors il en va de même de $aX + bY$. Dans ce cas

$$E(aX + bY) = aE(X) + bE(Y).$$

Théorème (positivité de l'espérance, croissance de l'espérance)

- 1** *Soit X une v.a. sur un espace (Ω, P) à valeurs réelles admettant une espérance. Si X est positive, alors $E(X) \geq 0$.*
- 2** *Soient X et Y deux v. a. sur un espace (Ω, P) à valeurs réelles, admettant toutes deux une espérance. Si $X \leq Y$ (c'est-à-dire si $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$), alors $E(X) \leq E(Y)$.*

6.3. Variance d'une v.a. réelle discrète infinie

Définition

Soit X une v.a. r. d. et $p \in \mathbb{N}^$. Si X^p admet une espérance, i.e., $E(|X|^p) < \infty$, alors on appelle moment d'ordre p de X le réel $E(X^p)$. On note $\mathcal{L}^p(\Omega, \mathcal{P})$ l'ensemble des v.a. r. d. définies sur l'espace probabilisé $(\Omega, \mathcal{P})$ admettant un moment d'ordre p .*

Théorème

Si une v.a.r. d. X admet un moment d'ordre 2 alors elle admet un moment d'ordre 1. En terme d'ensembles $\mathcal{L}^2(\Omega, \mathcal{P}) \subset \mathcal{L}^1(\Omega, \mathcal{P})$.

dém.

Soit $X(\Omega) = \{x_n, n \in \mathbb{N}\}$. For tout $x \in \mathbb{R}$ on a $|x| \leq 1 + x^2$, donc

$$\forall n \in \mathbb{N} \quad |x_n|P(X = x_n) \leq (1 + x_n^2)P(X = x_n)$$

Par linéarité, la série à termes positifs $(1 + x_n^2)P(X = x_n)$ est convergente. D'après le théorème de comparaison, la série $E(|X|)$ est convergente. $\square$

Variance d'une v.a.r.d. infinie

Déf. et thm.

Soit X une v.a.r.d. admettant un moment d'ordre 2. La variable aléatoire réelle discrète $(X - E(X))^2$ admet une espérance qu'on appelle la variance de X . De plus

$$V(X) = E(X^2) - E(X)^2.$$

dém.

Comme X admet un moment d'ordre 2, elle admet un moment d'ordre 1. On a

$$(X - E(X))^2 = X^2 - 2E(X)X + E(X)^2$$

Donc par linéarité $(X - E(X))^2$ admet une espérance. Par calcul simple, on trouve la formule ci-dessus. □

On vérifie aussi que pour tous $a, b \in \mathbb{R}$

$$V(X + b) = V(X) \quad V(aX) = a^2 V(X).$$

6.4. Fonctions génératrices

Définition

Soit X une v.a. sur un espace $(\Omega, \mathcal{P})$ à valeurs dans $\mathbb{N}$. On définit sa fonction génératrice (fg) g_X par la série entière

$$g_X(t) = \sum_{n=0}^{\infty} t^n P(X = n).$$

Comme la série est convergente en $t = 1$, le rayon de convergence $R \geq 1$.
On rappelle le résultat suivant du cours d'Analyse.

Théorème

La fg g_X est définie au moins sur un intervalle $] - R, R[$ avec $R \geq 1$. De plus, la convergence de la série est normale sur tout intervalle compact inclus dans $] - R, R[$ et g_X est de classe $\mathbb{C}^\infty$ sur $] - R, R[$ et ses dérivées successives sont obtenues par dérivation terme à terme de la série entière.

Théorème

Soit X une v.a. sur $(\Omega, \mathcal{P})$, à valeurs dans $\mathbb{N}$, et soit g_X sa fonction génératrice avec $R > 1$. Alors g_X est indéfiniment dérivable en 1, et pour tout entier k , la variable aléatoire $X(X-1)\dots(X-k+1)$ admet une espérance donnée par

$$g_X^{(k)}(1) = E(X(X-1)\dots(X-k+1)).$$

En particulier $g_X'(1) = E(X)$, $g_X''(1) = E(X(X-1))$ et

$$g_X''(1) + g_X'(1) = E(X^2).$$

Ce résultat est une conséquence immédiate du théorème de dérivation des séries entières

$$g_X(t) = \sum_{n \geq 0} P(X = n)t^n$$

$$g'_X(t) = \sum_{n \geq 1} P(X = n)nt^{n-1}$$

$$\dots = \dots$$

$$\forall k \in \mathbb{N} \quad g_X^{(k)}(t) = \sum_{n \geq k} P(X = n)n(n-1)\cdots(n-k+1)t^{n-k}.$$

On en déduit que

$$\begin{aligned} g_X^{(k)}(1) &= \sum_{n \geq k} P(X = n)n(n-1)\cdots(n-k+1) \\ &= E(X(X-1)\cdots(X-k+1)). \end{aligned}$$



Exemple (**Variance d'une loi de Poisson**) Soit $X \sim \mathcal{P}(\lambda)$. Alors

$$g_X(t) = \sum_{n \geq 0} e^{-\lambda} \frac{\lambda^n t^n}{n!} = e^{\lambda(t-1)}.$$

avec $R = +\infty$. On a

$$g'_X(t) = \lambda e^{\lambda(t-1)}, \quad g''_X(t) = \lambda^2 e^{\lambda(t-1)}$$

Il s'en suit que $E(X) = g'_X(1) = \lambda$,

$$E(X^2) = g'_X(1) + g''_X(1) = \lambda^2 + \lambda,$$

et $V(X) = E(X^2) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

Variance de la loi géométrique $\mathcal{G}^*(p)$

Soit $X \sim \mathcal{G}^*(p)$, i.e., $X(\Omega) = \mathbb{N}^*$ et $P(X = k) = pq^{k-1}$ avec $q = 1 - p$.

