

Chapitre 3

Les distributions statistiques

Les distributions statistiques découlent des lois de probabilité. Elles ne concernent que les **variables quantitatives** ou, éventuellement, les **variables semi-quantitatives**.

Pour trouver un modèle décrivant un ensemble de données, il est nécessaire de connaître parfaitement les lois statistiques les plus utilisées. Le choix de la loi est lié :

1. à la nature du phénomène étudié afin de choisir entre loi discrète et loi continue ;
2. à la forme de la distribution empirique (même s'est possible de la tester comme cela sera expliqué au chapitre concernant les statistiques inférentielles) ;
3. à la connaissance et à l'interprétation des principales caractéristiques de l'ensemble de données : espérance, médiane, variance, écart type, coefficients d'asymétrie et de dissymétrie, *etc.* ;
4. au nombre de paramètres des lois, une loi dépendant de plusieurs paramètres pouvant s'adapter plus facilement à une distribution.

Pour B. V. Gnedenko et A. N. Kolmogorov, « l'action collective régularise tout » [Gnedenko et Kolmogorov, 1954]. Cette citation illustre le fait que, dans une certaine situation bien précise, on finit par obtenir une loi de probabilité qui l'est tout autant. Boris Vladimirovitch Gnedenko (1912-1995)

3.1 Les distributions statistiques des variables discrètes

Les lois discrètes sont utilisées pour modéliser :

1. les résultats des jeux de hasard ;

2. les sondages d'opinion ;
3. les phénomènes biologiques ;
4. les processus aléatoires tels que les files d'attente, l'évolution de l'état de matériels.

La fonction de répartition est une fonction en escalier, constante sur tout intervalle $[x_i, x_{i+1}[$, admettant en chaque point x_i un saut égal à $p_{i+1} = \Pr(X = x_{i+1})$.

Les moments d'une variable aléatoire discrète sont la moyenne et la variance.

$$\mathbb{E}(X) = \sum_{i=1}^n x_i p_i \quad (3.1)$$

$$\mathbb{V}(X) = \left(\sum_{i=1}^n p_i x_i^2 \right) - [\mathbb{E}(X)]^2 \quad (3.2)$$

3.1.1 Loi de Dirac

Paul Dirac (1937-1984) La loi de Dirac au point a de \mathbb{R} est la loi de probabilité δ_a , définie sur \mathbb{R} par :

$$\begin{cases} \delta_a(x) = 1 & \text{si } x = a \\ \delta_a(x) = 0 & \text{si } x \neq a \end{cases} \quad (3.3)$$

Remarque. Il s'agit de la loi la plus simple, associée à un phénomène déterministe X dont le résultat de toute expérience est égale à a .

L'espérance vaut :

$$\mathbb{E}(X) = a \quad (3.4)$$

La variance vaut :

$$\mathbb{V}(X) = 0 \quad (3.5)$$

Soit A un ensemble de n nombres réels distincts a_i et n nombres réels p_i tels que :

$$\begin{cases} 0 \leq p_i \leq 1 \\ \sum_{i=1}^n p_i = 1 \end{cases} \quad (3.6)$$

La combinaison linéaire $\sum_{i=1}^n p_i \delta_{a_i}$ définit une loi de probabilité associée à une variable aléatoire discrète X telle que tout indice i :

$$\Pr(X = a_i) = p_i \quad (3.7)$$

3.1. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES DISCRÈTES 3

Remarque 1. Si, pour toutes les valeurs d'indice i , $p_i = \frac{1}{n}$, c'est-à-dire que les probabilités sont équiprobables, alors la loi de la variable aléatoire X est la **loi de probabilité uniforme sur A**.

Remarque 2. Pour toutes les variables discrètes, la densité de probabilité est composée d'une fonction de Dirac.

3.1.2 Loi uniforme discrète

La loi uniforme sur $[1, n]$ est la loi de probabilité d'une variable aléatoire X prenant chaque valeur de l'ensemble $\{1, 2, \dots, n\}$ avec la même probabilité :

$$\Pr(X = k) = \frac{1}{n} \text{ compris entre 1 et } n \quad (3.8)$$

Plus généralement, soit Ω un ensemble fini de cardinal n . La **loi de probabilité équadistribuée** (ou loi uniforme) sur Ω est la loi définie sur Ω par la probabilité :

$$\forall \omega \in \Omega, \Pr(A) = \frac{1}{n} \quad (3.9)$$

Pour toute partie finie A , on a :

$$\Pr(A) = \frac{\text{card}A}{\text{card}\Omega} \quad (3.10)$$

Tous les événements $(X = k)$ sont **équiprobables**.

L'espérance vaut :

$$\mathbb{E}(X) = \frac{1}{n} \sum_{k=1}^n k = \frac{n+1}{2} \quad (3.11)$$

La variance vaut :

$$\mathbb{V}(X) = \frac{1}{n} \sum_{k=1}^n k^2 - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12} \quad (3.12)$$

Les domaines d'utilisation sont :

- des jeux de pile ou face ;
- des jeux de dé non pipé ;
- des jeux de cartes ;
- des loteries ;
- des sondages.

3.1.3 Loi de Bernoulli et variable de Bernoulli

La loi de J. Bernoulli correspond à la **loi d'une variable aléatoire indicatrice**. Jacques Bernoulli (1654-1705) On suppose que l'on joue **une fois** à pile ou face. La règle du jeu est que l'on gagne 1 € lorsqu'on obtient « pile », et 0 € lorsqu'on obtient « face ». On note W le gain.

W est une **variable de Bernoulli**. Par convention, on définit un « succès » (ici pile) et un « échec » (ici face), soit une épreuve de J. Bernoulli. W peut ainsi prendre deux valeurs, soit 1 en cas de « succès », soit 0 en cas d'« échec ». De fait,

$$\Pr(W = 1) = p \text{ et } \Pr(W = 0) = 1 - p \quad (3.13)$$

On dit que loi de W est la loi de Bernoulli de paramètre p avec $0 \leq p \leq 1$. On la note $\mathcal{B}(p)$.

$$\Pr(W = 0) + \Pr(W = 1) = 1 \quad (3.14)$$

En appliquant la formule générale vue précédemment, l'espérance vaut :

$$\mathbb{E}(W) = \sum_{i=1}^2 p_i w_i = 1 \times p + 0 \times (1 - p) = p \quad (3.15)$$

N.B. $\mathbb{E}(W) = \mathbb{E}(W^2) = p$.

En appliquant la formule générale vue précédemment, la variance vaut :

$$\mathbb{V}(W) = \sum_{i=1}^2 [p_i (w_i - \mathbb{E}(W))^2] = p \times (1 - p)^2 + (1 - p) (0 - p)^2 = p(1 - p) \quad (3.16)$$

3.1.4 Loi binomiale ou loi des tirages avec remise

La loi binomiale concerne un **tirage non exhaustif avec remise** entre deux cas possibles, par exemple, une urne contenant des boules rouges et des boules noires. Quelle est la probabilité de piocher une boule rouge ? Il n'existe que deux choix possibles, donc il s'agit d'une variable de J. Bernoulli. Ainsi, on suppose que l'on tire n fois une boule avec les mêmes règles que la loi de J. Bernoulli. Les tirages successifs sont supposés être **indépendants**. On dit que l'on réalise n épreuves indépendantes de « succès » et d'« échecs ». De ce cadre, il s'agit d'une **loi des épreuves répétées d'un tirage avec remise** ; on appellera la variable aléatoire, **variable binomiale** K . Il s'agit d'une épreuve de J. Bernoulli, puisque les épreuves sont répétées et indépendantes à partir d'une même expérience.

Le gain cumulé K correspond à la somme des variables aléatoires de Bernoulli W_i de paramètre p avec $i = \{1, 2, \dots, n\}$.

$$K = W_1 + W_2 + \dots + W_n \quad (3.17)$$

3.1. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES DISCRÈTES 5

Les boules rouges sont l'épreuve à succès. On pose la variable aléatoire K correspondant le nombre de boules rouges, k les réalisations possibles de la variable aléatoire K , et p la proportion de boules rouges, et n le nombre de tirages total.

Pour tout $0 \leq k \leq n$, $K = k$ si et seulement si on observe k réalisations parmi les W_i prenant la valeur 1, tandis que l'on observe $n - k$ autres de valeur 0. Il existe ainsi $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ façons d'obtenir W_i valant 1. La probabilité dépend par conséquent du nombre de tirages total, c'est-à-dire de la taille de l'échantillon, de la population. On note p la probabilité de tirer une boule rouge qui dépend de ϖ , et $(1 - p)$ la probabilité de tirer une boule noire. Bien entendu, la probabilité de réalisation d'un événement doit être constante. En combinant les deux variables, la probabilité $\Pr(K = k)$ vaut :

$$\Pr(K = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (3.18)$$

avec $0 \leq k \leq n$. p^k représente la probabilité de n'avoir aucun succès, tandis que $(1 - p)^{n-k}$ représente la probabilité d'avoir **au moins** un succès. On note la loi binomiale $\mathcal{B}(n, p)$

La loi de la variable aléatoire K suit une loi binomiale β de paramètre n et p , notée $K \sim \mathcal{B}(n, p)$.

Si n et p sont constants, la fonction $P(*^k)$ est une distribution de probabilité discrète (Tab. 3.1)

k	0	1	2	...	n
$P(k)$	$(1 - p)^k$	$\binom{n}{p} (1 - p)^{n-1} p$	$\binom{n}{p} (1 - p)^{n-2} p^2$...	p^n

TABLE 3.1 – Loi de probabilité de la variable aléatoire binomiale avec n et p constants

La distribution binomiale correspond aux termes successifs du développement du binôme de I. Newton. On appelle également cette distribution, distribution de Bernoulli.

En appliquant la formule générale vue précédemment, l'espérance vaut :

$$\mathbb{E}(K = k) = \sum_{k=0}^n k \binom{n}{k} p^k (1 - p)^{n-k} = np \quad (3.19)$$

En appliquant la formule générale vue précédemment, la variance vaut :

$$\mathbb{V}(K = k) = \left[\sum_{k=0}^n k^2 \left[\binom{n}{k} p^k (1 - p)^{n-k} \right] \right] - n^2 p^2 = np(1 - p) \quad (3.20)$$

Le coefficient d'asymétrie vaut :

$$\gamma_1 = \frac{1 - 2p}{\sqrt{np(1-p)}} \quad (3.21)$$

Le coefficient d'aplatissement vaut :

$$\gamma_2 = \frac{1 - 6p}{np(1-p)} + 3 \quad (3.22)$$

Les domaines d'utilisation sont tous les phénomènes ne pouvant prendre que deux états s'excluant mutuellement, c'est-à-dire incompatibles :

- bonne pièce vs. pièce défectueuse ;
- fonctionnement vs. défaillance ;
- pile vs. face ;
- succès vs. échec

Propriété importante. La somme de n variables aléatoires binomiales indépendantes X et Y et de même paramètre p est une variable aléatoire binomiale.

$$\left. \begin{array}{l} X \sim \mathcal{B}(n_1, p) \\ Y \sim \mathcal{B}(n_2, p) \end{array} \right\} (X + Y) \sim \mathcal{B}(n_1 + n_2, p) \quad (3.23)$$

Pour finir, lorsque $n \rightarrow \infty$ et $p \rightarrow 0$, $\lambda = np$, on obtient la loi de Poisson.

3.1.5 Loi de Poisson

Siméon Denis Poisson (1781-1840) Une loi binomiale converge souvent vers une loi de S. D. Poisson lorsque le nombre de boules est grand ($n \rightarrow +\infty$), lorsque la proportion est faible ($p \rightarrow 0$), et lorsque le produit $\lambda = np$ est fini et positif avec $\lambda \in \mathbb{N}$.

$$p = \frac{\lambda}{n} \quad (3.24)$$

Pour tout $k \geq 0$, la probabilité $\Pr(X = k)$ vaut :

$$\Pr(X = k) = \left(1 - \frac{\lambda}{n}\right)^{-k} \frac{(n - k + 1) \times \dots \times n}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \frac{\lambda^k}{k!} \quad (3.25)$$

Lorsque $n \rightarrow +\infty$, les deux premiers facteurs tendent vers 1, tandis que le troisième converge vers $e^{-\lambda}$. Dans ce cas, la variable binomiale disparaît.

$$\Pr(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad (3.26)$$

3.1. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES DISCRÈTES 7

avec $k \in \mathbb{N}$ et λ une **constante quelconque**. La variable aléatoire X suit une loi de S. D. Poisson de paramètre λ , notée $X \sim \mathcal{P}(\lambda)$. Elle sert d'approximation de la loi binomiale $\mathcal{B}(n, p)$. La distribution de S. D. Poisson est **infiniment dénombrable**.

N.B. On a $\sum_{k=1}^n \frac{\lambda^k}{k} = e^\lambda$ ce qui implique que $\sum_{k=1}^n \Pr(X = k) = 1$.

La moyenne et la variance sont confondues.

$$\mathbb{E}(X) = \mathbb{V}(X) = \lambda \quad (3.27)$$

Propriété importante. La somme de deux variables aléatoires de Poisson, indépendantes et de paramètres $X \sim \mathcal{P}(\lambda_1)$ et $Y \sim \mathcal{P}(\lambda_2)$, est une variable aléatoire de Poisson de paramètre $(X + Y) \sim \mathcal{P}(\lambda_1 + \lambda_2)$.

La loi de S. D. Poisson correspond à la **loi des événements rares** (ou la loi des petites probabilités). Le domaine d'utilisation est la réalisation d'événements peu probables dans une succession d'épreuves très nombreuses (au moins 50). Elle permet d'obtenir une approximation de la loi de distribution binomiale pour k petit, pourvu que $\lambda = np$, et que p soit petit.

3.1.6 Loi géométrique

On répète la **même épreuve** de J. Bernoulli, les répétitions étant **indépendantes** les unes des autres jusqu'à l'**obtention du premier succès**.

Au jeu de pile ou face, on fixe une probabilité p de gagner à chaque coup. On définit une variable aléatoire $X \in \mathbb{N}^*$ comme étant le nombre d'essais qu'il faut jouer pour obtenir un « pile ». Sa loi est une loi géométrique de paramètre p , notée $X \sim \mathcal{G}(p)$, dont la probabilité $\Pr(X = k)$ vaut :

$$\Pr(X = k) = p(1 - p)^{k-1} \quad (3.28)$$

avec $k \in \mathbb{N}^*$. La loi géométrique de la loi de la variable X correspond à la « loi du nombre d'essais nécessaires » pour qu'un événement de probabilité p apparaisse pour la première fois, les hypothèses étant les mêmes que pour la loi binomiale, en particulier, la probabilité p reste constante au cours des essais. La loi précédente se lit : « il y a eu $k - 1$ échecs avant d'obtenir le k -ième essai ».

On note R_i l'événement « succès à la i -ième épreuve ». L'événement $\{X = k\}$ peut s'écrire :

$$\{X = k\} = \left(\bigcap_{i=1}^{k-1} (R_i)^c\right) \cap R_k \quad (3.29)$$

Par indépendance des épreuves, on obtient facilement l'expression de $\Pr(X = k)$.

$$\sum_{k=1}^{+\infty} \Pr(X = k) = p \sum_{k=1}^{+\infty} (1 - p)^{k-1} = p \sum_{l=0}^{+\infty} (1 - p)^l = p \frac{1}{1 - (1 - p)} = 1 \quad (3.30)$$

L'espérance de la loi géométrique vaut :

$$\mathbb{E}(X) = \frac{1}{p} \quad (3.31)$$

La variance de la loi géométrique vaut :

$$\mathbb{V}(X) = \frac{1-p}{p^2} \quad (3.32)$$

Le coefficient d'asymétrie vaut :

$$\gamma_1 = \frac{2-p}{\sqrt{1-p}} \quad (3.33)$$

Le coefficient d'aplatissement vaut :

$$\gamma_2 = \frac{(p-3)^2 - 3}{1-p} + 3 \quad (3.34)$$

3.1.7 Loi binomiale négative

On généralise l'exemple précédent et dans le même jeu de pile ou face. On désigne par la variable aléatoire X cette fois le nombre de coups qu'il faut jouer pour obtenir une fois pile. X prend alors ses valeurs dans les entiers plus grands ou égaux à m , sa loi correspond à la loi binomiale négative de paramètres m et p .

$$\forall k \geq m, \Pr(X = k) = C_{k-1}^{m-1} p^m (1-p)^{k-m} \quad (3.35)$$

En effet, on note la variable aléatoire Y_i si et seulement si le i -ième coup donne pile :

$$\{X = k\} = \{Y_1 + \dots + Y_{k-1} = m-1\} \cap \{Y_m = 1\} \quad (3.36)$$

L'espérance vaut :

$$\mathbb{E}(X) = m \frac{1-p}{p} \quad (3.37)$$

La variance vaut :

$$\mathbb{V}(X) = m \frac{1-p}{p} \frac{1}{p} \quad (3.38)$$

Le coefficient d'asymétrie vaut :

$$\gamma_1 = \frac{\frac{1-p}{p} + \frac{1}{p}}{\sqrt{m \frac{1-p}{p} \frac{1}{p}}} \quad (3.39)$$

Le coefficient d'aplatissement vaut :

$$\gamma_2 = \frac{1 + 6 \frac{1-p}{p} \frac{1}{p}}{m \frac{1-p}{p} \frac{1}{p}} + 3 \quad (3.40)$$

3.1.8 Loi hypergéométrique ou loi du tirage exhaustif

La loi hypergéométrique concerne un **tirage sans remise**, c'est-à-dire pour lequel les proportions changent à chaque tirage. Soit une urne contenant N boules de deux couleurs différentes. On note m le nombre de boules blanches et $N - m$ le nombre de boules noires. On tire n boules. Avec remise, la variable aléatoire X représentant les boules blanches suit une loi binomiale de type $X \sim \beta(n, \frac{m}{N})$. Sans remise, la loi de X est une loi hypergéométrique de paramètre N , m et n , notée $H(N, m, n)$. La probabilité $\Pr(X = k)$ vaut :

$$\Pr(X = k) = \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} \quad (3.41)$$

avec $m = Np$. Il s'agit d'un tirage équiprobable sans remise (ou tirage exhaustif). Les valeurs extrêmes sont : $\min k = \max\{0, n - N(1 - p)\}$ et $\max k = \min\{N, Np\}$. Le quotient $\frac{n}{N}$ est appelé **taux de sondage**.

