

Statistique descriptive et Probabilités

10 mai 2019

Table des matières

1	Statistique descriptive	2
1.1	Vocabulaire de la statistique	2
1.1.1	Statistique	2
1.1.2	Statistique descriptive	3
1.1.3	Statistique inférentielle	3
1.1.4	Population (ou population statistique)	3
1.1.5	Individu (ou unité statistique)	3
1.1.6	Échantillon	3
1.1.7	Enquête (statistique)	4
1.1.8	Recensement	4
1.1.9	Sondage	4
1.1.10	Caractères (ou variables) statistiques	4
1.1.11	Séries statistiques	5
1.1.12	Effectifs (cumulés), fréquences (cumulées)	5
1.2	Représentations graphiques	6
1.2.1	Diagramme en bâtons	6
1.2.2	Histogramme	6
1.2.3	Polygone des effectifs ou des fréquences	6
1.2.4	Courbes de fréquences cumulées	7
1.3	Caractéristiques de tendance centrale (ou de position)	7
1.3.1	Le mode	8
1.3.2	La médiane	8
1.3.3	La moyenne	8
1.3.4	Les quantiles	9
1.4	Paramètres de dispersion	10
1.4.1	Étendue	10
1.4.2	Variance et écart-type	10
1.4.3	Moments et moments centrés.	11
1.4.4	Coefficient de dispersion ou coefficient de variation	11

1.5	Paramètres de forme	11
1.5.1	Coefficient d'asymétrie de Fisher	11
1.5.2	Coefficient d'asymétrie de Pearson	12
1.5.3	Coefficient d'asymétrie de Yule	12
1.5.4	Coefficient d'aplatissement de Pearson	13
2	Introduction à la théorie des probabilités	14
2.1	Introduction	14
2.2	Analyse combinatoire	15
2.2.1	Le langage des ensembles	15
2.2.2	Dénombrement de listes	16
2.2.3	Dénombrement d'arrangements	17
2.2.4	Combinaisons	17
2.3	Mesure de probabilité et espace probabilisé	18
2.3.1	Notion de tribu	18
2.3.2	Mesures	19
2.4	Calcul de probabilités	19
2.4.1	Vocabulaire des probabilités	19
2.4.2	Définition de la probabilité et propriétés	20
2.4.3	Equiprobabilité	21
2.5	Probabilité conditionnelle- Indépendance	22
2.5.1	Probabilité conditionnelle	22
2.5.2	Indépendance	23
3	Variables aléatoires réelles	25
3.1	Tribu des boréliens	25
3.1.1	Ouverts et fermés de \mathbb{R} , tribu des boréliens	25
3.2	Notion de Variable aléatoire réelle	26
3.2.1	Loi de probabilité d'une variable aléatoire	26
3.2.2	Fonction de répartition d'une variable aléatoire	26
3.3	Variable aléatoire réelle discrète	27
3.3.1	Moments d'une variable aléatoire	27
3.4	Variable aléatoire réelle continue	29
3.4.1	Densité de probabilité et fonction de répartition	30
3.4.2	Moments d'une variable aléatoire	30
3.5	Fonction génératrice des moments et fonction caractéristique	32
3.5.1	Fonction génératrice des moments	32
3.5.2	Fonction caractéristique	32
3.6	Inégalités intéressantes	33
3.6.1	Inégalité de Markov	33
3.6.2	Inégalité de Bienaymé-Tchebychev	33

4 Lois classiques de probabilité	35
4.1 Lois de probabilité discrètes	35
4.1.1 Loi de Dirac	36
4.1.2 Loi uniforme discrète	36
4.1.3 Loi de Bernoulli	36
4.1.4 Loi binomiale	37
4.1.5 Loi géométrique	38
4.1.6 Loi de Pascal	38
4.1.7 Loi binomiale négative	38
4.1.8 Loi hypergéométrique	39
4.1.9 Loi de Poisson	39
4.2 Lois de probabilité continues	41
4.2.1 Loi uniforme	41
4.2.2 La loi exponentielle	41
4.2.3 Loi normale	42
4.2.4 La loi gamma	44
Bibliographie	45

STATISTIQUE DESCRIPTIVE

Nous donnons ici une (brève) présentation de ce qu'est la première étape de l'analyse des données issues d'une enquête : la description. Une enquête commence habituellement s'il y a (au niveau d'un organisme) un besoin d'informations et s'il n'y a pas de données ou si elles sont insuffisantes. On organise alors méthodiquement une activité de collecte de données sur des caractéristiques jugées d'intérêt au niveau des unités d'une population (ou d'une partie de cette dernière). S'ensuit alors une compilation et une présentation (description) des données recueillies sous une forme récapitulative exploitable (tableaux, graphiques, nombres ou indices). L'analyse des données rend compte de ce résumé et présente l'interprétation de leur signification pour obtenir des réponses claires aux questions qui ont motivé l'enquête.