Alors

$$g_X(t) = \sum_{n \geq 1} q^{n-1} p t^n = \frac{pt}{1 - qt} \quad (|t| < 1/q)$$

On a

$$g'_X(t) = \frac{p}{(1 - qt)^2} \implies E(X) = g'_X(1) = \frac{1}{p};$$

$$g''_X(t) = \frac{2pq}{(1 - qt)^3} \implies E(X(X - 1)) = g''_X(1) = \frac{2q}{p^2}$$

Donc $E(X^2) = g'_X(1) + g''_X(1) = \frac{1}{p} + \frac{2q}{p^2}$ et

$$V(X) = \frac{1}{p} + \frac{2q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

Un exemple type : le nombre d'essais pour le premier succès dans les épreuves de Bernoulli.

Variance de la loi géométrique $\mathcal{G}_0(p)$

Soit $X \sim \mathcal{G}_0(p)$, i.e., $X(\Omega) = \mathbb{N}$ et $P(X = k) = pq^k$ avec $q = 1 - p$. Alors

$$g_X(t) = \sum_{n \geq 0} q^n p t^n = \frac{p}{1 - qt} \quad (|t| < 1/q)$$

On a

$$g'_X(t) = \frac{qp}{(1 - qt)^2} \implies E(X) = g'_X(1) = \frac{q}{p};$$

$$g''_X(t) = \frac{2pq^2}{(1 - qt)^3} \implies E(X(X - 1)) = g''_X(1) = \frac{2q^2}{p^2}$$

Donc $E(X^2) = g'_X(1) + g''_X(1) = \frac{q}{p} + \frac{2q^2}{p^2}$ et

$$V(X) = \frac{q}{p} + \frac{2q^2}{p^2} - \frac{q^2}{p^2} = \frac{q}{p^2}.$$

Un exemple type : le nombre d'échecs avant le premier succès dans les épreuves de Bernoulli.

Théorème (Caractérisation par la fonction génératrice)

La loi d'une variable aléatoire à valeurs dans $\mathbb{N}$ est entièrement déterminée par la donnée de sa fonction génératrice. Plus précisément, si X est une v.a. sur (Ω, P) telle que $X(\Omega) \subset \mathbb{N}$ et si g_X est sa fonction génératrice alors

$$\forall n \in \mathbb{N} \quad P(X = n) = \frac{g_X^{(n)}(0)}{n!}.$$

Théorème (Somme de deux v.a. a. indépendantes)

Soient X et Y deux v.a. sur (Ω, P) , dont les supports sont inclus dans $\mathbb{N}$, g_X et g_Y leurs fonctions génératrices, D_X et D_Y leurs domaines de définition respectifs. Si X et Y sont indépendants, alors g_{X+Y} est défini au moins sur $D_X \cap D_Y$ et

$$\forall t \in D_X \cap D_Y \quad g_{X+Y}(t) = g_X(t)g_Y(t).$$

Exemple 1– Somme de deux lois de Poisson indépendantes

Dans le cas où X et Y suivent des lois de Poisson de paramètres respectifs $\lambda > 0$ et $\mu > 0$, la fonction génératrice de $Z = X + Y$ s'écrit

$$\forall t \in \mathbb{R} \quad g_Z(t) = g_X(t)g_Y(t) = e^{\lambda(t-1)}e^{\mu(t-1)} = e^{(\lambda+\mu)(t-1)}$$

On reconnaît la fonction génératrice d'une loi de Poisson de paramètre $\lambda + \mu$. Ainsi, en vertu du théorème de caractérisation par la fg, nous pouvons affirmer que Z suit la loi de Poisson de paramètre $\lambda + \mu$.

On retrouve aussi la loi de la somme de deux lois binomiales indépendantes, etc.

La formule du binôme négatif $(1 - x)^{-\alpha}$

$$\forall n \in \mathbb{N} \quad \forall x \in]-1, 1[\quad \frac{1}{(1-x)^{n+1}} = \sum_{k=n}^{+\infty} \binom{k}{n} x^{k-n} = \sum_{k=0}^{+\infty} \binom{k+n}{n} x^k$$

Démonstration.

En dérivant successivement la série géométrique $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$:

$$\frac{1}{(1-x)^2} = \sum_{k=0}^{\infty} (k+1)x^k$$

$$\frac{2!}{(1-x)^3} = \sum_{k=0}^{\infty} (k+2)(k+1)x^k$$

$$\dots = \dots$$

$$\frac{n!}{(1-x)^{n+1}} = \sum_{k=0}^{\infty} (k+n)(k+n-1)\dots(k+1)x^k.$$

Par récurrence on démontre la formule. □

Loi de Pascal $\mathcal{P}(r, p)$

Une variable aléatoire X suit une loi de Pascal de paramètres r, p , où $r \in \mathbb{N}^*$, $0 < p < 1$ si

- $X(\Omega) = [r, r + 1, \dots, +\infty[$
- $P(X = k) = \binom{k-1}{r-1} p^r q^{k-r}$

Un exemple type : On joue à pile (succès) ou face (échec) avec une pièce truquée telle que la probabilité de sortir pile est p . Alors la variable aléatoire X définie comme le nombre de lancers nécessaires avant d'obtenir le $r^{\text{ième}}$ pile suit une loi de Pascal $\mathcal{P}(r, p)$.

Pour $r = 1$ on retrouve la loi géométrique sur $\mathbb{N}^*$. On vérifie que

$$g_X(t) = \frac{t^r p^r}{(1 - qt)^r} \implies E(X) = \frac{r}{p}, \quad V(X) = \frac{rq}{p^2}$$

Loi binomiale négative $\mathcal{J}(r, p)$

Une variable aléatoire X suit une loi binomiale négative de paramètres r, p , où $r \in \mathbb{N}^*$, $0 < p < 1$ si

- $X(\Omega) = \mathbb{N} \cup \{+\infty\}$
- $P(X = k) = \binom{k+r-1}{r-1} p^r q^k$

Un exemple type : On joue à pile ou face avec une pièce truquée telle que la probabilité de sortir pile est p . Alors la variable aléatoire X définie comme le nombre d'échecs nécessaires avant d'obtenir le $r^{\text{ième}}$ pile suit une loi binomiale négative $\mathcal{J}(r, p)$.