L'espérance vaut :

$$\mathbb{E}(X) = \frac{nm}{N} \quad (3.42)$$

Soit X_1 la variable aléatoire du premier individu, on a alors :

$$\left. \begin{array}{l} \Pr(X_1 = 1) = p \\ \Pr(X_1 = 0) = 1 - p \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \mathbb{E}(X_1) = p \\ \mathbb{V}(X_1) = p(1 - p) \end{array} \right. \quad (3.43)$$

Soit X_2 la variable aléatoire du deuxième individu, on a alors :

$$\left. \begin{array}{l} \Pr(X_2 = 1) = p \\ \Pr(X_2 = 0) = 1 - p \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \mathbb{E}(X_2) = p \\ \mathbb{V}(X_2) = p(1 - p) \end{array} \right. \quad (3.44)$$

Pour X_i , $\mathbb{E}(X) = \sum_{i=1}^n X_i = np$

La variance vaut :

$$\mathbb{V}(X) = \frac{nm}{N} \left(1 - \frac{n}{N}\right) \left(\frac{N-n}{N-1}\right) \quad (3.45)$$

ou

$$\mathbb{V}(X) = \frac{N-n}{N-1} np(1-p) \quad (3.46)$$

$\frac{N-n}{N-1}$ est appelé le **facteur d'exhaustivité**. Les variables aléatoires X_i sont des variables aléatoires de Bernoulli non indépendantes. Le facteur d'exhaustivité tend vers 1 si N est grand devant n , et la variance de la variable X tend alors vers la variance d'une variable suivant la loi binomiale $\beta(n, p)$.

Les contrôles de qualité dans lesquels on retire de la population étudiée les éléments défectueux, sont l'un des domaines d'utilisation de la loi hypergéométrique

3.1.9 Loi multinominale

Il s'agit d'une généralisation de la loi binomiale. Soit un univers $\Omega = \{1, 2, \dots, n\}$ muni d'une probabilité. On tire n fois de suite un élément de Ω avec la probabilité $p_i = 1, 2, \dots, n$. Quelle est la probabilité d'obtenir $\sum_{i=1}^n = 1$ le nombre 1, k_1 fois le nombre 2, k_2 fois, sur une suite d'un **tirage équiprobable et indépendant** de n éléments ? Il s'agit d'un tirage avec remise.

Si l'univers d'une certaine expérience soit partitionné, c'est-à-dire, soient s événements A_1, A_2, \dots, A_s , s'excluant mutuellement, ayant pour probabilités respectives p_1, p_2, \dots, p_s , ce qui signifie que $p_1 + p_2 + \dots + p_s = 1$, on a alors le théorème suivant.

Théorème. Pour n épreuves répétée la probabilité pour que A_1 se produise k_1 fois, que A_2 se produise k_2 fois, \dots , et que A_s se produise k_s fois est égale à :

$$\frac{n!}{k_1! k_2! \dots k_s!} p_1^{k_1} p_2^{k_2} \dots p_s^{k_s} \quad (3.47)$$

avec $k_1 + k_2 + \dots + k_s = n$.

Ces nombres forment ce que l'on appelle la **loi de distribution multinomiale**, car ils sont précisément les termes du développement de $(p_1 + p_2 + \dots + p_s)^n$.

N.B. Si $s = 2$, on retrouve la loi de distribution binomiale.

Si l'on prend un ensemble d'événements ayant plusieurs issues, alors les différentes combinaisons de suites que l'on peut obtenir en prenant k_1 éléments choisis parmi n est :

$$\binom{n}{p} = \frac{n!}{p! (n-p)!} \quad (3.48)$$

Par conséquent, il y a :

$$\binom{n}{k_1} = \frac{n!}{k_1! (n-k_1)!} \quad (3.49)$$

façons différentes d'obtenir k_1 fois un premier événement. De même avec k_2 ,

$$\binom{n}{k_2} = \frac{n!}{k_2! (n-k_2)!} \quad (3.50)$$

façons différentes d'obtenir k_1 fois un deuxième événement.

Par récurrence, on a $\frac{n!}{\prod_{i=1}^n k_i!}$ combinaisons de sortir k_i fois quelques événements dans l'ordre donné.

La probabilité p recherchée vaut :

$$p = \frac{n!}{\prod_{i=1}^n k_i!} \prod_{i=1}^n p_i^{k_i} = \Pr(k_1, k_2, \dots, k_n) \quad (3.51)$$

3.1. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES DISCRÈTES 11

$\frac{n!}{\prod_{i=1}^n k_i!}$ correspond au nombre de partitions d'un échantillon de taille n en sous populations d'effectifs k_i .

Soit X_i la variable aléatoire représentant le nombre d'individus de type i dans l'échantillon. Par définition, le vecteur $X = \{X_1, X_2, \dots, X_i\}$ est un vecteur aléatoire suivant une loi multinomiale de paramètres (n, p_1, \dots, p_k) , notée $M(n, p_1, \dots, p_k)$.

Propriété 1. La loi marginale de X_i est une loi binomiale $\beta(n, p)$. Un élément tiré est :

- soit du type i avec la probabilité p_i ;
- soit de n'importe quel autre type avec la probabilité $(1 - p_i)$.

Le nombre d'individus du type i dans l'échantillon suit alors la loi binomiale $\beta(n, p_i)$:

- $\mathbb{E}(X_i) = np_i$;
- $\mathbb{V}(X_i) = np_i(1 - p_i)$.

Propriété 2. Soient deux variables X_i et X_k deux variables dépendantes, Le couple (X_i, X_k) suit une loi multinomiale de dimension 3. Un élément tiré est :

- soit du type i avec une probabilité p_i ;
- soit du type k avec une probabilité p_k ;
- soit de n'importe quel autre type avec une probabilité $1 - p_i - p_k$.

Propriété 3. Soient deux variables X_i et X_k deux variables dépendantes, $\mathbb{E}(X_i X_k) = n(n-1)p_i p_k$.

Propriété 4. Soient deux variables X_i et X_k deux variables dépendantes, $\text{cov}(X_i X_k) = \mathbb{E}(X_i X_k) - \mathbb{E}(X_i) \mathbb{E}(X_k) = -np_i p_k$.

Exercice type On jette 8 fois un dé bien équilibré. La probabilité p d'obtenir 5 et 6 deux fois, chacun des autres résultats une fois est :

$$p = \frac{8!}{2!2!1!1!1!1!} \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) = \frac{35}{5832} \approx 0,006 \quad (3.52)$$

avec $n = 8$, $k_1 = 2$, $k_2 = 2$, $k_3 = 1$, $k_4 = 1$, $k_5 = 1$ et $k_6 = 1$.

3.1.10 Loi de Pascal

La loi de Pascal de la variable Z recherche la « loi du nombre d'essais nécessaires » pour obtenir exactement k fois un événement de probabilité p , les hypothèses étant les mêmes que celles de la loi binomiale. La probabilité p est constante au cours des essais.

$$\Pr(Z = z) = p C_{z-1}^{k-1} p^{k-1} (1-p)^{z-k} \quad (3.53)$$

$z \geq k$ et $z \in \mathbb{N}^*$. Cette loi se lit « on a obtenu un succès à l'essai numéro z de probabilité p et $k-1$ succès au cours de $z-1$ essais précédents ».

L'espérance vaut :

$$\mathbb{E}(X) = \frac{k}{p} \quad (3.54)$$

La variance vaut :

$$\mathbb{V}(X) = \frac{k(1-p)}{p^2} \quad (3.55)$$

Le coefficient d'asymétrie vaut :

$$\gamma_1 = \frac{2-p}{\sqrt{k(1-p)}} \quad (3.56)$$

Le coefficient d'aplatissement vaut :

$$\gamma_2 = \frac{p^2 + 6(1-p)}{k(1-p)} + 3 \quad (3.57)$$

3.1.11 Lois de Zipf et de Zipf-Mandelbrot

George Kingsley Zipf (1902-1950) Dans les années 1930, G. K. Zipf observa que la fréquence d'utilisation d'un mot dans un texte volumineux était inversement proportionnelle à son rang. Par exemple, si le mot le plus fréquent apparaît 1 000 fois dans un texte (rang 1), les mots se classant aux deuxième et troisième rangs en ce qui concerne leur fréquence apparaîtront respectivement $\frac{1000}{2} = 500$ fois et $\frac{1000}{3} = 333$ fois. Dans ce cadre, on suppose que les événements ne sont pas équiprobables.

Selon G. K. Zipf, si n correspond au n ième mot le plus fréquent dans un texte, il doit alors avoir une fréquence $f(n) = \frac{K}{n}$ avec K une constante indépendante de n .

Benoît Mandelbrot (1924-2010) B. Mandelbrot a proposé une généralisation de la loi de G. K. Zipf. La fréquence du mot ayant le n ième rang vaut dans ce cas $f(n) = \frac{K}{(a+bn)^c}$ avec a, b, c et K des constantes.

Étant donné la distribution des fréquences, il est parfois difficile de dire s'il s'agit d'une distribution de Zipf, car d'autres distributions y ressemblent beaucoup comme la loi log-normale. Il existe une astuce pour vérifier s'il s'agit bel et bien d'une distribution de Zipf. Lorsque n est grand, on peut calculer la somme des fré-

Leonhard Euler (1707-1783) quences si elle est proche de la constante d'Euler-Mascheroni ($\gamma = 0,5772156649^1$).

Lorenzo Mascheroni (1750-1800)

$$\sum_{k=1}^n \frac{1}{k} \approx \gamma + \ln n + \frac{1}{2}n^{-1} - \frac{1}{12}n^{-2} + \frac{1}{120}n^{-4} + \dots = 1 \quad (3.58)$$

Cette somme est l'équivalent d'une somme de probabilité, donc elle vaut un. À partir de là, il est possible de contraindre la valeur des constantes.

En géographie, cette loi se rencontre dans les lois rang-taille confrontant, au sein d'un territoire, le nombre d'habitants d'une ville avec son rang.

3.2 Les distributions statistiques des variables continues

La fonction de répartition de la variable aléatoire X est définie par :

$$F(a) = \Pr(X < a) = \int_{-\infty}^a f(x) dx \quad (3.59)$$

Pour toutes les valeurs a et b appartenant à \mathbb{R} , on a la relation :

$$\Pr(a \leq X \leq b) = F(b) - F(a) \quad (3.60)$$

On en déduit :

$$\begin{cases} \Pr(X = x) = 0 \\ \Pr(x \leq X \leq x + dx) = f(x) dx \end{cases} \quad (3.61)$$

La densité de probabilité ρ d'une variable continue, c'est-à-dire sa fonction de répartition, vaut :

$$\rho(x) = \lim_{\Delta x \rightarrow 0} \frac{\Pr(x < X < x + \Delta x)}{\Delta x} \quad (3.62)$$

Les moments de la variable aléatoire continue X sont :

- l'espérance mathématique valant : $\mathbb{E}(X) = \int x f(x) dx$;
- la variance valant : $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \left(\int x^2 f(x) dx \right) - \left(\int x f(x) dx \right)^2$.

Il est important de comprendre que l'espérance et la variance n'existent que si les intégrales sont définies.

1. La constante d'Euler-Mascheroni est définie de la manière suivante : $\gamma = \lim_{n \rightarrow +\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n} - \ln n \right)$ ou $\gamma = \lim_{n \rightarrow +\infty} \left(\left(\sum_{k=1}^n \frac{1}{k} \right) - \ln n \right)$.

3.2.1 Loi de Poisson continue

Il existe une très grande quantité de fonctions mathématiques qui décrivent le comportement d'observations d'**événements aléatoires, indépendants et identiquement distribués**. Pour certains événements forts rares, la probabilité \Pr est très faible, et tend vers zéro. Toutefois, la valeur moyenne np tend vers une valeur fixe lorsque n tend vers l'infini. Soit une distribution binomiale de moyenne $\mu = np$ que l'on suppose finie lorsque n tend vers l'infini.

La probabilité de k réussites lors de n épreuves vaut :

$$\Pr(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (3.63)$$

En posant $p = \frac{m}{n}$, cette expression peut s'écrire :

$$\Pr(k) = \binom{n}{k} \left(\frac{m}{n}\right)^k \left(1 - \frac{m}{n}\right)^{n-k} \quad (3.64)$$

En regroupant les termes, on peut mettre la valeur de $\Pr(k)$ sous la forme :

$$\Pr(k) = \underbrace{\frac{m^k}{k!}}_{1^{\text{er}} \text{ facteur}} \underbrace{\left(1 - \frac{m}{n}\right)^n}_{2^{\text{e}} \text{ facteur}} \underbrace{\frac{n!}{(n-k)! n^k \left(1 - \frac{m}{n}\right)^k}}_{3^{\text{e}} \text{ facteur}} \quad (3.65)$$

Lorsque n tend vers l'infini, le deuxième facteur du produit a pour limite $e^{-\mu}$. Pour le troisième facteur, puisque l'on s'intéresse aux petites valeurs de k , c'est-à-dire la probabilité de réussite est très faible), sa limite pour n tendant vers l'infini vaut 1. On obtient la **loi de distribution de Poisson**.

$$\Pr(k) = \frac{\mu^k}{k!} e^{-\mu} \quad (3.66)$$

Cette distribution est fondamentale, car elle décrit beaucoup de processus dont la probabilité est petite et constante. Elle est souvent utilisée dans la « *queuing theory* » (temps d'attente), test d'acceptabilité et fiabilité, et contrôles statistiques de qualité. La distribution de Poisson donne la probabilité de trouver un nombre donné d'événements en fonction d'une variable donnée, si les événements sont indépendants, et ont un taux de répétition (correspondant à une probabilité) constant (c'est-à-dire équiprobable).

L'espérance de la fonction de distribution de Poisson est :

$$\mu = \mathbb{E}(k) = \sum_{k=0}^{+\infty} k \Pr(k) = \sum_{k=0}^{+\infty} k \frac{\mu^k}{k!} e^{-\mu} = \mu e^{-\mu} \sum_{k=0}^{+\infty} \frac{\mu^{k-1}}{(k-1)!} = \mu e^{-\mu} e^{\mu} = \mu \quad (3.67)$$

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 15

La variance de la fonction de distribution de Poisson est donnée par :

$$\begin{aligned}\mathbb{V}(k) &= \sum_{k=0}^{+\infty} (k - \mu)^2 \Pr(k) \\ \mathbb{V}(k) &= \sum_{k=0}^{+\infty} (k - \mu)^2 \frac{\mu^k}{k!} e^{-\mu} \\ \mathbb{V}(k) &= \sum_{k=0}^{+\infty} \left[k^2 \frac{\mu^k}{k!} e^{-\mu} - 2k\mu \frac{\mu^k}{k!} e^{-\mu} + \mu^2 \frac{\mu^k}{k!} e^{-\mu} \right]\end{aligned}\quad (3.68)$$

après quelques calculs, on trouve :

$$\mathbb{V}(k) = \mu \quad (3.69)$$

Pour la distribution de Poisson, on a le cas particulier dans lequel :

$$\mathbb{E}(k) = \mathbb{V}(k) = \sigma^2 \quad (3.70)$$

donc :

$$\sigma = \pm \sqrt{\mu} \quad (3.71)$$

Les lois de distribution statistiques sont établies en supposant la réalisation d'un nombre infini de mesures. Il est évident que l'on ne peut en effectuer qu'un nombre fini N .

En théorie :

$$\begin{aligned}\Pr(k) &= \frac{\mu^k}{k!} e^{-\mu} \\ \mathbb{E}(k) &= \mu \\ \mathbb{V}(k) &= \sigma^2 \\ \Pr_{\text{théorique}}(k) &= \frac{\mu^k}{k!} \\ \mu_{\text{théorique}} &= \sum_{k=0}^{+\infty} k \Pr_{\text{théorique}}(k) = \mu \\ \sigma_{\text{théorique}}^2 &= \sum_{k=0}^{+\infty} (k - \mu_{\text{théorique}})^2 \Pr_{\text{théorique}}(k) = \mu_{\text{théorique}}\end{aligned}\quad (3.72)$$

En pratique :

$$\begin{aligned}\Pr_{\text{expérimentale}}(k) &= \frac{N_k}{N} \\ \mu_{\text{expérimentale}} &= \sum_{k=0}^{+\infty} k \Pr_{\text{expérimentale}}(k) = \frac{1}{N} \sum_{k=0}^{+\infty} k N_k \\ \sigma_{\text{expérimental}}^2 &= \sum_{k=0}^{+\infty} (k - \mu_{\text{expérimentale}})^2 \Pr_{\text{expérimentale}}(k) \approx \mu_{\text{expérimentale}} \approx \mu_{\text{théorique}}\end{aligned}\quad (3.73)$$

Dit autrement, à chaque fois qu'une mesure est opérée, il s'ensuit une incertitude sur le résultat unique obtenu. On suppose avoir compté (c'est-à-dire observé) N événements en fonction d'une certaine variable. L'écart entre N et la moyenne μ sera en moyenne σ . Il est par conséquent naturel d'écrire :

$$N = \mu \pm \sigma \quad (3.74)$$

De plus $\sigma_{\text{expérimental}} \approx \sqrt{N} \approx \sqrt{\mu_{\text{expérimentale}}} \approx \sqrt{\mu}$, on peut de fait écrire l'observation de N événements comme étant :

$$N = \mu \pm \sqrt{N} \quad (3.75)$$

3.2.2 Loi normale

La loi normale concerne les variables continues. Elle s'appelle ainsi, car sa paternité est difficile à établir entre Pierre-Simon de Laplace qui a découvert le théorème centrale limite et Carl Friedrich Gauss qui en a fait sa formalisation. Parfois, on l'appelle de fait loi de Gauss-Laplace, ou plus simplement la loi de Gauss, ce dernier ayant à son compte les travaux les plus récents sur cette loi.

Pierre-Simon de Laplace (1746-1827)
 Carl Friedrich Gauss (1777-1855)
 Abraham de Moivre (1667-1774)

Il est compliqué de ne pas commencer les distributions statistiques par la loi normale. Pourtant, ce n'est pas la distribution statistique la plus simple à comprendre. Elle n'en demeure pas moins celle qui est la plus fréquente, celle qui donne notamment un sens à la moyenne arithmétique. De plus, la loi normale est bien souvent la distribution limite à l'infini d'autres distributions statistiques. Tel est le cas de la loi binomiale, comme cela fut démontré par A. de Moivre.

Si on fait tendre le nombre n d'épreuves d'une loi binomiale vers l'infini, et si Pr est fixé au départ, la moyenne $\mu = np$ tend également vers l'infini et l'écart type $np(1-p)$ tend vers l'infini.

Remarque. Si Pr varie et tend vers zéro tout en laissant fixe la moyenne μ , alors il s'agit d'un processus de Poisson.

Étape 1. Comment varie $\text{Pr}(k)$ en fonction de k ?

Pour y répondre, il suffit d'étudier $\text{Pr}(k+1) - \text{Pr}(k)$.

$$\begin{aligned} \text{Pr}(k+1) - \text{Pr}(k) &= \binom{n}{k+1} p^{k+1} (1-p)^{n-k-1} - \binom{n}{k} p^k (1-p)^{n-k} \\ \text{Pr}(k+1) - \text{Pr}(k) &= \text{Pr}(k) \frac{np-k-(1-p)}{(k+1)(1-p)} \end{aligned} \quad (3.76)$$

$\text{Pr}(k)$ est une fonction croissante de k tant que $np - k - (1-p)$ est positif.

Remarque 1. Si $1-p < 1$, alors la valeur k voisine de la moyenne $\mu = np$ constitue le maximum de $\text{Pr}(k)$.

Remarque 2. La différence $\text{Pr}(k+1) - \text{Pr}(k)$ est le taux d'accroissement de la fonction $\text{Pr}(k)$. De fait, on peut écrire :

$$\frac{\Delta \text{Pr}(k)}{\Delta k} = \frac{\text{Pr}(k+1) - \text{Pr}(k)}{(k+1) - k} \quad (3.77)$$

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 17

Si n devient grand, il est clair que k l'est également. Dès lors, pour autant que Pr soit fixé et que l'on s'intéresse aux valeurs de la fonction k proches de la moyenne $\mu = np$ dès lors les variations de k peuvent s'écrire :

$$\Delta k = k - np \quad (3.78)$$

la valeur de la nouvelle variable tend vers zéro lorsque l'on effectue un passage à la limite pour n dont la moyenne est nulle (c'est-à-dire avec des variations négligeables) et dont son écart type est unitaire. Ainsi,

$$x = \frac{k - np}{\sqrt{np(1-p)}} \quad (3.79)$$

Étape 2. Étude des variations

On appelle $F(x)$ l'expression de $\text{Pr}(k)$, calculée en fonction de la nouvelle variable. Le taux d'accroissement de $F(x)$ vaut alors :

$$\frac{\Delta F(x)}{\Delta x} = \frac{\Delta \text{Pr}(k) \Delta k}{\Delta k \Delta x} = \frac{\Delta \text{Pr}(k)}{\Delta k} \frac{k - np}{\frac{k - np}{\sqrt{np(1-p)}}} = \frac{\Delta \text{Pr}(k)}{\Delta k} \sqrt{np(1-p)} \quad (3.80)$$

En appliquant le résultat précédemment trouvé, on obtient :

$$\frac{\Delta F(x)}{\Delta x} = - \frac{[(k - np) + (1 - p)] np}{(k + 1) \sqrt{np(1-p)}} \text{Pr}(k) \quad (3.81)$$

Comme on a posé $F(x) = \text{Pr}(k)$ et $x = \frac{k - np}{\sqrt{np(1-p)}}$, l'expression devient :

$$\begin{aligned} \frac{\Delta F(x)}{\Delta x} &= - \frac{[(k - np) + (1 - p)] np}{(k + 1) \sqrt{np(1-p)}} F(k) \\ \frac{\Delta F(x)}{\Delta x} &= -x F(x) - \frac{np(1-p)}{(k + 1) \sqrt{np(1-p)}} F(x) \end{aligned} \quad (3.82)$$

Étape 3. Passage à la limite

Après un passage à la limite pour n tendant vers l'infini, on obtient :

$$\frac{dF(x)}{dx} = -x F(x) \quad (3.83)$$

Cette équation peut encore s'écrire :

$$\frac{dF(x)}{F(x)} = -x dx \quad (3.84)$$

En intégrant les deux membres de cette égalité, on obtient :

$$\ln (F(x)) = -\frac{1}{2}x^2 + K \quad (3.85)$$

dont une des solutions est la fonction suivante :

$$F(x) = \lambda e^{-\frac{1}{2}x^2} \quad (3.86)$$

La constante λ est déterminée par la condition que $\int_{-\infty}^{+\infty} F(X) dx$ qui représente la somme de toutes les probabilités, vaut 1. On peut alors montrer que :

$$\lambda = \frac{1}{\sqrt{2\pi}} \quad (3.87)$$

On obtient ainsi la loi normale centrée réduite :

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (3.88)$$

En revenant aux variables non normées, on obtient :

$$\Pr(k, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{k-\mu}{\sigma}\right)^2} \quad (3.89)$$

Cette loi régit sous des conditions très générales, et souvent rencontrées, beaucoup de phénomènes aléatoires. Elle est symétrique par rapport à la moyenne μ .

Étape 4. Quelle est l'espérance de la loi normale ?

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad (3.90)$$

On pose $u = \frac{x-\mu}{\sigma}$.

$$\begin{aligned} \mathbb{E}(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma u + \mu) e^{-\frac{1}{2}u^2} \sigma du \\ \mathbb{E}(X) &= \sigma \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sigma u e^{-\frac{1}{2}u^2} du + \sigma \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mu e^{-\frac{1}{2}u^2} du \\ \mathbb{E}(X) &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u e^{-\frac{1}{2}u^2} du + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}u^2} du \end{aligned} \quad (3.91)$$

On calcule la première intégrale directement :

$$J = \int_{-\infty}^{+\infty} u e^{-\frac{1}{2}u^2} du = \left[-e^{-\frac{u^2}{2}} \right]_{-\infty}^{+\infty} = \left[e^{-\frac{u^2}{2}} \right]_{-\infty}^{+\infty} = 0 \quad (3.92)$$

donc l'équation à résoudre se réduit à :

$$\mathbb{E}(X) = \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}u^2} du = \frac{\mu}{\sqrt{2\pi}} \sqrt{2\pi} = \mu \quad (3.93)$$

Étape 5. Calcul de la variance et de l'écart type

$$\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \quad (3.94)$$

On calcule $[\mathbb{E}(X)]^2$ directement :

$$[\mathbb{E}(X)]^2 = \mu^2 \quad (3.95)$$

Il ne reste plus qu'à calculer $\mathbb{E}(X^2)$.

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad (3.96)$$

On change de variable en posant $y = \frac{x-\mu}{\sqrt{2}\sigma}$.

$$\begin{aligned} \mathbb{E}(X^2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (y\sqrt{2}\sigma + \mu)^2 e^{y^2} dy \\ \mathbb{E}(X^2) &= \frac{2}{\sqrt{\pi}} \sigma^2 \int_{-\infty}^{+\infty} y^2 e^{-y^2} dy + 2\mu\sigma \sqrt{\frac{2}{\pi}} \int_{-\infty}^{+\infty} y e^{-y^2} dy + \frac{\mu^2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-y^2} dy \end{aligned} \quad (3.97)$$

On sait que :

$$\begin{cases} \int_{-\infty}^{+\infty} y e^{-y^2} dy = 0 \\ \int_{-\infty}^{+\infty} e^{-y^2} dy = \sqrt{\pi} \end{cases} \quad (3.98)$$

Il ne reste plus qu'à calculer la première intégrale $\int_{-\infty}^{+\infty} y^2 e^{-y^2} dy$ par une intégration par parties², justifiée et facilitée par les exponentielles de la formule.

$$\int_{-\infty}^{+\infty} y^2 e^{-y^2} dy = \int_{-\infty}^{+\infty} y (y e^{-y^2}) dy = \left[-y \frac{e^{-y^2}}{2} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{2} dy = \frac{\sqrt{\pi}}{2} \quad (3.99)$$

Il est désormais possible de calculer $\mathbb{E}(X^2)$.

$$\mathbb{E}(X^2) = \frac{2}{\sqrt{\pi}} \sigma^2 \frac{\sqrt{\pi}}{2} + 2\mu\sigma \sqrt{\frac{2}{\pi}} \times 0 + \frac{\mu^2}{\sqrt{\pi}} \sqrt{\pi} = \sigma^2 + \mu^2 \quad (3.100)$$

Il est alors possible de calculer la variance :

$$\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \quad (3.101)$$

2. Formule de l'intégration par parties : $\int_a^b f(t) g'(t) dt = [f(t) g(t)]_a^b - \int_a^b f'(t) g(t) dt$

Étape 6. Les coefficients d'asymétrie et d'aplatissement

Le coefficient d'asymétrie $\gamma_1 = 0$. C'est de fait une distribution symétrique.

Le coefficient d'aplatissement $\gamma_2 = 3$.

Caractéristiques de la loi normale

La loi normale caractérise une variable aléatoire X qui subit un grand nombre de facteurs non tous identifiés, indépendant, contribuant très faiblement à faire varier cette grandeur, et n'intervenant pas en même temps.

Les valeurs de la médiane, du mode et de la moyenne sont confondues.

Il existe une infinité de lois normales, notée $X \sim N(\mu, \sigma)^3$. Plus l'écart type σ est petit, plus la distribution est resserrée par rapport à la moyenne μ .

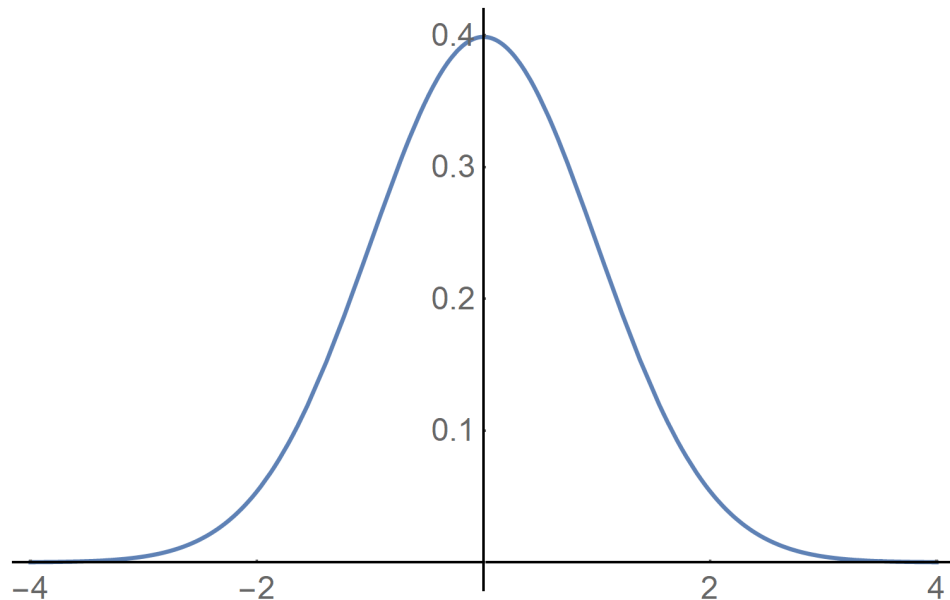


FIGURE 3.1 – Loi normale centrée réduite

La loi normale (ici centrée et réduite) est représentée sous la forme d'une cloche (Fig .3.1) que l'on appelle l'**ogive de Galton**.

Francis Galton
(1822-1911)

Sur une loi normale (ici centrée et réduite) (Fig .3.1), la variable suit une répartition en fonction de sa moyenne μ et de son écart type σ .

- 68,3 % des valeurs de la distribution se trouve entre $\mu - \sigma$ et $\mu + \sigma$.
- 95,5 % des valeurs de la distribution se trouve entre $\mu - 2\sigma$ et $\mu + 2\sigma$.
- 99,7 % des valeurs de la distribution se trouve entre $\mu - 3\sigma$ et $\mu + 3\sigma$.

3. Cela se lit : X suit une loi normale de paramètres μ et σ .

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 21

La largeur de l'intervalle a une très grande importance dans l'interprétation des incertitudes d'une mesure. Par exemple, $\mu \pm \sigma$ signifie que la valeur moyenne possède environ 68,3 % de chance de se trouver entre $\mu - \sigma$ et $\mu + \sigma$. De même, $\mu \pm 2\sigma$ signifie que la valeur moyenne possède environ 95,5 % de chance de se trouver entre $\mu - 2\sigma$ et $\mu + 2\sigma$. Et ainsi de suite... Toutefois, il est important de noter, notamment pour construire un intervalle de confiance, que :

- $\Pr [(X < \mu - 1,65\sigma) \text{ ou } (X > \mu + 1,65\sigma)] = 10 \%$
- $\Pr [(X < \mu - 1,96\sigma) \text{ ou } (X > \mu + 1,96\sigma)] = 5 \%$
- $\Pr (\mu - \sigma < X < \mu + \sigma) = 68,3 \%$
- $\Pr [(X < \mu - \sigma) \text{ ou } (X > \mu + \sigma)] = 33,7 \%$
- $\Pr (\mu - 2,58\sigma < X < \mu + 2,58\sigma) = 99 \%$

La **largeur de la distribution à mi-hauteur**⁴ w , c'est-à-dire entre ses deux points d'inflexion, vaut :

$$w = 2\sigma\sqrt{2\ln 2} \quad (3.102)$$

La **hauteur maximale** h de la distribution dépend de l'écart type :

$$h = \frac{1}{\sigma\sqrt{2\pi}} \quad (3.103)$$

La loi normale n'est pas tabulée puisqu'il faudrait autant de tables numériques que de valeurs possibles pour la moyenne μ et l'écart-type σ . C'est pourquoi, en opérant un changement de variable, la loi normale générale utilisée est la **loi normale centrée réduite**, qui, elle, est tabulée. On la note $N(0, 1)$.