Pour $r = 1$ on retrouve la loi géométrique sur $\mathbb{N}$. On vérifie que

$$g_X(t) = \frac{p^r}{(1 - qt)^r} \implies E(X) = \frac{rq}{p}, \quad V(X) = \frac{rq}{p^2}$$

6.5. Covariances

Théorème

Soient X et Y deux v.a. r. d. sur (Ω, P) . Si X et Y admettent un moment d'ordre 2, alors XY admet une espérance et

$$E(XY) \leq \sqrt{E(X^2)}\sqrt{E(Y^2)}.$$

Dém. Soit $\Omega = \{\omega_n, n \in \mathbb{N}\}$, $P(\{\omega_n\}) = p_n$, $X(\omega_n) = x_n$ et $Y(\omega_n) = y_n$ pour $n \in \mathbb{N}$ (N.B. $XY(\omega_n) = x_n y_n$). Si l'un des $E(X^2)$ et $E(Y^2)$ est nul, par exemple, le premier, alors pour tout $n \in \mathbb{N}$, on a $x_n^2 p_n = 0$ et donc $x_n p_n = 0 \implies E(XY) = 0$.

Si ces deux termes sont non nuls, il résulte de l'inégalité

$$\forall a, b \in \mathbb{R}_+ \quad \sqrt{ab} \leq \frac{a+b}{2}$$

que

$$\frac{|x_n|}{\sqrt{E(X^2)}} \frac{|y_n|}{\sqrt{E(Y^2)}} \leq \frac{1}{2} \left(\frac{x_n^2}{E(X^2)} + \frac{y_n^2}{E(Y^2)} \right)$$

Après multiplication par p_n et sommations sur $\mathbb{N}$ on obtient

$$\frac{E(|XY|)}{\sqrt{E(X^2)} \cdot \sqrt{E(Y^2)}} \leq \frac{1}{2}(1 + 1) = 1.$$

Ce qui démontre l'inégalité demandée.



Déf. et thm.

Soit (X, Y) un couple de v. a. r. sur (Ω, P) . Si X et Y admettent un moment d'ordre 2, alors la v.a.r d. $(X - E(X))(Y - E(Y))$ admet une espérance. On l'appelle covariance de X et de Y et définit

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$$

On a aussi

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y).$$

Sous réserve d'existence, les propriétés de la covariance établies dans les univers finis restent vraies.

Théorème

Soient X , Y et Z trois v.a. r. d. sur $(\Omega, \mathcal{P})$ et $(a, b) \in \mathbb{R}^2$. On suppose que X , Y et Z admettent des moments d'ordre 2.

- i) $V(aX + bY) = a^2 V(X) + b^2 V(Y) + 2ab \operatorname{cov}(X; Y);$
- ii) $\operatorname{cov}(X, X) = V(X);$
- iii) $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X);$
- iv) $\operatorname{cov}(aX + bY, Z) = a \operatorname{cov}(X, Z) + b \operatorname{cov}(Y, Z);$
- v) *si X et Y sont indépendantes, alors $\operatorname{cov}(X, Y) = 0$.*

Coefficients de corrélation linéaire

Déf. et thm.

Soient X et Y deux v.a. r. d. sur (Ω, P) . On suppose que X et Y admettent un moment d'ordre 2 et que leur variances ne sont pas nulles. On appelle coefficient de corrélation linéaire de X et de Y le réel

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

On a $|\rho(X, Y)| \leq 1$. de plus si $|\rho(X, Y)| = 1$ alors il existe deux réels a et b tels que l'évènement $(Y = aX + b)$ soit certain : $P(Y = Ax + b) = 1$.

Les démonstrations sont identiques au cas fini.

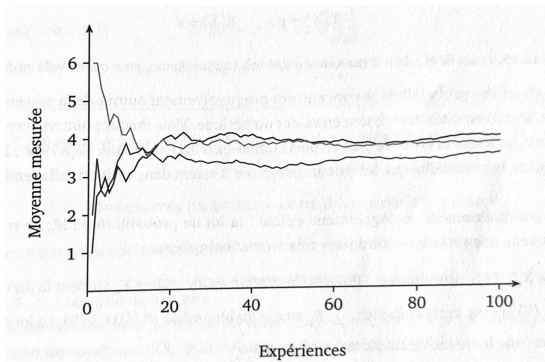
Chapitre 7. Convergence et approximations

7.1 Lois faibles des grands nombres

Réalisons une expérience aléatoire dont le résultat est une v.a. X . La loi de X est inconnue, mais on a bon espoir de la connaître de mieux en mieux au fur et à mesure de la répétition des expériences. A chacune des expériences correspond une v.a. : X_1, X_2, X_3 etc. jusqu'à X_n . On suppose que les n v.a. sont indépendantes et suivant une même loi (mais c'est une hypothèse !). En particulier, tout expérimentateur qui se respecte songe spontanément à calculer la moyenne des valeurs observées

$$M_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n).$$

Notez, et c'est important, que M_n est elle-même une v.a., somme de n variables aléatoires indépendantes. Il n'y a aucune raison de penser que M_n suit la même loi que X , et ce n'est d'ailleurs généralement pas le cas !



Evolution de la moyenne lors de 3 répétitions d'une centaine d'expériences *lancer d'un dé de 6 faces honnête*. La moyenne semble s'approcher de l'espérance $7/2$ dans chacun des cas. Nous démontrons dans cette section que l'espérance d'une v.a. correspond bien à la valeur moyenne qu'on observe empiriquement lorsqu'on effectue un grand nombre de tirages de la variable.

Théorème (Inégalité de Markov)

Soit X une v.a. à valeurs positives d'espérance finie. Soit $a > 0$. Alors

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Dém.

Notons x_i les valeurs prises par X . Séparons-les en deux parties : notons $I_{<a}$ l'ensemble des indices i tels que $x_i < a$ et $I_{\geq a}$ l'ensemble des indices i tels que $x_i \geq a$. Alors

$$E(X) = \sum_{i \in I} x_i P(X = x_i) = \sum_{i \in I_{<a}} x_i P(X = x_i) + \sum_{i \in I_{\geq a}} x_i P(X = x_i)$$

Comme les valeurs x_i sont positives, on peut minorer la première somme par 0 et minorer tous les x_i de la seconde somme par a . □

Théorème (Inégalité de Bienaymé-Tchebychev)

Soit X une v.a. ayant une espérance et une variance finie, et $a > 0$. Alors

$$P(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}.$$

Dém. La v.a. $(X - E(X))^2$ est positive et admet une espérance. On peut lui appliquer l'inégalité de Markov

$$P((X - E(X))^2 \geq a^2) \leq \frac{V(X)}{a^2}$$

L'évènement $[(X - E(X))^2 \geq a^2]$ est exactement $|X - E(X)| \geq a$, on obtient l'inégalité de B-T. □

Théorème (Loi faible des grands nombres)

Si $(X_k)_{k \in \mathbb{N}}$ une suite de v.a. indépendantes et de même loi, d'espérance m et de variance σ^2 finies. Soit $M_n = (X_1 + \dots + X_n)/n$. Alors, pour tout $\epsilon > 0$,

$$\lim_{n \rightarrow +\infty} P(|M_n - m| \geq \epsilon) = 0.$$

Dém.

Par linéarité de l'espérance, $E(M_n) = m$. De plus, comme les variables X_i sont indépendantes, $V(M_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \sigma^2/n$.

Appliquons l'inégalité de B-T à la variable M_n et $a = \epsilon$:

$$P(|M_n - m| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2} \xrightarrow{n \rightarrow \infty} 0$$



Exemple ou exercice de référence

On lance une pièce de monnaie supposée honnête $n \geq 1$ fois de suite.

- 1) À l'aide de l'inégalité de B-T, trouver une condition suffisante sur l'entier n pour que la fréquence de Face obtenus soit strictement comprise entre 0,4 et 0,6 avec une probabilité supérieure ou égale à 0,95.
- 2) Au bout de 1000 lancers, on observe une proportion de Pile de 0,65. La pièce est-elle vraiment honnête ?

Solution

- 1) Le nombre X_n de Face obtenus suit une loi binomiale $\mathcal{B}(n, 1/2)$. La fréquence de Face F_n étant égale à X_n/n , on a $E(F_n) = 1/2$ et $V(F_n) = 1/(4n)$. On cherche à estimer la probabilité de l'évènement

$$(F_n \in]0,4; 0,6[) = (|F_n - 0,5| < 0,1)$$

Or, d'après l'inégalité de B-T

$$P(|F_n - 0,5| < 0,1) = 1 - P(|F_n - 0,5| \geq 0,1) \geq 1 - \frac{V(F_n)}{0,1^2} \geq 1 - \frac{25}{n}$$

Pour que l'évènement soit de proportion supérieure à 0,95 il suffit que $\frac{25}{n} \leq 0,05$ soit $n \geq 500$.

2) Ici $n \geq 500$, donc on peut appliquer le résultat de la question précédente. Comme

$$P(F_n \geq 0,65) \leq P(F_n \notin]0,4; 0,6]) \leq 0,05$$

on peut penser que la pièce n'est pas parfaitement équilibrée.

7.2. Les mesures de probabilité continues

Soit Ω un intervalle de $\mathbb{R}$ et soit $\mathcal{B}$ l'ensemble de toutes des parties de Ω que l'on peut former à partir des intervalles inclus dans Ω , leurs intersections et leurs réunions finies ou dénombrables). Cet ensemble, appelé la **tribu des boréliens**, contient toutes les parties qu'on peut aisément imaginer mais est néanmoins strictement inclus dans $\mathcal{P}(\Omega)$. Définir une mesure de probabilité sur $(\Omega, \mathcal{B})$ signifie la définir pour tous les éléments de $\mathcal{B}$ avec les mêmes axiomes que dans le cas discret. Pour ce faire, il suffit de la définir sur les intervalles de Ω .

Définition

Soit P une mesure de probabilité définie sur $(\Omega, \mathcal{B})$. On dit que P est une mesure à densité s'il existe une fonction réelle f telle que

$$\forall b \geq a, \quad P([a, b]) = \int_a^b f(x) dx.$$

Propriétés de fonction de densité

- La fonction de densité f doit être telle que les intégrales $\int_a^b f(x)dx$ soient toutes bien définies. C'est le cas lorsque f est continue ou encore continue par morceaux.
- Comme P est une mesure de probabilité, elle doit aussi satisfaire

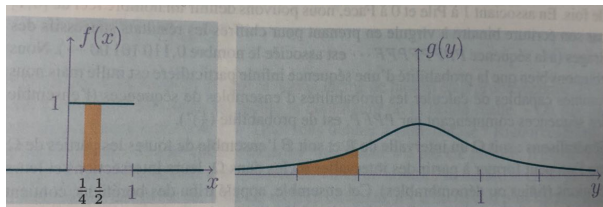
$$\forall x \in \mathbb{R}, \quad f(x) \geq 0 \quad \text{et} \quad \int_{-\infty}^{+\infty} f(x)dx = 1$$

où

$$\int_{-\infty}^{+\infty} f(x)dx = \lim_{A \rightarrow -\infty, B \rightarrow +\infty} \int_A^B f(x)dx.$$

Pour des fonctions positives, la limite (finie ou infinie) de l'intégrale est garantie.