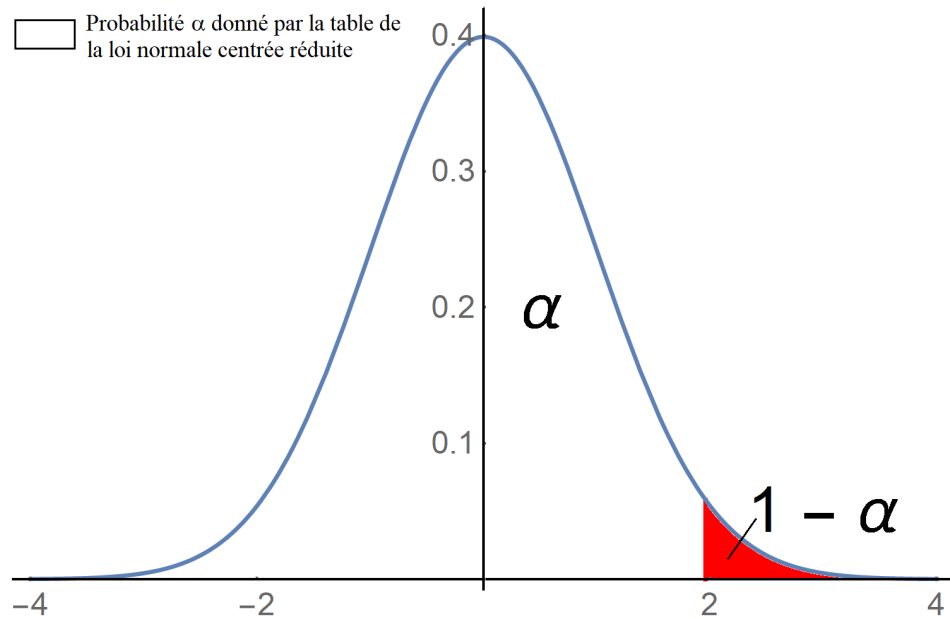
$$N(0, 1) = \Pr(u, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (3.104)$$

La table donne les valeurs pour une probabilité α d'être inférieure à un certain quantile ε . La table fournit le quantile positif et donne la probabilité à gauche (Fig. 3.2).

$$\alpha = \Pr(Z < \varepsilon) = F(z_\alpha) \quad (3.105)$$

Par exemple, on souhaite calculer la probabilité que la variable z soit inférieure à 2,5 %. Comme vu dans la description de la loi normale, il faut prendre $z_\alpha = 1,96$. Pour lire la valeur dans la table de la loi normale centrée réduite, il faut décomposer la valeur en $1,96 = 1,90 + 0,06$. En colonne, on cherche 1,90, et, en ligne, on cherche 0,06. On trouve dans la table que $\alpha = 0,9750 = \Pr(Z < 1,96)$.

4. *Full Width at Half Maximum* (F.W.H.M.)

FIGURE 3.2 – Loi normale centrée réduite et probabilité critique α

De fait, $1 - \alpha = 0,025 = \Pr(Z > 1,96)$, qui correspond bien à la probabilité recherchée.

On travaille souvent sur les fréquences cumulées (donc la fonction de répartition) avec l'anamorphose galtonienne dans laquelle les fréquences cumulées forment une droite (ou droite de Henry).

Pierre Jean Paul
Henry (1848-
1907)

Propriétés

D'après C. F. Gauss, la limite d'une somme de variables aléatoires tend vers une loi normale.

Les lois normales peuvent s'additionner. Soient $X \sim N(\mu_1, \sigma_1)$ et $Y \sim N(\mu_2, \sigma_2)$ avec X et Y indépendantes, alors :

$$X + Y \sim N(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}) \quad (3.106)$$

Ce résultat se généralise à la somme de n variables aléatoires normales indépendantes.

Si les lois sont centrées réduites, alors on $X \sim N(0, 1)$ et $Y \sim N(0, 1)$ avec X et Y indépendantes. Le résultat se réduit à :

$$X + Y \sim N(0, \sqrt{2}) \quad (3.107)$$

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 23

Pour un **mouvement brownien** (ou un mouvement gaussien), les accroissements $B(t_2) - B(t_1)$ et $B(t_4) - B(t_3)$ sont indépendants. Pour tout t , $B(t) - B(0)$ suit une loi normale de variance égale à t . Si cet accroissement a une espérance nulle, la distribution suit une loi d'échelle telle que :

$$\frac{B(t + \Delta t) - B(t)}{\sqrt{\Delta t}} \quad (3.108)$$

qui est indépendante de t et de Δt

Théorème central limite de P.-S. de Laplace

La loi de la somme de n variables aléatoires indépendantes et de variance finie tend vers une loi normale lorsque n augmente.

Comparaison de deux distributions avec une loi normale

La probabilité d'un intervalle, par exemple, sur l'intervalle de données : $[-1, 3]$, est donnée par une loi normale $X \sim N(1, 18; 1, 69)$.

On commence par centrer et réduire les valeurs bornant l'intervalle.

$$X_1 = \frac{3 - 1,18}{1,69} = 1,08 \quad (3.109)$$

et

$$X_2 = \frac{-1 - 1,18}{1,69} = -1,29 \quad (3.110)$$

Dans la table de probabilité de la loi normale centrée réduite, on regarde les valeurs respectives p_1 et p_2 de X_1 et de X_2 .

$$X_1 - (1 - X_2) = n \quad (3.111)$$

On transforme n en un pourcentage. $n = 0,7615$, soit 76,15 %.

On peut alors comparer ce pourcentage avec la répartition étudiée. On pose $N = \frac{\text{effectif de l'intervalle}}{\text{effectif total}} \times 100$. Ici, $N = 76$ %, donc la loi normale est bien sa distribution de probabilité. Par contre, si $N = 50$ % ou $N = 90$ %, alors la loi normale ne correspond pas à la distribution de probabilité.

Notes au sujet de la loi normale

Ce n'est pas parce que la loi est dite normale qu'elle est toujours valable. « Rappelons le vieux cliché de l'époque 1900 : on disait que les physiciens utilisaient la distribution gaussienne pour les « erreurs », parce qu'ils croyaient que les

mathématiciens en avaient établi la nécessité, tandis que les mathématiciens étudiaient seulement la distribution gaussienne, parce qu'ils croyaient que les physiciens en avaient établi l'universelle applicabilité » [Mandelbrot, 1997, p. 88]. Dès le début du XX^e siècle, on savait que les variables aléatoires concernant la fréquence du dénombrement de certains objets géographiques (lacs, montagnes, etc.) ne suivait pas une loi normale [Fréchet, 1941] [Korčák, 1940].

On qualifie souvent une variable suivant une loi normale de « **bruit blanc** ». Un « bruit » est une manière de désigner une variable aléatoire. Le terme « blanc » est utilisé pour une variable indépendante. Cela étant, il faut bien faire attention un « bruit blanc » peut de façon grossière être distribué par une autre loi que la loi normale, d'où la nécessité de connaître d'autres distributions.

3.2.3 Loi log-normale

Robert Gibrat
(1904-1980)

La loi log-normale s'appelle aussi : loi de Galton ou loi de Gibrat.

Soit Y une variable aléatoire suivant la loi normale $N(\mu, \sigma)$. La variable aléatoire X définie par $X = e^Y$ suit, par définition, une loi log-normale.

La densité de la loi de la variable X se déduit de celle de la variable Y par le changement de variable : $x \rightarrow e^y$ (Fig. 3.3).

$$\begin{cases} f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2} & \text{si } x \geq 0 \\ f(x) = 0 & \text{si } x < 0 \end{cases} \quad (3.112)$$

Le facteur $\frac{1}{x}$ dans l'expression de la densité est un **facteur de pondération**.

L'espérance de la loi log-normale vaut :

$$\mathbb{E}(X) = e^{\mu + \frac{\sigma^2}{2}} \quad (3.113)$$

La variance de la loi log-normale vaut :

$$\mathbb{V}(X) = e^{\sigma^2 - 1} e^{2\mu + \sigma^2} \quad (3.114)$$

Remarque. Si l'écart type σ est petit, alors la loi de la variable aléatoire X est proche de la loi normale.

Propriété. Le produit de n variables aléatoires indépendantes suivant une loi log-normale suit lui aussi une loi log-normale.

3.2.4 Loi uniforme continue

Soient $a < b$, la fonction de distribution d'une loi uniforme continue est :

$$\Pr_{a,b}(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]} \quad (3.115)$$

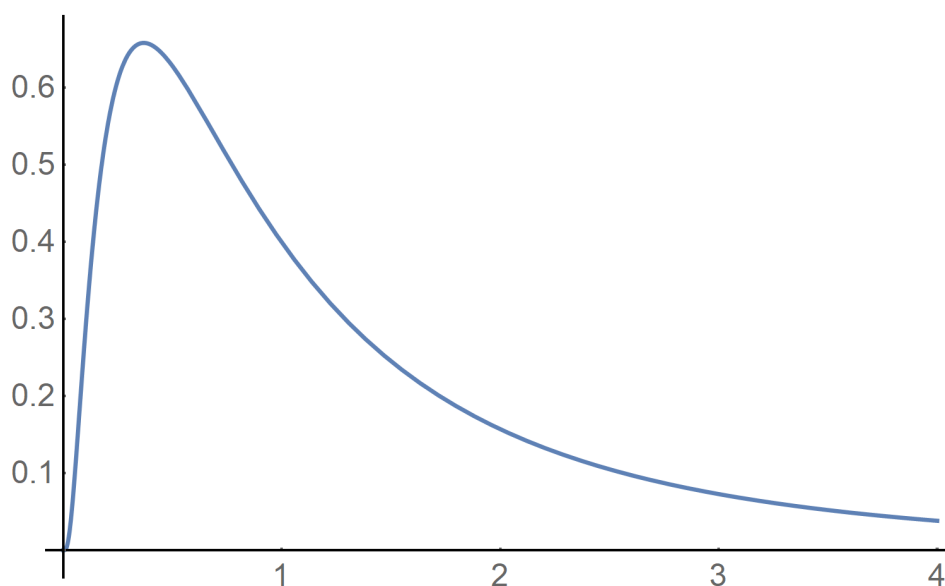
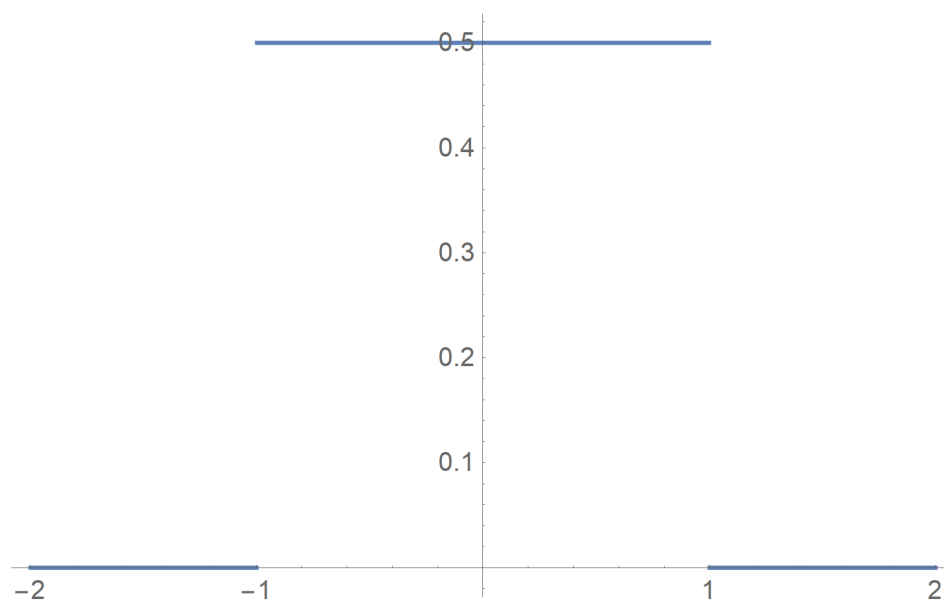


FIGURE 3.3 – Loi log-normale centrée réduite

avec $\mathbf{1}_{[a,b]}$ signifiant que, en dehors du domaine de définition $[a, b]$ la fonction de distribution est nulle (Fig. 3.4). $\mathbf{1}_{[a,b]}$ est la **fonction caractéristique** de l'ensemble $[a, b]$.

$\mathbf{1}_{[a,b]}$ correspond au cardinal d'un ensemble infini. Ce nombre est appelé **aleph**.


FIGURE 3.4 – Loi uniforme avec $a = -1$ et $b = 1$

L'espérance de la fonction uniforme est :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_a^b x \frac{1}{b-a} dx = \left[\frac{1}{2(b-a)} x^2 \right]_a^b = \frac{b-a}{2} \quad (3.116)$$

La variance de la fonction uniforme est :

$$\begin{aligned} \mathbb{V}(X) &= \int_{-\infty}^{+\infty} f(x) f(x - \mathbb{E}(X))^2 dx = \int_a^b \frac{1}{b-a} \left(x - \frac{b-a}{2}\right)^2 dx \\ \mathbb{V}(X) &= \left[\frac{1}{3(b-a)} (x - \mathbb{E}(X))^3 \right]_a^b \\ \mathbb{V}(X) &= \frac{\frac{1}{4}(b-a)^3}{3(b-a)} = \frac{(b-a)^2}{12} \end{aligned} \quad (3.117)$$

Propriété 1. La somme de deux variables aléatoires, indépendantes ou non, suivant une loi uniforme sur $[a, b]$, ne suit pas une loi uniforme sur $[a, b]$.

Propriété 2. L'image, par sa fonction de répartition, de toute variable aléatoire réelle continue, est une variable aléatoire réelle suivant la loi uniforme sur $[0, 1]$. Cette propriété est utilisée, pour simuler ou engendrer des échantillons de la loi de la variable aléatoire X .

Propriété 3. La loi uniforme sur $[a, b]$ traduit l'hypothèse d'équirépartition, ou répartition indifférente, sur $[a, b]$.

Propriété 4. La loi uniforme est utilisée en statistique bayésienne, pour déterminer les lois de probabilité *a priori*, dans le cas de l'ignorance totale, dans l'intervalle $[0, 1]$ (elle est dite non informative), ou dans l'intervalle $[a, b]$, en utilisant les résultats de l'expert (elle est dite informative).

3.2.5 Loi triangulaire

Soient $a < c < b$, la fonction triangulaire est définie par une construction basée sur les deux fonctions de distributions suivantes :

$$\Pr_{a,c}(x) = \frac{2(x-a)}{(b-a)(c-a)} \mathbf{1}_{[a,c]} \quad (3.118)$$

et

$$\Pr_{c,b}(x) = \frac{2(b-x)}{(b-a)(b-c)} \mathbf{1}_{[c,b]} \quad (3.119)$$

a est assimilée à la **valeur optimiste**, c , la **valeur attendue** (c'est-à-dire le mode) et b , la **valeur pessimiste**.

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 27

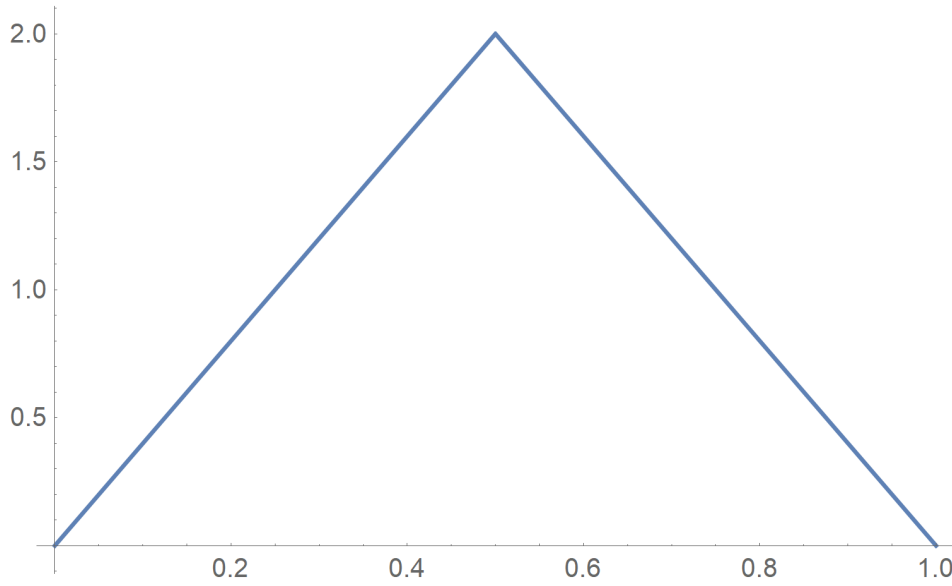


FIGURE 3.5 – Exemple de loi triangulaire

Le triangle de base $(c - a)$ doit avoir une hauteur $h = \frac{2}{c-a}$ telle que sa surface totale soit égale à l'unité.

$$\mathcal{A}_{\Delta} = \frac{(c - a) \times h}{2} = \frac{c - a}{c - a} = 1 \quad (3.120)$$

La pente de la première droite (croissante à gauche) est :

$$\Delta_1 = \frac{2}{(b - a)(c - a)} \quad (3.121)$$

et la pente de la seconde droite (décroissante à droite) est :

$$\Delta_2 = -\frac{2}{(b - a)(c - a)} \quad (3.122)$$

La fonction triangulaire est une fonction de distribution si :

$$\int_{-\infty}^{+\infty} [\text{Pr}_{a,c}(x) + \text{Pr}_{c,b}(x)] dx = 1 = \mathcal{A}_{\Delta} \quad (3.123)$$

puisque l'intégrale correspond à l'aire du triangle.