Exemple : Loi uniforme et loi de Cauchy



1) La fonction f est la fonction de densité de la loi uniforme sur $[0, 1]$. Elle est bien positive et l'aire totale sous la courbe est 1. Comme f est constante, nous lisons instantanément le fait que les probabilités sont équiréparties dans l'intervalle $[0, 1]$.

2) La fonction g est définie sur $\mathbb{R}$ par $g(y) = \frac{1}{\pi} \frac{1}{1+y^2}$. Vérifions qu'il s'agit bien d'une fonction de densité : elle est positive et sa primitive $\frac{1}{\pi} \arctan y$ est une fonction usuelle, donc

$$\int_{-\infty}^{+\infty} \frac{1}{\pi} \frac{1}{1+y^2} dy = \frac{1}{\pi} [\arctan y]_{-\infty}^{+\infty} = 1$$

Les v.a. à densité

Définition

Soit X une v.a. r. définie sur $(\Omega, \mathcal{B})$. On dit que X est une variable continue ou variable à densité si sa mesure de probabilité P_X est une mesure à densité, c'est-à-dire s'il existe une fonction de densité f_X telle que pour tout intervalle $[a, b]$ de $\bar{\mathbb{R}}$

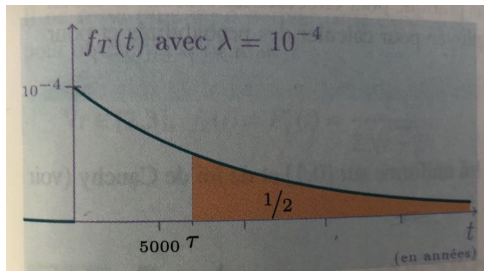
$$P(X \in [a, b]) = P_X([a, b]) = \int_a^b f_X(x) dx.$$

On dit alors que X suit la loi de densité f_X .

Remarque

La fonction f_X représente la loi de la variable : les valeurs qu'elle prend et leurs probabilités associées. Son graphe se lit presque comme la loi d'une variable discrète. Les valeurs de X se lisent en abscisse, mais attention, les probabilités ne se lisent pas sur l'axe des ordonnées. On ne calcule que des probabilités d'intervalles et celles-ci se représentent comme des aires sous la courbe de f_X .

Loi exponentielle de paramètre λ



Exemple. Le temps que met un atome radioactif à se désintégrer peut être modélisé par une v.a. T de fonction de densité $f_T(t) = \lambda e^{-\lambda t}$ pour $t > 0$, avec $\lambda > 0$. La demi-vie τ , temps au bout duquel la moitié des atomes est désintégrée, est caractérisée par

$$P(T > \tau) = \frac{1}{2} \quad (\tau \text{ est la médiane de la variable } T).$$

Ainsi $\int_{\tau}^{+\infty} \lambda e^{-\lambda t} dt = 1/2$, et on obtient $\tau = \frac{\ln(2)}{\lambda} \approx 7000$ ans.

Fonction de répartition

Définition

Soit X une v.a. de densité f_X . On appelle fonction de répartition de X la fonction

$$\begin{aligned} F_X : \mathbb{R} &\rightarrow [0, 1] \\ x &\mapsto P_X([-\infty, x]). \end{aligned}$$

Autrement dit $F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t) dt$.

Remarque

La densité est la bonne façon de représenter la loi d'une v.a., son graphe permet de bien la visualiser. Mais elle n'est pas pratique pour effectuer les calculs. C'est en pratique la fonction de répartition qui est employée pour calculer des probabilités et pour simuler des variable aléatoire.

Proposition

Soit X une v.a. de densité f_X et de fonction de répartition F_X . Reformulons notre définition.

- *La fonction F_X est la primitive de f_X qui s'annule en $-\infty$. En particulier $F'_X = f_X$.*
- *$\forall b > a, P(X \in [a, b]) = F_X(b) - F_X(a)$.*
- *Comme f_X est positive et d'intégrale 1, la fonction F_X est croissante et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.*

Exemple. On tire un nombre ω dans $[0, 1]$ selon la loi uniforme et on définit $Z(\omega) = \omega^2 + 2$. Quelle est la densité de Z ?

Cherchons d'abord F_Z : soit $t \in [2, 3]$. Alors

$$F_Z(t) = P(Z \leq t) = P(\omega \leq \sqrt{t-2}) = \sqrt{t-2}.$$

Il ne reste plus qu'à dériver cette fonction :

$$\forall t \in [2, 3], \quad f_Z(t) = F'_Z(t) = \frac{1}{2\sqrt{t-2}}.$$

Remarque

On pourra attribuer une valeur quelconque pour $f_Z(2)$ car cela ne modifie pas la loi de probabilité.

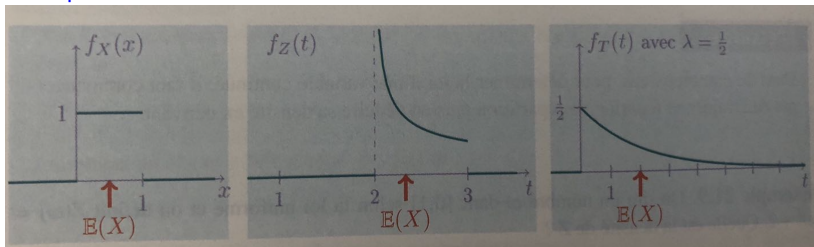
Espérance, variance et propriétés

Définition

Soit X une v.a. de densité f . Alors, si les intégrales suivantes sont convergentes, son espérance et sa variance sont :

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx \quad \text{et} \quad V(X) = \int_{-\infty}^{+\infty} (x - E(x))^2 f(x)dx.$$

Exemples dans le tableaunoir !



Définition

La loi normale de paramètre $m \in \mathbb{R}$ et $\sigma^2 > 0$, notée $\mathcal{N}(m, \sigma^2)$, est la loi de densité

$$\forall x \in \mathbb{R}, f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

On admet que

$$\int_{-\infty}^{+\infty} f = 1 \iff I := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-t^2/2} dt = 1$$

On vérifie d'abord la deuxième identité. Notons

$$\begin{aligned} I^2 &= \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2} dx \right) \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-y^2/2} dy \right) \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \exp\left(-\frac{x^2+y^2}{2}\right) dx dy \end{aligned}$$

Le passage aux coordonnées polaires en coordonnées cartésiennes est un changement de variables $(r, \theta) \rightarrow (x, y)$ défini par l'application suivante :

$$(r, \theta) \mapsto (r \cos \theta, r \sin \theta).$$

La matrice jacobienne au point (r, θ) est :

$$\begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

Le jacobien du passage en coordonnées polaires est donc :
 $r \cos^2 \theta + r \sin^2 \theta = r$ et on obtient

$$\begin{aligned} I^2 &= \frac{1}{2\pi} \int_0^{+\infty} \int_0^{2\pi} r \exp(-r^2/2) dr d\theta \\ &= \int_0^{+\infty} r \exp(-r^2/2) dr = 1. \end{aligned}$$

Proposition

La v.a. $X \sim \mathcal{N}(0, 1)$ est centrée, c'est-à-dire de l'espérance nulle, et réduite, c'est-à-dire de variance 1. De plus, la v.a. $-X \sim \mathcal{N}(0, 1)$.

Démonstration.

Le calcul de l'espérance est immédiate quand on a observé que $xf(x)$ est une fonction impaire. Le calcul de la variance se fait par une intégration par parties. Enfin, montrons que X et $-X$ ont la même loi. Pour tout $a, b \in \mathbb{R}$,

$$P[a \leq -X \leq b] = P[-b \leq X \leq -a] = \int_{-b}^{-a} f(x)dx,$$

qui est $P[a \leq X \leq b]$ par un changement de variable : $y = -x$ et le fait que f est une fonction paire. □

Utilisation des tables

Pour calculer $P[a \leq X \leq b]$ ou $P[X \leq a]$, on a recours au calcul numérique sur ordiinateur ou, plus simplement, à une table qui donne $P[X \leq a]$ pour tout décimal positif x à deux chiffres après la virgule. Puis il faut remarquer que

$$P[a \leq X \leq b] = P[X \leq b] - P[X \leq a]$$

puis on la table si a et b sont positifs. Pour trouver $P[X \leq -x]$ avec $x > 0$, on utilise le fait que X et $-X$ ont même loi (ou par la symétrie de la courbe) :

$$P[X \leq -x] = P[-X \geq x] = P[X \geq x] = 1 - P[X \leq x]$$

Exemple : On cherche à calculer $P[-1 \leq X \leq 1]$.

$$\begin{aligned} P[-1 \leq X \leq 1] &= P[X \leq 1] - P[X \leq -1] = P[X \leq 1] - (1 - P[X \leq 1]) \\ &= 2P[X \leq 1] - 1 = 2 \times 0.8413 - 1 = 0.6826 \end{aligned}$$

Exemple : On cherche $u \in \mathbb{R}$ tel que $P[-u \leq X \leq u] = 0.90$.

$$P[-u \leq X \leq u] = 2P[X \leq u] - 1$$

D'où $P[X \leq u] = 0.95$ et $u = 1.6446$.

Loi normale : cas général

Définition

Soient $m \in \mathbb{R}$ et $\sigma > 0$. Une v.a. X suit une loi normale $\mathcal{N}(m, \sigma^2)$ si sa densité f est donnée par

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad \text{pour tout } x \in \mathbb{R}.$$

Proposition

Si X suit une loi normale $\mathcal{N}(m, \sigma^2)$, $Z = \frac{X-m}{\sigma}$ suit une loi normale centrée réduite $\mathcal{N}(0, 1)$. De plus,

$$E[Z] = m, \quad \text{Var}(X) = \sigma^2.$$

Preuve.

soient deux réels $a < b$.

$$\begin{aligned} P[a \leq Z \leq b] &= P[a\sigma + m \leq X \leq b\sigma + m] \\ &= \int_{m+a\sigma}^{m+b\sigma} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx \end{aligned}$$

Changement de variable : $z = \frac{x-m}{\sigma}$

$$P[a \leq Z \leq b] = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$$

Donc $Z \sim \mathcal{N}(0, 1)$.

L'espérance et variance de X se découlent des propriétés par changement de variable linéaire : $X = \sigma Z + m$. $\square$

On peut ainsi facilement se ramener à la loi normale centrée réduite. Comme la fonction de répartition φ de la loi normale centrée réduite est tabulée, on se ramènera même systématiquement à cette loi pour calculer des probabilités. Plus précisément, si X est de loi $\mathcal{N}(m, \sigma)$, et a et b sont deux réels, on calcule $P[a \leq X \leq b]$ ainsi :

$$P[a \leq X \leq b] = P\left[\frac{a-m}{\sigma} \leq \frac{X-m}{\sigma} \leq \frac{b-m}{\sigma}\right] = P\left[\frac{a-m}{\sigma} \leq Z \leq \frac{b-m}{\sigma}\right]$$

où Z est de loi $\mathcal{N}(0, 1)$. Donc

$$P[a \leq X \leq b] = \varphi\left(\frac{b-m}{\sigma}\right) - \varphi\left(\frac{a-m}{\sigma}\right)$$