L'espérance de la fonction triangulaire est :

$$\mathbb{E}(X) = \frac{a + b + c}{3} \quad (3.124)$$

Ce résultat se démontre en faisant une intégration directe :

$$\int_{-\infty}^{+\infty} x f(x) dx = \int_a^c x \frac{2(x-a)}{(b-a)(c-a)} dx + \int_c^b x \frac{2(b-x)}{(b-a)(b-c)} dx \quad (3.125)$$

La variance de la fonction triangulaire est :

$$\mathbb{V}(X) = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18} \quad (3.126)$$

Ce résultat se démontre en intégrant directement :

$$\begin{aligned} \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x) dx &= \int_a^c (x - \mathbb{E}(X))^2 \frac{2(x-a)}{(b-a)(c-a)} dx \\ &+ \int_c^b (x - \mathbb{E}(X))^2 \frac{2(b-x)}{(b-a)(b-c)} dx \end{aligned} \quad (3.127)$$

Propriété. Sous condition que leurs bornes soient identiques, la somme de deux variables aléatoires indépendantes chacune de loi uniforme $[a, b]$ suit une loi triangulaire sur $[2a, 2b]$

3.2.6 Loi exponentielle

Soit $\lambda > 0$, la fonction :

$$\begin{cases} \Pr(x) = \lambda e^{-\lambda x} \mathbf{1}_{[0, +\infty[} & \text{si } x \geq 0 \\ \Pr(x) = 0 & \text{si } x < 0 \end{cases} \quad (3.128)$$

est une fonction de distribution, car elle vérifie :

$$\Pr(X < t) = \int_{-\infty}^{+\infty} \Pr_\lambda(X) dx = \int_{-\infty}^{+\infty} \lambda e^{-\lambda x} dx = \lim_{t \rightarrow +\infty} (-e^{-\lambda t} + 1) = 1 \quad (3.129)$$

L'espérance de la fonction exponentielle est :

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} x \Pr_\lambda(X) dx = \int_{-\infty}^{+\infty} x \lambda e^{-\lambda x} dx \\ \mathbb{E}(X) &= \left[-e^{-\lambda x} \right]_a^b - \int_a^b -e^{-\lambda x} dx = \int_0^{+\infty} e^{-\lambda x} dx \\ \mathbb{E}(X) &= \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_0^{+\infty} = \frac{1}{\lambda} \end{aligned} \quad (3.130)$$

La variance de la fonction exponentielle est :

$$\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \quad (3.131)$$

On calcule $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = \int_0^{+\infty} \lambda x^2 e^{-\lambda x} dx \quad (3.132)$$

On pose $y = \lambda x$.

$$\mathbb{E}(X^2) = \frac{1}{\lambda^2} \int_0^{+\infty} y^2 e^{-y} dy \quad (3.133)$$

Une double intégration par parties donne :

$$\int_0^{+\infty} y^2 e^{-y} dy = [-y^2 e^{-y}]_0^{+\infty} + 2 \int_0^{+\infty} y e^{-y} dy = 2 [-y e^{-y}]_0^{+\infty} + 2 \int_0^{+\infty} e^{-y} dy = 2 \quad (3.134)$$

d'où

$$\mathbb{E}(X^2) = \frac{2}{\lambda^2} \quad (3.135)$$

Dit autrement,

$$\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \quad (3.136)$$

Le coefficient d'asymétrie de la loi exponentielle est :

$$\gamma_1 = 2 \quad (3.137)$$

Le coefficient d'aplatissement de la loi exponentielle est :

$$\gamma_2 = 9 \quad (3.138)$$

Propriété 1. La somme de deux variables aléatoires indépendantes, suivant des lois exponentielles de paramètres respectifs λ_1 et λ_2 , est une variable aléatoire suivant une loi exponentielle de paramètre $\lambda_1 + \lambda_2$.

Propriété 2. La loi exponentielle est qualifiée de loi « sans mémoire ». Elle permet la modélisation du comportement des matériels fonctionnant avec un taux de défaillance constant.

Propriété 3. On considère un matériel ayant fonctionné sans défaillance pendant le temps x_1 et on cherche la probabilité qu'il soit encore en état de marche au temps $x + x_1$. La définition de la probabilité conditionnelle donne :

$$\Pr(X \geq x + x_1 \mid X \geq x_1) = \frac{\Pr(X \geq x + x_1 \text{ et } X \geq x_1)}{\Pr(X \geq x_1)} = \frac{e^{-\lambda(x+x_1)}}{e^{-\lambda x_1}} = e^{-\lambda x} \quad (3.139)$$

Le matériel a « oublié » qu'il avait déjà fonctionné pendant le temps x_1 . Pour ce type de matériel, il est inutile de procéder à un remplacement préventif.

Les domaines d'utilisation de la loi exponentielle sont :

- pour la distribution exponentielle qui est associée aux processus de Poisson ;
- pour la loi exponentielle qui est utilisée en fiabilité. Le paramètre λ représente le taux de défaillance alors que son inverse $\theta = \frac{1}{\lambda}$ est le temps moyen de bon fonctionnement⁵. Avec le paramètre θ , la densité de probabilité s'écrit :

$$f(x) = \frac{1}{\theta} e^{-\frac{1}{\theta}x} \quad (3.140)$$

et

$$\mathbb{E}(X) = \theta \text{ et } \mathbb{V}(X) = \theta^2 \quad (3.141)$$

- pour la loi exponentielle qui s'applique aux matériels électroniques.

3.2.7 Loi de Cauchy

Augustin Cauchy
(1789-1857)

La fonction

$$\Pr(x) = \frac{1}{\pi} \left(\frac{1}{1+x^2} \right) \quad (3.142)$$

est une fonction de distribution, car elle vérifie :

$$\int_{-\infty}^{+\infty} \Pr(x) dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{1}{1+x^2} dx = \frac{1}{\pi} (\arctan(+\infty) - \arctan(-\infty)) = 1 \quad (3.143)$$

La forme de la courbe est une ogive beaucoup plus pointue que la loi normale (Fig. 3.6).

L'espérance de la fonction de Cauchy vaut :

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} x \Pr(x) dx = \frac{1}{\pi} \left[\int_{-\infty}^0 \frac{x}{1+x^2} + \int_0^{+\infty} \frac{x}{1+x^2} \right] \\ \mathbb{E}(X) &= \frac{1}{2\pi} \left([\ln(1+x^2)]_{-\infty}^0 + [\ln(1+x^2)]_0^{+\infty} \right) = \frac{1}{2\pi} [\ln(1) - \ln(\infty) + \ln(1) - \ln(1)] = 0 \end{aligned} \quad (3.144)$$

La variance de la fonction de Cauchy vaut :

⁵ Mean Time Between Failure (M.T.B.F.)

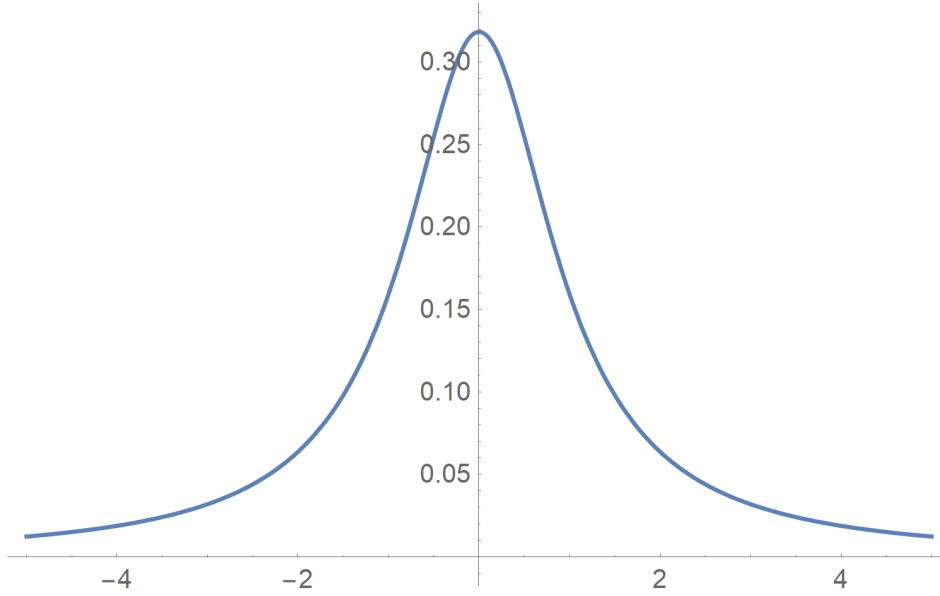


FIGURE 3.6 – Loi de Cauchy

$$\begin{aligned}
 \mathbb{V}(X) &= \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 \Pr(x) dx \\
 \mathbb{V}(X) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x^2}{1+x^2} dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} \left(1 - \frac{1}{1+x^2}\right) dx \\
 \mathbb{V}(X) &= \frac{2}{\pi} \lim_{t \rightarrow +\infty} \int_0^t \left(1 - \frac{1}{1+x^2}\right) dx = \frac{2}{\pi} \lim_{t \rightarrow +\infty} \int_0^t (t - \arctan t) dx = +\infty
 \end{aligned} \tag{3.145}$$

De manière plus générale, avec $a \in \mathbb{R}_+^*$ et $x \in \mathbb{R}$, la densité de probabilité est $\Pr(x) = \frac{1}{\pi} \left(\frac{a}{a^2 + x^2} \right)$. $-a$ et $+a$ correspondent aux points d'inflexion de la densité. La fonction de répartition est $F(x) = \frac{1}{\pi} \arctan \frac{x}{a} + \frac{1}{2}$. La loi de Cauchy est centrée sur $\Pr(X = 0) = \frac{1}{2}$, $\Pr(X = -a) = \frac{1}{4}$ et $\Pr(X = +a) = \frac{3}{4}$.

Tous les moments de la loi de Cauchy générale divergent. Pour $a > 0$, la moyenne n'existe pas ; elle est infinie. Il en va de même pour la variance.

$$\mathbb{E}(X) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x}{1+x^2} dx = +\infty \tag{3.146}$$

et

$$\mathbb{V}(X) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x^2}{1+x^2} dx = +\infty \tag{3.147}$$

Propriété 1. Si la loi X est la loi de Cauchy de paramètre 1, pour $a > 0$, la loi de aX est la loi de Cauchy de paramètre a .

Propriété 2. Si la loi X est la loi de Cauchy de paramètre a , alors $\frac{1}{X}$ a pour loi la loi de Cauchy de paramètre $\frac{1}{a}$.

Propriété 3. Le quotient de deux variables aléatoires X et Y de lois normales, de moyenne 0 et de variances respectives $\sigma_X^2 > 0$ et $\sigma_Y^2 > 0$. Le quotient $\frac{X}{Y}$ est presque sûrement bien défini. Sa loi est la loi de Cauchy de paramètre $a = \frac{\sigma_Y}{\sigma_X}$.

N.B. il arrive souvent que l'on centre la loi de Cauchy en un réel $x_0 \neq 0$.

$$\Pr(x) = \frac{1}{\pi} \left(\frac{a}{a^2 + (x - x_0)^2} \right) \quad (3.148)$$

$$a \in \mathbb{R}_+^* \text{ et } x \in \mathbb{R}.$$

Un **processus de Cauchy** (ou bruit de Cauchy blanc) est un processus à valeurs de Cauchy **indépendantes**. La moyenne est une variable réduite de Cauchy :

$$\mu = \frac{1}{T} \sum_{t=1}^T X(t) \quad (3.149)$$

La moyenne est identique aux $X(t)$ en distribution. Prendre sa moyenne n'offre aucun effet sur la variabilité. La loi des grands nombres ne s'appliquent pas. C'est un bruit de Cauchy blanc non ergodique. La moyenne ne converge pas. Elle reste aléatoire et suit elle-même une loi de Cauchy.

Le **mouvement de Cauchy** est une fonction $C(t)$ dont les incréments $C(t + \Delta t) - C(t)$ sont des variables aléatoires indépendantes de Cauchy. La distribution

$$\frac{C(t + \Delta t) - C(t)}{\Delta t} \quad (3.150)$$

est indépendante de t et de Δt .

La fonction caractéristique des moments est définie par la transformée de Fourier suivante :

$$\varphi(\theta) = \int_{-\infty}^{+\infty} e^{i\theta x} \times \frac{1}{\pi} \left(\frac{a}{a^2 + x^2} \right) dx = e^{-|a\theta|} \quad (3.151)$$

3.2.8 Loi bêta de types I et II

La fonction gamma d'Euler

On définit la fonction gamma d'Euler comme étant la relation :

$$\Gamma(z) = \int_0^{+\infty} e^{-x} x^{z-1} dx \quad (3.152)$$

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 33

avec $Z \in \mathbb{C}$.

On peut démontrer que cette fonction est bien définie, c'est-à-dire que l'intégrale existe pour $\Re(z) > 0$.

Propriété 1. $\Gamma(z+1) = z\Gamma(z)$ (ce qui se démontre en intégrant par parties).

Propriété 2. $\Gamma(a)\Gamma(b) = \lim_{R \rightarrow +\infty} \int_{A_R} \int e^{-x-y} x^{a-1} y^{b-1} dx dy$
avec $A_R = \{(x, y) \mid x \geq 0, y \geq 0, x+y \geq R\}$. On opère un changement de variable $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u-v \\ v \end{pmatrix}$ et on obtient

$$\Gamma(a)\Gamma(b) = \lim_{R \rightarrow +\infty} \int_0^R e^{-u} \left(\int_0^u (u-v)^{a-1} v^{b-1} dv \right) du \quad (3.153)$$

Pour l'intégrale interne, on utilise la substitution $v = ut$ avec $0 \leq t \leq 1$ et on trouve :

$$\Gamma(a)\Gamma(b) = \lim_{R \rightarrow +\infty} \int_0^R e^{-u} u^{a+b-1} du \int_0^1 (1-t)^{a-1} t^{b-1} dt \quad (3.154)$$

On pose $B(a, b) = \int_0^1 (1-t)^{a-1} t^{b-1} dt$

$$\Gamma(a)\Gamma(b) = B(a, b) \int_0^{+\infty} e^{-u} u^{a+b-1} du = B(a, b) \Gamma(a+b) \quad (3.155)$$

La fonction bêta de type I

La fonction B qui apparaît dans $\Gamma(a)\Gamma(b) = B(a, b)\Gamma(a+b)$ est la **fonction bêta**.

$$B(a, b) = \int_0^1 (1-t)^{a-1} t^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad (3.156)$$

Pour $a > 0$ et $b > 0$, on considère maintenant la fonction de distribution :

$$\Pr_{a,b}(x) = \frac{x^{a-1} (1-x)^{b-1}}{B(a, b)} \mathbf{1}_{]0,1[} \quad (3.157)$$

avec $B(a, b) = \int_0^{+\infty} x^{a-1} (1-x)^{b-1} dx$.

$$\begin{aligned} \int_{-\infty}^{+\infty} \Pr_{a,b}(x) dx &= \int_{-\infty}^{+\infty} \frac{x^{a-1} (1-x)^{b-1}}{B(a, b)} \mathbf{1}_{]0,1[} dx = \int_{-\infty}^{+\infty} \frac{x^{a-1} (1-x)^{b-1}}{K} \mathbf{1}_{]0,1[} dx \\ \int_{-\infty}^{+\infty} \Pr_{a,b}(x) dx &= \frac{1}{K} \int_{-\infty}^{+\infty} x^{a-1} (1-x)^{b-1} dx \end{aligned} \quad (3.158)$$

$$\int_{-\infty}^{+\infty} \text{Pr}_{a,b}(x) dx = \frac{1}{K} B(a, b) = \frac{1}{K} K = 1 \quad (3.159)$$

L'espérance de la fonction bêta de type I est :

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} x \text{Pr}_{a,b}(x) dx = \frac{1}{B(a,b)} \int_0^1 x^{a-1} (1-x)^{b-1} dx \\ \mathbb{E}(X) &= \frac{B(a+1,b)}{B(a,b)} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b+1)} = \frac{a}{a+b} \end{aligned} \quad (3.160)$$

La variance de la fonction bêta de type I est :

$$\begin{aligned} \mathbb{V}(X) &= \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x) dx = \frac{1}{B(a,b)} \int_0^1 (x - \mathbb{E}(X))^2 x^{a-1} (1-x)^{b-1} dx \\ \mathbb{V}(X) &= \frac{1}{B(a,b)} (B(a+2,b) - (\mathbb{E}(X))^2 B(a,b)) \end{aligned} \quad (3.161)$$

En sachant que $\Gamma(z+1) = z\Gamma(z)$ et que $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, on trouve :

$$B(a+2,b) = \frac{\mathbb{E}(X)(a+1)}{a+b+1} B(a,b) \quad (3.162)$$

donc

$$\begin{aligned} \mathbb{V}(X) &= \frac{1}{B(a,b)} \left(\frac{\mathbb{E}(X)(a+b)}{a+b+1} B(a,b) - (\mathbb{E}(X))^2 B(a,b) \right) \\ \mathbb{V}(X) &= \left(\frac{\mathbb{E}(X)(a+b)}{a+b+1} - (\mathbb{E}(X))^2 \right) = \frac{ab}{(a+b)^2(a+b+1)} \end{aligned} \quad (3.163)$$

La fonction bêta de type II

Soit X une variable aléatoire suivant une loi bêta de type I, $B(n, p)$. La variable Y , positive ou nulle, définie par $Y = \frac{X}{1-X}$, suit une loi bêta de type II dont la densité s'obtient facilement par le changement de variables :

$$\begin{cases} f(y) = \frac{1}{B(n,p)} \frac{y^{n-1}}{(1+y)^{n+p}} & \text{si } y \geq 0 \\ f(y) = 0 & \text{si } y < 0 \end{cases} \quad (3.164)$$

L'espérance de la fonction bêta de type II est :

$$\mathbb{E}(X) = \frac{n}{p-1} \quad (3.165)$$

La variance de la fonction bêta de type II est :

$$\mathbb{V}(X) = \frac{n(n+p-1)}{(p-1)^2(p-2)^2} \quad (3.166)$$

3.2.9 Loi gamma

Pour $a > 0$ et $\lambda > 0$, on considère la fonction de distribution :

$$\begin{cases} \Pr_{a,\lambda}(x) = \frac{x^{a-1}e^{-\lambda x}}{\int_0^{+\infty} x^{a-1}e^{-\lambda x}dx} \mathbf{1}_{]0,+\infty[} & \text{si } x \geq 0 \\ \Pr_{a,\lambda}(x) = 0 & \text{si } x < 0 \end{cases} \quad (3.167)$$

On effectue un changement de variable $t = \lambda x$ avec $\lambda \neq 1$, et on obtient :

$$\gamma(a) = \int_0^{+\infty} x^{a-1}e^{-\lambda x}dx = \frac{\Gamma(a)}{\lambda^a} \quad (3.168)$$

donc

$$\Pr_{a,\lambda}(x) = \frac{x^{a-1}e^{-\lambda x}}{\Gamma(a)} \lambda^a \mathbf{1}_{]0,+\infty[} \quad (3.169)$$

Γ est la **fonction eulérienne** définie par l'intégrale pour $t > 0$.

$$\Gamma(a) = \int_0^{+\infty} e^{-y} y^{a-1} dy \quad (3.170)$$

avec t le paramètre de forme et $\frac{1}{\lambda}$ le paramètre d'échelle. Il est à noter que si $a = 1$, on trouve la loi exponentielle.