On peut aussi remarquer que si $Z \sim \mathcal{N}(0, 1)$, alors $-Z$ aussi et donc, pour $\mu > 0$,

$$P[-u \leq Z \leq u] = P[Z \leq u] - P[Z \geq u] = 2P[Z \leq u] - 1.$$

Exemple : Soit $X \sim \mathcal{N}(15, 4^2)$. On a

$$P[10 \leq X \leq 22] = P\left[\frac{10 - 15}{4} \leq \frac{X - 15}{4} \leq \frac{22 - 15}{4}\right] = P[-1,25 \leq Y \leq 1,75]$$

où $Y \sim \mathcal{N}(0, 1)$. Ainsi,

$$\begin{aligned} P[10 \leq X \leq 22] &= P[Y \leq 1,75] + P[Y \leq 1,25] - 1 \\ &= 0,9599 + 0,8944 - 1 = 0,8543. \end{aligned}$$

Quelques valeurs importantes

Soit $Y \sim \mathcal{N}(0, 1)$. Alors

$$P[-1 \leq Y \leq 1] = 0,6826$$

$$P[-2 \leq Y \leq 2] = 0,9544$$

$$P[-3 \leq Y \leq 3] = 0,9973$$

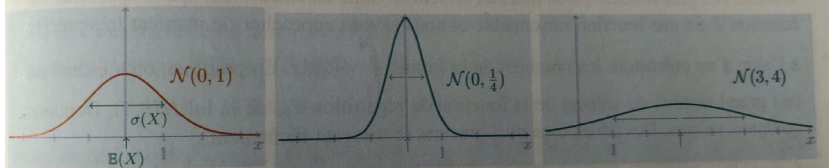
Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Alors $Y = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$ et

$$P[\mu - \sigma \leq X \leq \mu + \sigma] = 0,6826$$

$$P[\mu - 2\sigma \leq X \leq \mu + 2\sigma] = 0,9544$$

$$P[\mu - 3\sigma \leq X \leq \mu + 3\sigma] = 0,9973$$

Exemple !. Nous représentons les graphes des densités des lois normales $\mathcal{N}(0,1)$, $\mathcal{N}(0,\frac{1}{4})$ et $\mathcal{N}(3,4)$.



La première loi (en rouge) a été choisie comme loi normale de référence. Il s'agit de la loi normale centrée réduite, c'est-à-dire d'espérance nulle et de variance 1.

Proposition

Soit $X \sim \mathcal{N}(m, \sigma^2)$. Alors $E(X) = m$ et $V(X) = \sigma^2$.

Dém. Au Tableau.



Proposition

Soit $X \sim \mathcal{N}(m, \sigma^2)$ et $a, b \in \mathbb{R}$. Alors

$$\frac{X - m}{\sigma} \sim \mathcal{N}(0, 1).$$

Dém. Soit $X \sim \mathcal{N}(m, \sigma^2)$ et $Y = a + bX$ avec $b > 0$. Alors

$F_Y(t) = P(Y \leq t) = P(X \leq \frac{t-a}{b}) = F_X(\frac{t-a}{b})$. Dérivons :

$F'_Y(t) = f_Y(t) = \frac{1}{b} f_X(\frac{t-a}{b}) = \frac{1}{b\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-a-bm)^2}{2b^2\sigma^2}\right)$. En posant

$a = -m/\sigma$ et $b = 1/\sigma$ on trouve $F'_Y(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$



Calcul avec la loi $\mathcal{N}(0, 1)$

Nous ne savons pas intégrer la densité f de la loi $\mathcal{N}(0, 1)$. Or c'est indispensable pour calculer des probabilités. En pratique, on a estimé un très grand nombre de valeurs de la fonction de répartition $F_{\mathcal{N}}$ de la loi $\mathcal{N}(0, 1)$. Donnons quelques valeurs.

x	0	0,20	0,40	0,60	0,80	1,00	1,20	1,40	1,65	1,96	2,33	2,58	3,09
$F_{\mathcal{N}}(x)$	0,5	0,58	0,66	0,73	0,79	0,84	0,88	0,92	0,95	0,975	0,99	0,995	0,999

Les valeurs de $F_{\mathcal{N}}(x)$ pour $x < 0$ se déduisent par symétrie de f :

$$\forall x \in \mathbb{R}, \quad F_{\mathcal{N}}(x) = 1 - F_{\mathcal{N}}(-x).$$

Exemple

- Soit $X \sim \mathcal{N} \in (0, 1)$. Estimons $P(X < -1)$.
- Déterminons $\alpha \in \mathbb{R}$ tel que $P(X \in [-\alpha, \alpha]) = 0,95$. Il s'agit de l'intervalle de fluctuation de X à 95%.

Exemple

La taille (en centimètres) des Françaises est modélisée par une v.a. X de loi $\mathcal{N}(167, 40)$: elles mesurent en moyenne 1,67 m avec un écart-type de l'ordre de 0,063 m.

- *Quelle est la proportion de femmes de taille inférieure à 1,75 m ?
Si $X \sim \mathcal{N}(167, 40)$ alors $\frac{X-167}{\sqrt{40}} \sim \mathcal{N}(0, 1)$. Ainsi*

$$\begin{aligned} P(X < 175) &= P\left(\frac{X - 167}{\sqrt{40}} < \frac{175 - 167}{\sqrt{40}}\right) \\ &= F_N\left(\frac{8}{\sqrt{40}}\right) \approx F_N(1, 26) \approx 0,9 \end{aligned}$$

d'après notre table. La proportion est donc à l'ordre de 90%.

- Quel est l'intervalle de fluctuation de X à 99% ?
Cherchons β tel que $P(|X - 167| < \beta) = 0,99$ En s'y prenant comme ci-dessus :

$$P(|X - 167| < \beta) = 2F_N\left(\frac{\beta}{\sqrt{40}}\right) - 1$$

Nous en déduisons $F_N\left(\frac{\beta}{\sqrt{40}}\right) = 0,995$, donc $\frac{\beta}{\sqrt{40}} \approx 2,58$ et nous obtenons $\beta \approx 16$. Nous pouvons dire (avec un risque de 1% d'erreur) que les Françaises mesurent entre 1,51 m et 1,83 m.