Par un raisonnement similaire en tout point à celui de la fonction bêta, on montre que :

$$\int_{-\infty}^{+\infty} \Pr_{a,\lambda}(x) dx = 1 \quad (3.171)$$

L'espérance de la fonction gamma est :

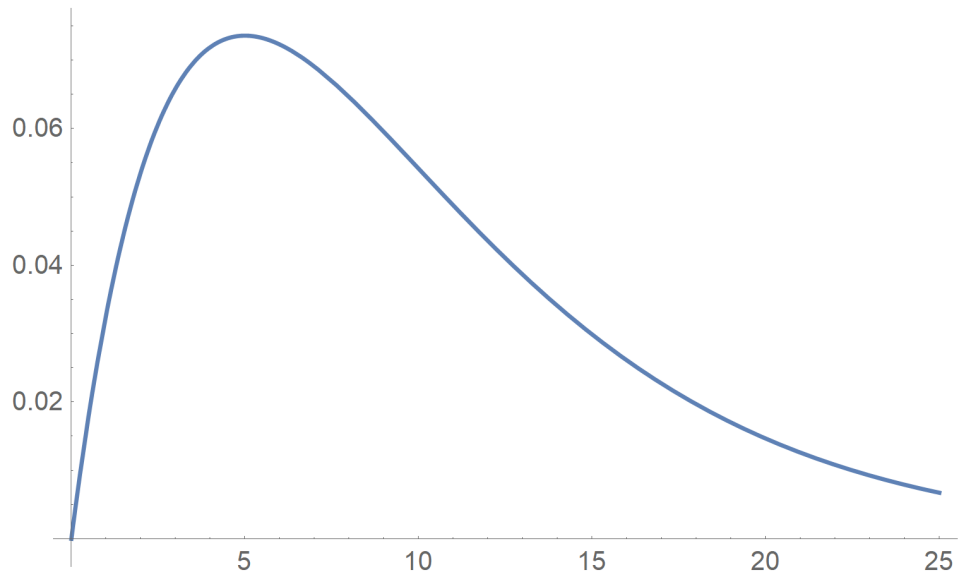
$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} x f(x) dx = \frac{\lambda^a}{\Gamma(a)} \int_0^{+\infty} x^a e^{-\lambda x} dx \\ \mathbb{E}(X) &= \frac{\lambda^a}{\Gamma(a)} \frac{a\Gamma(a)}{\lambda^{a+1}} = \frac{\lambda^a}{\Gamma(a)} \frac{a\Gamma(a)}{\lambda^{a+1}} \\ \mathbb{E}(X) &= \frac{a}{\lambda} \end{aligned} \quad (3.172)$$

La variance de la fonction gamma est :

$$\mathbb{V}(X) = \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x) dx = \frac{\lambda^a}{\Gamma(a)} \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 x^{a-1} e^{-\lambda x} dx = \frac{a}{\lambda^2} \quad (3.173)$$

Propriété 1 Selon les valeurs des paramètres, la loi gamma s'identifie à d'autres lois :

1. Si $t = 1$, alors on a la loi exponentielle.

FIGURE 3.7 – Loi gamma avec $a = 2$ et $\lambda = 5$

2. Si $t = n$, avec $n \in \mathbb{N} \setminus n > 1$ alors sa densité est :

$$\begin{cases} f(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{n-1}}{(n-1)!} & \text{si } x \geq 0 \\ f(x) = 0 & \text{si } x < 0 \end{cases} \quad (3.174)$$

avec $\Gamma(n) = (n-1)!$. Sa fonction de répartition, qui correspond à l'apparition de n événements en un temps inférieur à x , est donnée par l'expression :

$$\begin{cases} F(x) = 1 - e^{-\lambda x} \sum_{i=1}^n \left[\frac{(\lambda x)^{i-1}}{(i-1)!} \right] & \text{si } x \geq 0 \\ F(x) = 0 & \text{si } x < 0 \end{cases} \quad (3.175)$$

Agner Krarup Erlang (1878-1929)

On obtient la loi de A. K. Erlang.

3. La loi de la variable χ^2 à n degrés de liberté, $\chi^2(n)$ est obtenu si $\lambda = \frac{1}{2}$ et $t = \frac{n}{2}$ où $n \in \mathbb{N}$.

Propriété 2 La somme de deux variables aléatoires indépendantes, suivant des lois gamma $\gamma(t, \lambda)$ et $\gamma(u, \lambda)$, suit une loi gamma $\gamma(t + u, \lambda)$.

Les domaines d'utilisation de la fonction gamma sont :

- la théorie des files d'attentes ;
- la fiabilité.

3.2.10 Loi de χ^2

arl Pearson
(1857-1936)

La distribution du χ^2 fut introduite par K. Pearson en 1900.

On considère la fonction de distribution gamma dans le cas particulier $a = \frac{k}{2}$ avec $\lambda = \frac{1}{2}$ et $k \in \mathbb{N}$:

$$\text{Pr}_k(x) = \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} \mathbf{1}_{[0,+\infty[} \quad (3.176)$$

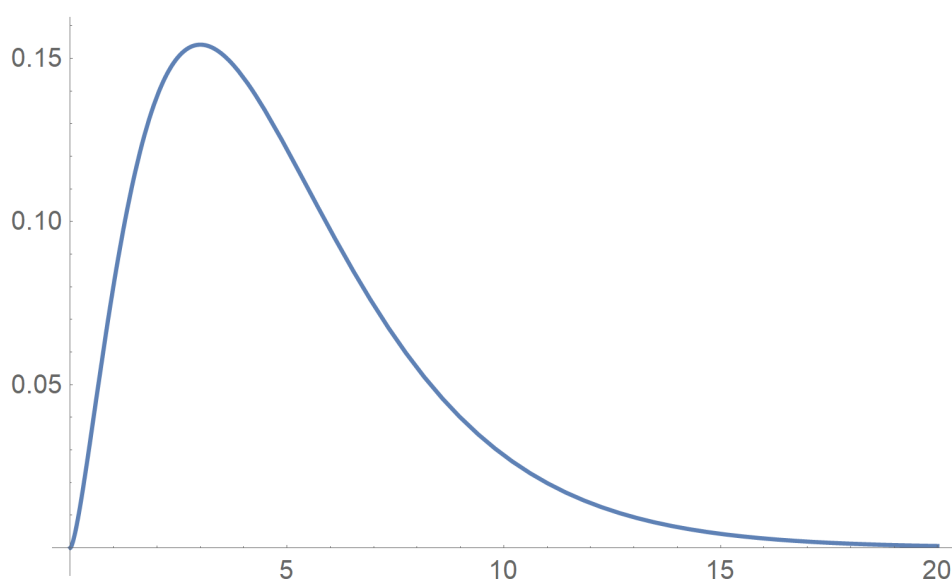


FIGURE 3.8 – Loi du χ^2 avec $k = 5$ pour une variable centrée réduite

Tous les calculs faits auparavant s'appliquent :

$$\begin{cases} \mathbb{E}(X) = k \\ \mathbb{V}(X) = 2k \end{cases} \quad (3.177)$$

Variable du χ^2 La variable aléatoire U valant la somme des carrés de v variables aléatoires indépendantes, centrées, réduites, suivant la loi normale, suit la loi du χ^2 , à v degrés de liberté :

$$\chi^2(v) = \sum_{i=1}^v U_i^2 \quad (3.178)$$

avec $U_i = \frac{X_i - \mu}{\sigma}$.

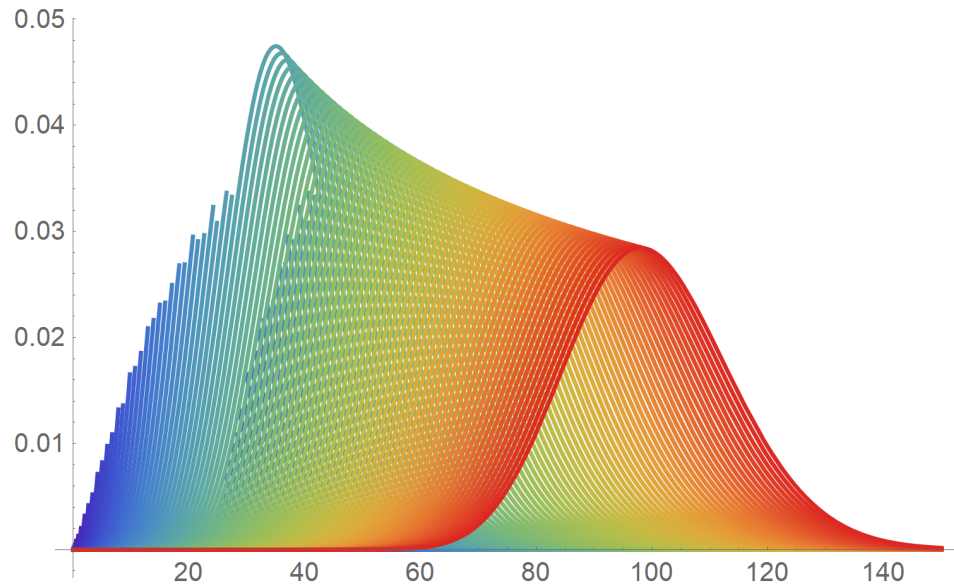


FIGURE 3.9 – Loi du χ^2 de $k = 1$ à $k = 100$ pour une variable centrée réduite : la trajectoire vers une courbe en cloche

Propriété de la variable du χ^2

- La variable aléatoire $\chi^2(v)$ varie de 0 à $+\infty$.
- Le paramètre v est le nombre de degrés de liberté de la variable. Il représente la dimension de l'espace dans lequel se trouve le point représentatif de l'échantillon. Si les variables aléatoires X_i vérifient k relations linéaires, le nombre de degrés de liberté diminue de k .
- La loi suivie par la somme de variables aléatoires indépendantes, suivant chacune des lois du χ^2 , est une loi du χ^2 dont le degré de liberté est la somme des degrés de liberté de chacune des lois.
- Les moments de la variable aléatoire du χ^2 sont :

$$\mathbb{E}(\chi^2(v)) = v \quad (3.179)$$

$$\mathbb{V}(\chi^2(v)) = 2v \quad (3.180)$$

- Le mode M_0 vaut :

$$M_0 = v - 2 \text{ si } v > 2 \quad (3.181)$$

- Selon les valeurs de v , la distribution du χ^2 possède différentes formes, mais la distribution de la variable du χ^2 tend à se rapprocher d'une distribution normale lorsque v augmente (Fig. 3.9).

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 39

- La densité de probabilité de la variable aléatoire $\chi^2(v)$ a une expression mathématique compliquée et peu maniable.

$$f(\chi^2(v)) = \frac{1}{2^{\frac{v}{2}} \Gamma(\frac{v}{2})} e^{-\frac{\chi^2(v)}{2}} (\chi^2(v))^{\frac{v}{2}-1} \quad (3.182)$$

avec Γ la fonction eulérienne.

- Les tables donnent les fractiles d'ordre α :

$$\alpha = \Pr(\chi^2(v) \leq \chi_{\alpha}^2(v)) = \int_0^{\chi_{\alpha}^2(v)} f(\chi^2(v)) d\chi^2(v) \quad (3.183)$$

- La fonction caractéristique de la variable aléatoire $\chi^2(v)$ a pour expression :

$$\varphi_{\chi^2(v)}(t) = \frac{1}{(1 - 2it)^{\frac{v}{2}}} \quad (3.184)$$

avec $i^2 = -1$.

- Pour les grandes valeurs de v , il existe plusieurs formes limites.

1. Une première forme est obtenue en appliquant le théorème central limite. La loi de la variable aléatoire $\frac{\chi^2(v)-v}{\sqrt{2v}}$ converge, lorsque v tend vers l'infini, vers la loi normale centrée réduite.
 2. La deuxième forme est due à R. A. Fisher. Pour $v \geq 30$, la loi de la variable aléatoire $\sqrt{2\chi^2(v)} - \sqrt{2v-1}$ est la loi normale centrée réduite.
 3. La formule d'approximation de Wilson-Hilferty, pour $v \geq 30$, $\chi_{\alpha}^2(v) \cong v \left(1 - \frac{2}{9v} + U_{\alpha} \sqrt{\frac{2}{9v}}\right)^3$ avec U_{α} le fractile d'ordre α de la loi normale centrée réduite.
- **Relation entre la loi du χ^2 et les lois gamma.** Si la variable aléatoire U est une variable centrée réduite et suivant une loi normale, la densité de la variable $\chi^2(1)$ définie par $T = U^2$ est :

$$g(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t}{2}} t^{-\frac{1}{2}} \quad (3.185)$$

On en déduit que la variable aléatoire $\frac{U^2}{2} = \frac{\chi^2(1)}{2}$ suit une loi $\gamma(\frac{1}{2}, 1)$. La variable $\frac{\chi^2(n)}{2}$ suit alors une loi $\gamma(\frac{n}{2}, 1)$. Cette propriété permet de retrouver la densité de la variable $\chi^2(n)$, ainsi que les différents moments. Si X est une variable aléatoire suivant la loi $\gamma(\frac{n}{2}, 1)$, la variable aléatoire $2X$ suit la loi $\chi^2(n)$.

Edwin Bidwell Wilson (1879-1964)
Margaret M. Hilferty (?-?)

3.2.11 Loi de Fisher-Snedecor

Ronald Aylmer Fisher (1890-1962) et George Snedecor (1881-1974) ont étudié cette distribution en 1924 et calculée en 1934 par G. Snedecor. On appelle aussi cette distribution la **distribution F de Fisher**. Elle joue un rôle dans l'analyse de la variance et dans la régression.

La loi de Fisher-Snedecor se définit par deux paramètres k et l qui caractérisent la relation suivante :

$$f(x) = \frac{\Gamma\left(\frac{k+l}{2}\right) \left(\frac{k}{l}\right)^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{l}{2}\right)} x^{\frac{k}{2}-1} \left(1 + \frac{k}{l}x\right)^{-\frac{k+l}{2}} \quad (3.186)$$

avec $x \geq 0$, $k \in \mathbb{N}$ et $l \in \mathbb{N}$.

Les paramètres k et l correspondent aux **degrés de liberté** des deux lois du χ^2 sous-jacentes. Cette distribution est souvent notée $F_{k,l}$ ou $F(k, l)$.

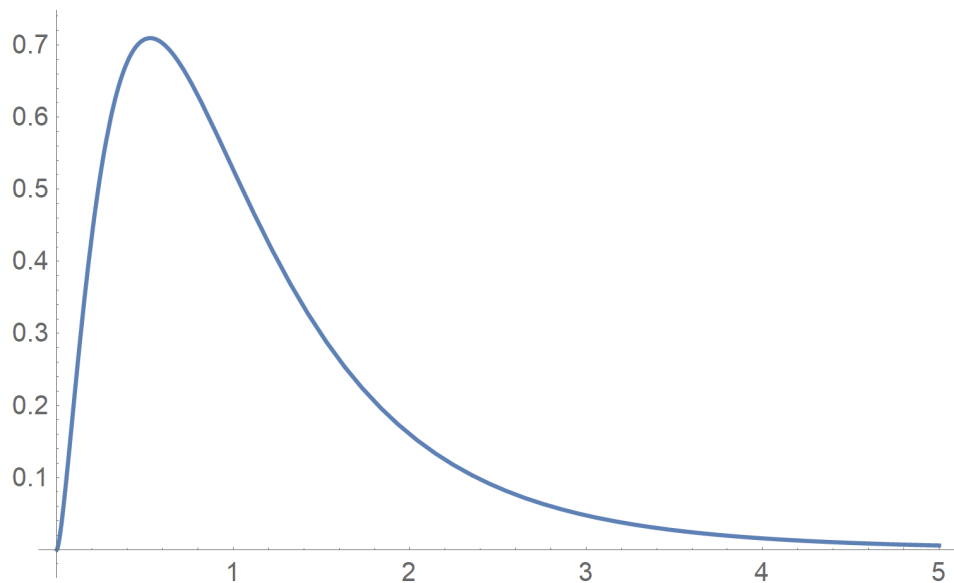


FIGURE 3.10 – Loi du χ^2 pour $k = 1$ à $l = 2$ pour une variable centrée réduite

Il est facile de démontrer qu'il s'agit bien d'une fonction de distribution. En effet, le produit de deux fonctions de distribution (ici celle du χ^2) est une fonction de distribution.

$$\int_{-\infty}^{+\infty} F_{k,l}(x) dx = 1 \quad (3.187)$$

Une démonstration plus longue peut être proposée :

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 41

1. La loi du χ_n^2 est donnée par :

$$f(y) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} e^{-\frac{y}{2}} y^{\frac{n}{2}-1} \quad (3.188)$$

avec $y \geq 0$ et $n \geq 1$.

2. La fonction Γ est définie pour tout $\alpha > 0$ par :

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx \quad (3.189)$$

Soient X et Y deux variables aléatoires indépendantes suivant respectivement les lois χ_n^2 et χ_m^2 . On considère la variable aléatoire :

$$F = \frac{\frac{X}{n}}{\frac{Y}{m}} \quad (3.190)$$

On note F et f les fonctions de répartition et de densité de T . f_X et f_Y sont respectivement les fonctions de densité de X et Y . f est la fonction de densité du couple (X, Y) . Pour tout $t \in \mathbb{R}$,

$$F(t) = \Pr(T \leq t) = \Pr\left(\frac{\frac{X}{n}}{\frac{Y}{m}} \leq t\right) = \int \int_D f(x, y) dx dy = \int \int_D f_X(x) f_Y(y) dx dy \quad (3.191)$$

avec $D = \{(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+ \mid x \leq \frac{nt}{m}y\}$ dans ce domaine de définition les valeurs positives imposées proviennent à l'origine d'une loi du χ^2 pour x et y

$$F(t) = \int \int_D f_X(x) f_Y(y) dx dy = \int_0^{+\infty} f_Y(y) \int_0^{\frac{nt}{m}y} f_X(x) dx dy \quad (3.192)$$

On obtient la fonction de densité de T en dérivant F :

$$f(t) = F'(t) = \frac{n}{m} \int_0^{+\infty} f_Y(y) f_X\left(\frac{nt}{m}y\right) y dy \quad (3.193)$$

or f est une loi du χ^2 :

$$\begin{cases} f(y) = \frac{1}{2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)} e^{-\frac{y}{2}} y^{\frac{m}{2}-1} \\ f\left(\frac{nt}{m}y\right) = \frac{1}{2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)} e^{-\frac{nt}{2m}y} \left(\frac{nt}{m}y\right)^{\frac{m}{2}-1} \end{cases} \quad (3.194)$$

Il ne reste plus qu'à remplacer $f_Y(y)$ et $f_X\left(\frac{nt}{m}y\right)$ par leurs expressions dans $f(t)$.

$$\begin{aligned} f(t) &= \frac{n}{m} \int_0^{+\infty} \frac{1}{2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)} e^{-\frac{y}{2}} y^{\frac{m}{2}-1} \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \left(\frac{nt}{m}y\right)^{\frac{n}{2}-1} e^{-\frac{nt}{2m}y} \\ f(t) &= \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) 2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)} \frac{n}{m} \int_0^{+\infty} e^{-\frac{y}{2}} y^{\frac{m}{2}-1} e^{-\frac{nt}{2m}y} \left(\frac{nt}{m}\right)^{\frac{n}{2}-1} y^{\frac{n}{2}-1} y dy \\ f(t) &= \frac{\frac{n}{m}}{2^{\frac{n+m}{2}} \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{nt}{m}\right)^{\frac{n}{2}-1} \int_0^{+\infty} y^{\frac{n+m}{2}-1} e^{-\frac{y}{2}\left(1+\frac{nt}{m}\right)} dy \end{aligned} \quad (3.195)$$

L'intégrale se résout par un changement de variable.

$$u = \frac{y}{2} \left(1 + \frac{nt}{m}\right) \quad (3.196)$$

d'où

$$\begin{aligned} y &= u \frac{2}{1+\frac{nt}{m}} \\ dy &= du \frac{2}{1+\frac{nt}{m}} \end{aligned} \quad (3.197)$$

Cela permet d'obtenir :

$$f(t) = \frac{\frac{n}{m}}{2^{\frac{n+m}{2}} \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{nt}{m}\right)^{\frac{n}{2}-1} \frac{2^{\frac{n+m}{2}}}{\left(1+\frac{nt}{m}\right)^{\frac{n+m}{2}}} \int_0^{+\infty} u^{\frac{n+m}{2}-1} e^{-u} du \quad (3.198)$$

mais $\int_0^{+\infty} u^{\frac{n+m}{2}-1} e^{-u} du = \Gamma\left(\frac{n+m}{2}\right)$, donc

$$\begin{aligned} f(t) &= \frac{\frac{n}{m} \Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{nt}{m}\right)^{\frac{n}{2}-1} \left(1 + \frac{nt}{m}\right)^{-\frac{n+m}{2}} \\ f(t) &= \frac{\Gamma\left(\frac{n+m}{2}\right) \left(\frac{n}{m}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} t^{\frac{n}{2}-1} \left(1 + \frac{nt}{m}\right)^{-\frac{n+m}{2}} \end{aligned} \quad (3.199)$$

Pour $m > 2$, l'espérance de la loi de Fisher vaut $\mathbb{E}(X) = \frac{m}{m-2}$.

Pour $m > 4$, la variance de la loi de Fisher vaut $\mathbb{V}(X) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}$.

Pour $m > 6$, le coefficient d'asymétrie vaut $\gamma_1 = \frac{(2n+m-2)\sqrt{8(m-4)}}{(m-6)\sqrt{n(n+m-2)}}$.

Pour $m > 8$, le coefficient d'aplatissement vaut $\gamma_2 = 12 \frac{n(5m-22)(n+m-2)+(m-4)(m-2)^2}{n(m-6)(m-8)(n+m-2)}$.

Variable aléatoire de Fisher On considère deux variables aléatoires indépendantes suivant des lois du χ^2 à v_1 et v_2 degrés de liberté respectivement. La variable aléatoire F de Fisher est définie par :

$$F = \frac{\frac{\chi^2(v_1)}{v_1}}{\frac{\chi^2(v_2)}{v_2}} \quad (3.200)$$

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 43

Propriétés de la variable aléatoire de Fisher

1. $F(v_1, v_2) = \frac{1}{F(v_1, v_2)}$
2. La variable $F(v_1, v_2)$ varie de 0 à $+\infty$.
3. La loi de probabilité de la variable F ne dépend que deux paramètres : les degrés de liberté, v_1 et v_2 .
4. Les moments de la variable F de Fisher sont :

$$\mathbb{E}(F) = \frac{v_2}{v_2 - 2} \text{ si } v_2 > 2 \quad (3.201)$$

et

$$\mathbb{V}(F) = \left(\frac{v_2}{v_2 - 2} \right)^2 \frac{2(v_1 + v_2 - 2)}{v_1(v_2 - 4)} \text{ si } v_2 > 4 \quad (3.202)$$

5. Le mode M_0 vaut :

$$M_0 = \frac{v_2(v_1 - 2)}{v_1(v_2 + 2)} \text{ si } v_2 > 2 \quad (3.203)$$

6. La densité a pour forme algébrique :

$$\begin{cases} f \leq 0, g(f) = 0 \\ f > 0, g(f) = \frac{\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}}}{B\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} f^{\frac{v_1}{2}-1} \frac{1}{\left(1 + \frac{v_1}{v_2}f\right)^{\frac{v_1+v_2}{2}}} \end{cases} \quad (3.204)$$

avec B la fonction eulérienne.

7. Les tables donnent les fractiles d'ordre α :

$$\alpha = \Pr(F(v_1, v_2) \leq F_\alpha(v_1, v_2)) = \int_0^{F_\alpha} g(f) df \quad (3.205)$$

$$8. F_\alpha(v_1, v_2) = \frac{1}{F_{1-\alpha}(v_1, v_2)}$$

9. Relation entre la loi de Fisher et les lois bêta.

- (a) Si une variable aléatoire X suit une loi bêta de type I, la variable $\frac{v_1}{v_2} \frac{X}{1-X}$ est une variable de Fisher $F(2v_1, 2v_2)$.
- (b) Si une variable aléatoire Y suit une loi bêta de type II, la variable $\frac{v_1}{v_2} Y$ est une variable de Fisher $F(2v_1, 2v_2)$.

3.2.12 Loi de Student

La distribution T de Student joue un rôle important dans l'étude de la statistique d'un échantillon pour une distribution normale dont on ne connaît pas la variance.

La loi de Student se définit par :

$$\text{Pr}_k(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \sqrt{k\pi}} \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}} \quad (3.206)$$

Il s'agit d'une fonction de distribution, car elle vérifie :

$$\int_{-\infty}^{+\infty} \text{Pr}_k(x) dx = 1 \quad (3.207)$$

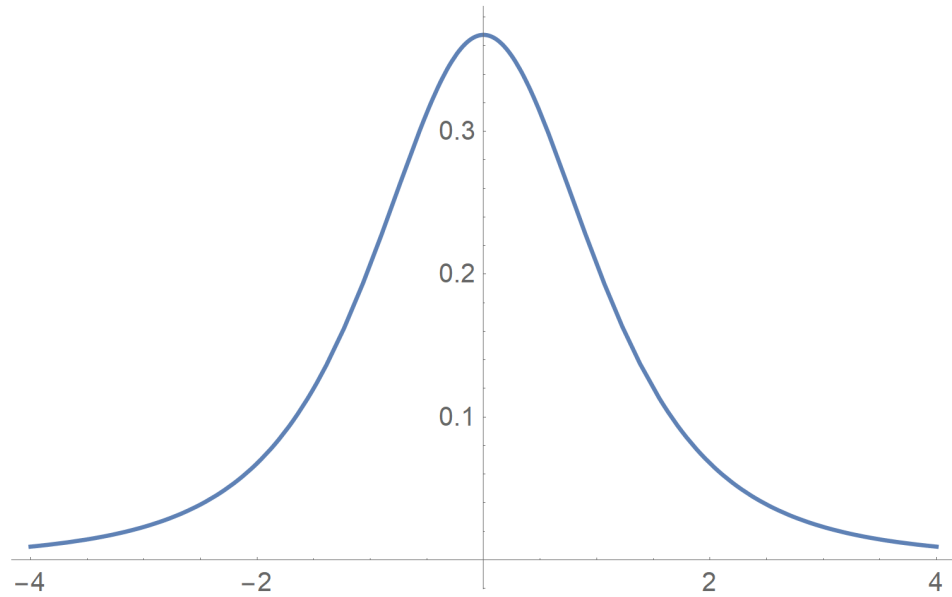


FIGURE 3.11 – Loi de Student avec $k = 1$ pour une variable centrée réduite

L'espérance de la loi de Student est :

$$\mathbb{E}(X) = 0 \quad (3.208)$$

La variance de la loi de Student est :

$$\mathbb{V}(X) = \frac{k}{k-2} \quad (3.209)$$

pseudonyme
du statisticien
William Seale
Gosset (1876-
1937)

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 45

Variable T de Student Soit une variable centrée et réduite U suivant une loi normale, la **variable aléatoire T de Student** est définie par :

$$T^2(v) = \frac{U^2}{\frac{\chi^2(v)}{v}} = F(1, v) \quad (3.210)$$

avec F la variable de Fisher, ou

$$T(v) = \frac{U}{\sqrt{\frac{\chi^2(v)}{v}}} \quad (3.211)$$

Propriétés de variable T de Student

- La variable $T(v)$ varie de $-\infty$ à $+\infty$.
- La loi de probabilité de la variable $T(v)$ dépend d'un seul paramètre, v , appelé le degré de liberté de la variable $\chi^2(v)$.
- Les moments de la variable T sont :

$$\mathbb{E}(T(v)) = 0 \quad (3.212)$$

et

$$\mathbb{V}(T(v)) = \frac{v}{v-2} \text{ si } v > 2 \quad (3.213)$$

- Le mode M_0 de la variable T est nul.

$$M_0 = 0 \quad (3.214)$$

- La densité de probabilité de variable T a une expression mathématique compliquée et peu utilisée.

$$f(t) = \frac{1}{\sqrt{v}} \frac{1}{B\left(\frac{1}{2}, \frac{v}{2}\right)} \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} \quad (3.215)$$

avec B la fonction eulérienne.

- La table de la variable T donne les fractiles d'ordre α de la loi de Student :

$$\alpha = \Pr(T(v) \leq T_\alpha(v)) = \int_{-\infty}^{t_\alpha(v)} f(t(v)) dt \quad (3.216)$$

Comme la fonction est symétrique, il suffit de prendre $\alpha > 50$.

- Selon les valeurs de v , la distribution T de Student possède des formes différentes. La courbe admet un axe de symétrie et ressemble à la « courbe en cloche ».

- Pour les grandes valeurs de v ($v > 100$), la loi de Student peut être remplacée par la loi normale réduite.
- **Relation entre la loi de Student et la loi de Cauchy.** Pour $v = 1$, la densité de la variable $T(1)$

$$f(t) = \frac{1}{\pi} \frac{1}{1+t^2} \quad (3.217)$$

est celle d'une variable suivant la loi de Cauchy.

3.2.13 Loi de Benford

Frank Benford (1883-1948) En 1938, F. Benford découvrit expérimentalement que la probabilité nulle qu'un nombre commence par le chiffre n , excepté 0, est :

$$\Pr(n) = \log_{10} \left(1 + \frac{1}{n} \right) \quad (3.218)$$

S'il existe une distribution de probabilité universelle $\Pr(n)$ sur de tels nombres, ils doivent être invariants sous un changement d'échelle tel que :

$$\Pr(kn) = f(k) \Pr(n) \quad (3.219)$$

Si $\int_{-\infty}^{+\infty} \Pr(n) dn = 1$, alors $\int_{-\infty}^{+\infty} \Pr(kn) dn = \frac{1}{k}$

La normalisation de la distribution donne :

$$f(k) = \frac{1}{k} \quad (3.220)$$

si on dérive $\Pr(kn) = f(k) \Pr(n)$ par rapport à k et que l'on pose $k = 1$, on obtient :

$$\frac{d\Pr(kn)}{dk} = \frac{d\left(\frac{1}{k}\right)}{dk} \Pr(n) = -\frac{1}{k^2} \Pr(n) \quad (3.221)$$

$$n\Pr'(kn) = -\frac{1}{k^2} \Pr(n) \quad (3.222)$$

En posant $k = 1$, on a :

$$n\Pr'(n) = -\Pr(n) \quad (3.223)$$

Cette équation différentielle a pour solution :

$$\Pr(n) = \frac{1}{n} \quad (3.224)$$

Remarque 1. Cette fonction n'est pas à proprement parler une fonction de distribution de probabilité, car elle diverge.

Remarque 2. Les lois de la physique et humaine imposent des limites.

La loi de Benford fut testée :

- sur la longueur des fleuves du globe ;
- sur la superficie des pays ;
- sur le résultat des élections.