7.3. Le théorème central limite (TCL)

La loi forte des grands nombres affirme que la suite des moyennes de tirages aléatoires indépendantes d'une même loi converge vers la moyenne théorique de cette loi. Le TCL permet de préciser la vitesse de convergence de cette suite. Plus précisément, il donne une estimation probabiliste de l'écart entre les termes de la suite et leur limite. C'est ce théorème qui permet, en statistiques, de définir la notion d'intervalle de confiance.

Nous considérons dans cette partie une suite X_k de v.a. Nous notons S_n et M_n la somme et la moyenne des n premiers termes :

$$S_n = \sum_{k=1}^n X_k \quad \text{et} \quad M_n = \frac{S_n}{n}.$$

Théorème (central limite, TCL)

Soit $(X_k)_{k \geq 1}$ une suite de v.a.i. (variables aléatoires indépendantes), de même loi, d'espérance m et de variance σ^2 finies. Quand n tend vers $+\infty$, la suite de v. a. $\frac{M_n - m}{\frac{\sigma}{\sqrt{n}}}$ converge en loi vers la loi normale $\mathcal{N}(0, 1)$, c'est-à-dire

$$\forall a < b, \quad \lim_{n \rightarrow +\infty} P \left(a \leq \frac{M_n - m}{\frac{\sigma}{\sqrt{n}}} \leq b \right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

En pratique, quand n est assez grand (i.e. $n \approx 30$) on approche les lois de M_n et S_n par une loi normale :

$$M_n \sim \mathcal{N}\left(m, \frac{\sigma^2}{n}\right), \quad (0.1)$$

$$S_n \sim \mathcal{N}(nm, n\sigma^2); \quad (0.2)$$

et

$$\frac{M_n - m}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1), \quad (0.3)$$

$$\frac{S_n - nm}{\sigma\sqrt{n}} \sim \mathcal{N}(0, 1). \quad (0.4)$$

Théorème (Théorème de Moivre-Laplace)

Si n est assez grand, il est possible d'approcher la loi binomiale $\mathcal{B}(n, p)$ par une loi normale :

$$\mathcal{B}(n, p) \sim \mathcal{N}(np, np(1 - p))$$

dém.

Rappelons qu'une v.a. de loi binomiale peut être vue comme une somme de v.a. de Bernoulli indépendantes. Soit (X_k) une suite de v.a.i. de même loi $\mathcal{B}(p)$. Ces variables sont d'espérance p et de variance $p(1 - p)$. En notant $S_n = \sum_{i=1}^n X_i$, alors $S_n \sim \mathcal{B}(n, p)$ et d'après TCL, lorsque n est assez grand, on peut considérer $S_n \sim \mathcal{N}(np, np(1 - p))$. $\square$

N.B. Dans la pratique l'approximation de la loi $\mathcal{B}(n, p)$ par la loi normale $\mathcal{N}(np, np(1 - p))$ est bonne si

$$n \geq 30, \quad np \geq 5, \quad n(1 - p) \geq 5.$$

Exemple

- *On lance une pièce équilibrée 100 fois. Quelle est la probabilité d'obtenir entre 47 et 53 Piles ?*
- *On lance maintenant la pièce 1000 fois et on cherche $P(470 \leq S_{1000} \leq 530)$.*

Exemple

Une mutuelle couvre 100 000 personnes contre un certain risque. La description de la loi de probabilité du risque est bien trop complexe mais une étude a permis de montrer que les dégâts annuels s'élèvent en moyenne à 37 euros par personne avec un écart-type de 250 euros (cela correspond à un risque provoquant de forts dégâts mais chez peu de personnes). A combien d'euros doit être fixée la cotisation pour que la mutuelle ait une probabilité inférieure à 0,001 de perdre de l'argent ?

Application de TCL

On suppose que le montant des dégâts par personne est une variable aléatoire. Le dégât X_i du mutualiste i suit la même loi que les autres et toutes ces variables sont supposées indépendantes (cela est pertinent pour des accidents de voiture mais est discutable si on parle de catastrophes naturelles). On sait que ces variables sont d'espérance 37 et d'écart-type 250. Notons k le montant de la cotisation.

Soit $n = 10^5$. La mutuelle perçoit chaque année nk mais doit reverser $S_n = \sum_{i=1}^n X_i$. On cherche donc la valeur de k pour laquelle $P(nk - S_n > 0) \geq 0,999$.

Le TCL (0.4) nous permet d'affirmer que $Y = \frac{S_n - 37n}{250\sqrt{n}} \sim \mathcal{N}(0, 1)$ à peu près. Ainsi

$$P(nk - S_n > 0) = P\left(Y < \frac{nk - 37n}{250\sqrt{n}}\right) \geq 0,999$$

On en déduit $\frac{nk - 37n}{250\sqrt{n}} \simeq 3,09$ et finalement $k \simeq 39,44$ euros.

Exemple : Théorème de Moivre-Laplace-approximation d'une loi binomiale par une loi normale

Un avion peut embarquer 300 passagers. Lors de l'embarquement, la probabilité qu'un passager ayant réservé ne se présente pas est de 0,1. On note n le nombre de réservations proposées par la compagnie et X la v.a. indiquant le nombre de passagers ayant réservé et se présentant à l'embarquement. Les comportements des passagers sont indépendants les uns des autres.

- 1 Quelle est la loi de X ?
- 2 Justifier que cette loi binomiale peut être approchée par une loi normale dont les paramètres sont à préciser.
- 3 Si la compagnie accepte 330 réservations, quelle est la probabilité, à 10^{-1} près, que tous les passagers ne puissent embarquer.
- 4 La compagnie souhaite limiter à 2,5% le risque de ne pas pouvoir embarquer tous les passagers ayant réservé. Déterminer, à l'aide d'une approximation par une loi normale, le nombre maximum n de places qu'elle peut proposer à la réservation.