3.2.14 Loi de Weibull

La fonction de répartition vaut :

$$\begin{cases} F(t) = 0 & \forall t < \gamma \\ F(t) = 1 - e^{(-\frac{t-\gamma}{\eta})^\beta} & \forall t \geq \gamma \end{cases} \quad (3.225)$$

Ernst Hjalmar
Waloddi Weibull
(1887-1979)

La densité de probabilité vaut :

$$\begin{cases} f(t) = 0 & \forall t < \gamma \\ f(t) = \frac{\beta}{\eta} \left(\frac{t-\gamma}{\eta}\right)^{\beta-1} e^{(-\frac{t-\gamma}{\eta})^\beta} & \forall t \geq \gamma \end{cases} \quad (3.226)$$

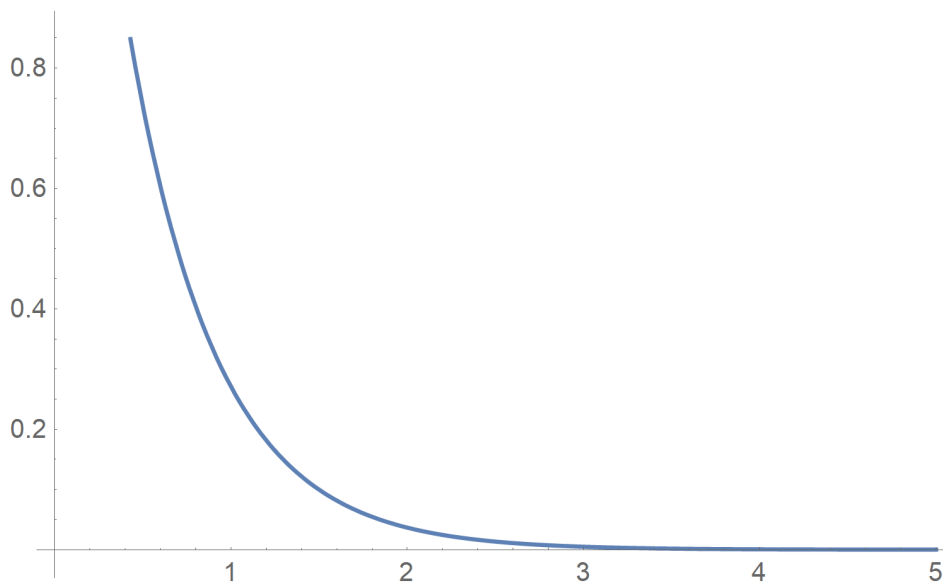


FIGURE 3.12 – Loi de Weibull avec $\beta = 1$ et $\eta = 0,5$

3.2.15 Loi de Gumbel ou loi des valeurs extrêmes

Emil Julius Gumbel (1891-1966)

La fonction de répartition est :

$$F(x) = e^{-e^{-\frac{x-x_0}{a}}} \quad (3.227)$$

avec $a \neq 0$

La densité de probabilité est :

$$f(x) = \frac{1}{a} e^{-\frac{x-x_0}{a}} e^{-e^{-\frac{x-x_0}{a}}} \quad (3.228)$$

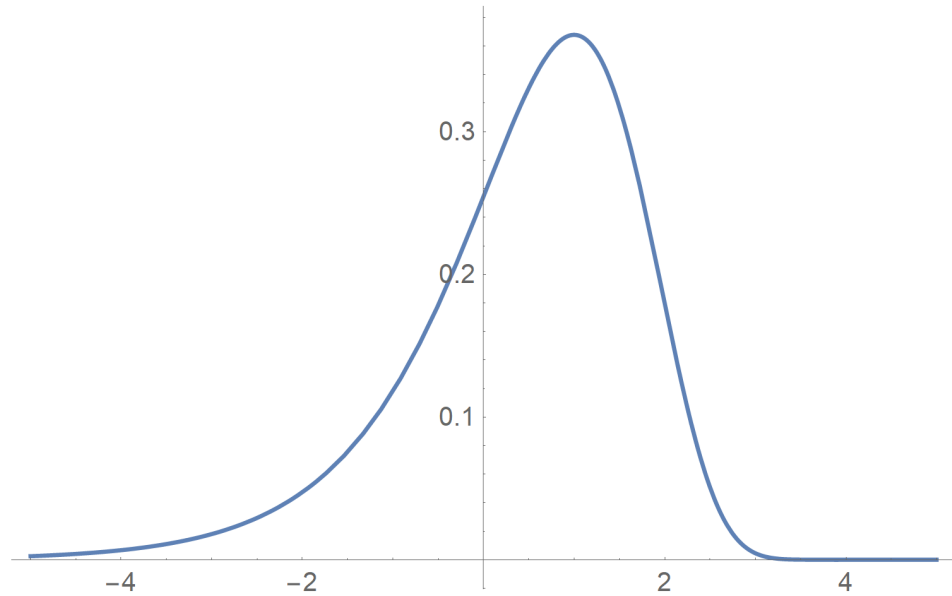


FIGURE 3.13 – Loi de Gumbel avec $x_0 = 1$ et $a = 1$

L'espérance mathématique de la loi de Gumbel vaut :

$$\mathbb{E}(X) = x_0 + \gamma a \approx x_0 + 0,577216a \quad (3.229)$$

où γ est la constante d'Euler-Mascheroni.

La variance mathématique de la loi de Gumbel vaut :

$$\mathbb{V}(X) = \frac{\pi^2}{6} a^2 \quad (3.230)$$

Les probabilités correspondant à la moyenne ont pour valeurs :

$$F(\mu) = F(x_0 + \gamma a) = \begin{cases} 0,570 & \text{si } a > 0 \\ 0,430 & \text{si } a < 0 \end{cases} \quad (3.231)$$

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 49

La médiane vaut :

$$F(M_e) = \frac{1}{2} \Rightarrow M_e = x_0 + 0,366513a \quad (3.232)$$

Le mode vaut :

$$F(x_0) = \begin{cases} 0,368 & \text{si } a > 0 \\ 0,632 & \text{si } a < 0 \end{cases} \quad (3.233)$$

Remarque. On peut réécrire toutes les formules précédentes en posant une variable centrée réduite : $u = \frac{x-x_0}{a}$.

$$\begin{aligned} F(u) &= e^{-e^{-u}} \\ f(u) &= e^{-u} e^{-e^{-u}} \\ \mathbb{E}(X) &= \gamma \approx 0,577216 \\ \mathbb{V}(X) &= \frac{\pi^2}{6} \end{aligned} \quad (3.234)$$

3.2.16 Loi de Fréchet

La fonction de répartition vaut :

$$F(x) = e^{-e^{-\frac{\ln(x-x_0)}{a}}} \quad (3.235)$$

La densité de probabilité vaut :

$$f(x) = \frac{1}{ax} e^{-\frac{\ln(x-x_0)}{a}} e^{-e^{-\frac{\ln(x-x_0)}{a}}} \quad (3.236)$$

René Maurice
Fréchet (1878-
1973)

3.2.17 Loi de Pareto

La fonction de répartition vaut :

$$\Pr(X \leq x) = 1 - C \left(\frac{x}{x_0} \right)^{-\alpha} = 1 - C \left(\frac{x_0}{x} \right)^{\alpha} \quad (3.237)$$

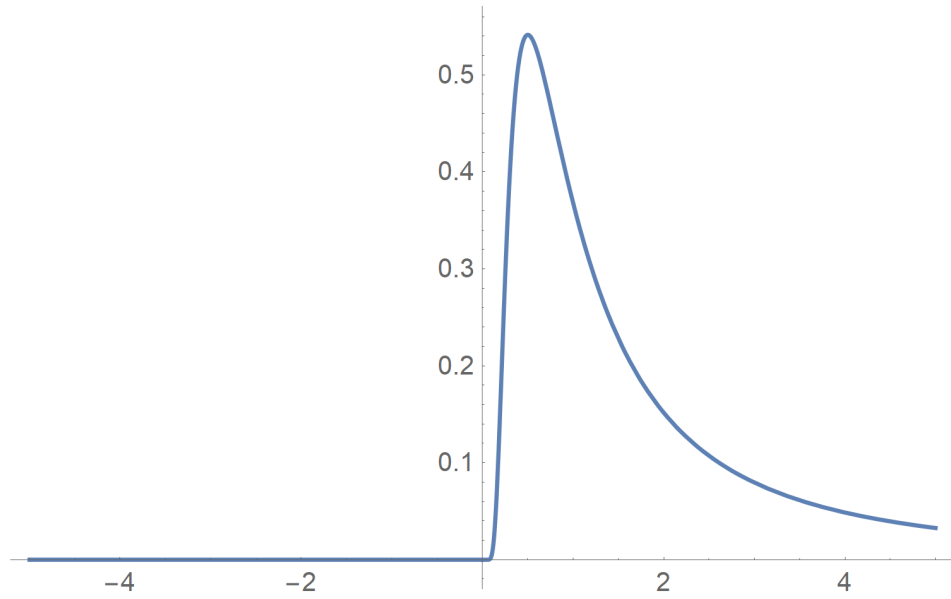
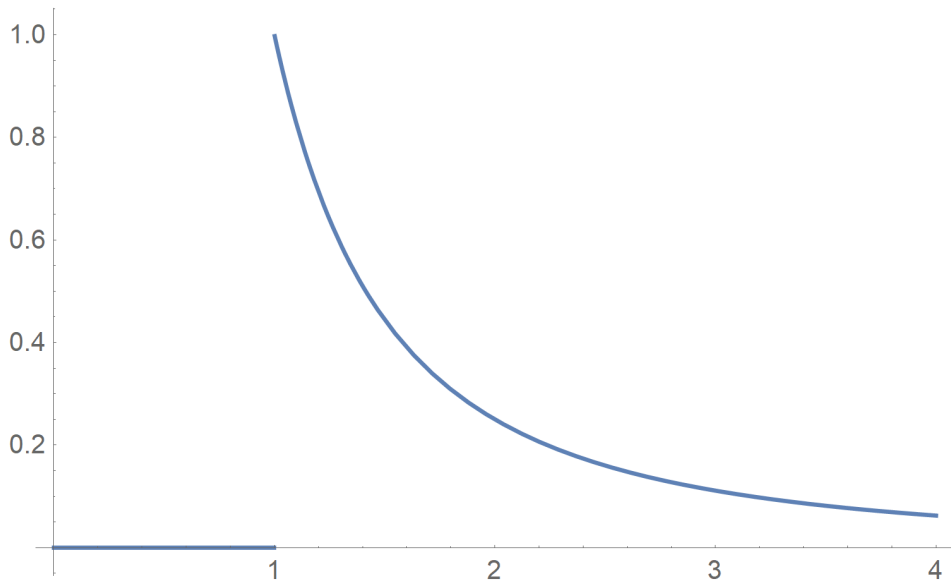
avec $x \geq x_0$ et C un paramètre d'échelle.

Pour $C = 1$, la densité de probabilité vaut :

$$\begin{cases} \Pr(X = x) = \frac{d}{dx} \left[1 - \left(\frac{x_0}{x} \right)^{\alpha} \right] = \alpha \frac{x_0^{\alpha}}{x^{\alpha+1}} & \text{avec } x \geq x_0 \\ \Pr(X = x) = 0 & \text{avec } x < x_0 \end{cases} \quad (3.238)$$

La distribution de Pareto est définie par deux paramètres x_0 et α . La constante α est nommée **index de Pareto**.

Vilfredo Pareto
(1848-1923)

FIGURE 3.14 – Loi de Fréchet avec $x_0 = 0$ et $a = 1$ FIGURE 3.15 – Loi de Pareto avec $x_0 = 1$ et $\alpha = 1$

Propriété importante. Soit a une constante.

$$\begin{aligned}
 f(ax) &= \alpha x_0^\alpha (ax)^{-\alpha-1} \\
 f(ax) &= a^{-\alpha-1} \alpha x_0^\alpha x^{-\alpha-1} \\
 f(ax) &= a^{-\alpha-1} \alpha \frac{x_0^\alpha}{x^{\alpha+1}} \\
 f(ax) &= a^{-\alpha-1} f(x) \propto f(x)
 \end{aligned} \tag{3.239}$$

3.2. LES DISTRIBUTIONS STATISTIQUES DES VARIABLES CONTINUES 51

La loi de Pareto vérifie la définition d'une fonction de distribution.

$$\begin{aligned}\int_0^{+\infty} f(x) dx &= \left[1 - \left(\frac{x_0}{x}\right)^\alpha\right]_{x_0}^{+\infty} \\ \int_0^{+\infty} f(x) dx &= \left[1 - \left(\frac{x_0}{+\infty}\right)^\alpha\right] - \left[1 - \left(\frac{x_0}{x_0}\right)^\alpha\right] \\ \int_0^{+\infty} f(x) dx &= 1 - 0 - (1 - 1^\alpha) \\ \int_0^{+\infty} f(x) dx &= 1\end{aligned}\quad (3.240)$$

Les caractéristiques de la loi de Pareto sont peu nombreuses.

- Il s'agit d'une loi extrêmement dissymétrique.
- Lorsque $\alpha \leq 1$, la moyenne n'existe pas. Dit autrement, les valeurs extrêmes ont une probabilité plus élevée de se réaliser par rapport à la loi normale. L'espérance de la fonction de Pareto est :

$$\mathbb{E}(X) = \frac{\alpha x_0}{\alpha - 1} \quad (3.241)$$

Ce résultat s'obtient en intégrant directement la définition de l'espérance⁶.

- Lorsque $\alpha \leq 2$, la variance n'existe pas⁷. La variance de la fonction de Pareto est :

$$\mathbb{V}(X) = \frac{\alpha x_0^2}{(\alpha - 1)^2 (\alpha - 2)} \quad (3.244)$$

Remarque. Lorsque $\alpha \rightarrow +\infty$, la distribution de Pareto s'approche de $\delta(x - x_0)$ avec δ la fonction de Dirac.

6.

$$\begin{aligned}\mathbb{E}(X) &= \int_{x_0}^{+\infty} x f(x) dx \\ \mathbb{E}(X) &= \int_{x_0}^{+\infty} x \alpha \frac{x_0^\alpha}{x^{\alpha+1}} dx \\ \mathbb{E}(X) &= \alpha x_0^\alpha \int_{x_0}^{+\infty} \frac{1}{x^\alpha} dx \\ \mathbb{E}(X) &= \left[-\frac{\alpha x_0^\alpha}{\alpha-1} \frac{1}{x^{\alpha-1}} \right]_{x_0}^{+\infty} \\ \mathbb{E}(X) &= \frac{\alpha x_0}{\alpha-1}\end{aligned}\quad (3.242)$$

7.

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \\ \mathbb{E}(X^2) &= \int_{x_0}^{+\infty} x^2 f(x) dx \\ \mathbb{E}(X^2) &= \alpha x_0^\alpha \int_{x_0}^{+\infty} \frac{1}{x^{\alpha-1}} dx \\ \mathbb{E}(X^2) &= \alpha x_0^\alpha \left[-\frac{1}{\alpha-2} \frac{1}{x^{\alpha-2}} \right]_{x_0}^{+\infty} \\ \mathbb{E}(X^2) &= \frac{\alpha x_0^2}{\alpha-2} \\ \mathbb{V}(X) &= \frac{\alpha x_0^2}{\alpha-2} - \left(\frac{\alpha x_0}{\alpha-1} \right)^2\end{aligned}\quad (3.243)$$

Il existe deux lois de Pareto :

1. la loi forte de Vilfredo Pareto (1897) ;
2. la loi faible de Benoît Mandelbrot (1963).

Soit x_s un seuil au-delà duquel on calcule l'espérance de la quantité examinée, et $\mathbb{E}(Y)$ l'espérance au-delà de ce seuil x_s tel qu'il soit proportionnel au seuil choisi :

$$\mathbb{E}(Y) = ax_s + b \quad (3.245)$$

Dans ce cadre, la moyenne **conditionnelle** au-delà du seuil x_s est un multiple de ce seuil à une constante près, c'est-à-dire une fonction linéaire de ce seuil. Si elle existe et elle est vérifiée, on dit que la distribution de probabilité obéit à une **forme de Pareto généralisée**. De fait, la loi de Pareto définit comme $\Pr(X \leq x) = 1 - \left(\frac{x_0}{x}\right)^\alpha$ est appelée **loi de Pareto de type I**.

La loi de Pareto de type II fait intervenir un paramètre supplémentaire μ tel que :

$$\Pr(X \leq x) = 1 + \left(\frac{x_0}{x - \mu}\right)^\alpha \quad (3.246)$$

La loi de Pareto de type III fait intervenir deux paramètres supplémentaires μ et γ tels que :

$$\Pr(X \leq x) = \left[1 + \left(\frac{x - \mu}{x_0}\right)^{\frac{1}{\gamma}}\right]^{-1} \quad (3.247)$$

La loi de Pareto de type IV généralise le type III tel que :

$$\Pr(X \leq x) = \left[1 + \left(\frac{x - \mu}{x_0}\right)^{\frac{1}{\gamma}}\right]^{-\alpha} \quad (3.248)$$

Une autre façon d'obtenir une loi de Pareto est de reprendre les travaux de B. Mandelbrot [Mandelbrot, 1997]. Il rapproche cette distribution de la notion de fractale. Pour lui, il s'agit d'une **distribution scalante** définie telle que :

$$\Pr(U > u) = u^{-\alpha} \quad (3.249)$$

avec $\alpha = \frac{A}{A-1}$. Ce type de lois a été utilisé dès le début du XX^e siècle [Lotka, 1925] [Korčák, 1940].

Lorsque tout ce que l'on sait de U est que $U \geq h$, on obtient la probabilité conditionnelle suivante :

$$\Pr(U \geq u \mid U \geq h) = \left(\frac{U}{h}\right)^\alpha = h^\alpha U^{-\alpha} \quad (3.250)$$

On pose $C = h^\alpha$.

En conséquence,

$$\Pr(U - h) = \frac{h}{\alpha - 1} \quad (3.251)$$

« La distribution scalante n'a pas d'échelle propre, mais emprunte toujours celle de l'observateur » [Mandelbrot, 1997, p. 79].

Bibliographie

- [Fréchet, 1941] FRÉCHET, M. (1941). Sur la loi de répartition de certaines grandeurs géographiques. Journal de la société statistique de Paris, 82:114–122.
- [Gnedenko et Kolmogorov, 1954] GNEDENKO, B. V. et KOLMOGOROV, A. N. (1954). Limit Distributions for Sums of Independent Random Variable. Addisson-Wesley, Cambridge.
- [Korčák, 1940] KORČÁK, J. (1940). Deux types fondamentaux de distribution statistique. Bulletin de l’Institut international de statistique, 30(3):295–299. Rapports et communications présentés à la XXIVe session de l’Institut international de statistique, Prague, 1938, 2e partie.
- [Lotka, 1925] LOTKA, A. J. (1925). Frequency distribution of scientific productivity. Journal of the Washington Academy of Science, 16:317–323.
- [Mandelbrot, 1997] MANDELBROT, B. (1997). Fractales, hasard et finances (1959-1997). Flammarion, Paris.