1.1 Vocabulaire de la statistique

On précise ici un certain nombre de termes statistiques très courants qui seront régulièrement utilisés par la suite et qu'il convient de bien connaître.

1.1.1 Statistique

La statistique désigne à la fois un ensemble de données et l'ensemble des activités consistant à collecter ces données, à les traiter et à les interpréter.

1.1.2 Statistique descriptive

Le traitement des données, pour en dégager un certain nombre de renseignements qualitatifs ou quantitatifs à des fins de comparaison, s'appelle la statistique descriptive.

1.1.3 Statistique inférentielle

Un autre but de la statistique consiste à extrapoler à partir d'un échantillon de la population à étudier, le comportement de la population dans son ensemble (sondages, contrôle de qualité comportant un test destructif...). C'est la statistique inductive ou encore appelée statistique inférentielle.

1.1.4 Population (ou population statistique)

C'est l'ensemble concerné par une étude statistique. On parle aussi de champ de l'étude. Si l'on s'intéresse par exemple aux notes d'un groupe d'étudiants, ce groupe constitue la population.

1.1.5 Individu (ou unité statistique)

Les éléments qui constituent la population sont appelés les individus ou encore les unités statistiques.

1.1.6 Échantillon

Dans une étude statistique, il est fréquent que l'on n'observe pas la population tout entière. Les observations du phénomène considéré sont donc réalisées sur une partie restreinte de la population, appelée échantillon. On appelle donc échantillon le sous-ensemble de la population sur lequel sont effectivement réalisées les observations. La taille de l'échantillon est le cardinal de l'échantillon, autrement dit c'est le nombre d'individus qu'il contient (échantillon de taille 800, de taille 1000...). En général, on note n la taille de l'échantillon considéré.

1.1.7 Enquête (statistique)

C'est l'opération consistant à observer (ou mesurer, ou questionner,...) l'ensemble des individus d'un échantillon (ou, éventuellement, de la population complète).

1.1.8 Recensement

C'est une enquête dans laquelle l'échantillon observé est en fait la population tout entière (on parle aussi d'enquête exhaustive).

1.1.9 Sondage

C'est, au contraire, une enquête dans laquelle l'échantillon observé est un sous-ensemble strict de la population (on parle, dans ce cas, d'enquête non exhaustive).

1.1.10 Caractères (ou variables) statistiques

Les caractéristiques étudiées sur les individus d'une population sont appelées les caractères. Un caractère ou variable statistique est une application χ d'un ensemble Ω de cardinal fini N (la population) dans un ensemble C (l'ensemble des valeurs possibles du caractère), qui associe à chaque individu ω de Ω la valeur $\chi(\omega)$ que prend ce caractère sur l'individu ω . Les valeurs possibles de la variable, sont appelées modalités. L'ensemble des valeurs possibles ou des modalités est appelé le domaine de la variable. Le caractère statistique peut être quantitatif ou qualitatif.

- Les variables quantitatives expriment une quantité (taille, poids, salaire, température,...). Elles sont donc mesurables et sont numériques. Il y en a deux types : les variables quantitatives discrètes et les variables quantitatives continues. Une variable quantitative est discrète si son domaine est un ensemble discret (au plus dénombrable). Une variable quantitative continue peut prendre toutes les valeurs dans un certain intervalle de \mathbb{R} .

- Un caractère est dit qualitatif lorsqu'il est lié à une observation ne faisant pas l'objet d'une mesure (profession, sexe,...). Les variables qualitatives dont

les modalités ne peuvent être ordonnées sont dites nominales. Les variables qualitatives dont les modalités peuvent être ordonnées sont dites ordinales.

Remarque 1.1. 1. *Les caractères qualitatifs peuvent toujours être transformés en quantitatifs par codage. C'est ce qui se fait le plus généralement. Mais un tel codage est purement conventionnel et n'a pas vraiment un sens quantitatif.*

Exemple : Nous ne pouvons pas calculer le sexe moyen.

2. *Certains caractères qualitatifs s'expriment à l'aide de nombres.*

Exemple : Un numéro de téléphone.

Mais ils n'ont pas non plus de sens quantitatif.

Exemple : Calculer un numéro de téléphone moyen n'est pas pertinent.

1.1.11 Séries statistiques

On appelle série statistique la suite des valeurs prises par une variable statistique sur les unités d'observation. Si x est une telle variable statistique et N est l'effectif (la taille) de l'échantillon (population) considéré(e), les observations sont notées : $x_1, x_2, \dots, x_i, \dots, x_N$.

1.1.12 Effectifs (cumulés), fréquences (cumulées)

Si X est un caractère quantitatif (simple ou univarié), la suite finie $X(\Omega) = (X_1, \dots, X_N)$ des valeurs atteintes par le caractère est un ensemble fini $\{x_1, \dots, x_p\}$. Nous supposons que ces valeurs sont ordonnées : $x_1 < \dots < x_p$. Le fait que telle valeur soit relative à tel individu est un renseignement qui n'intéresse pas le statisticien. Seul l'ensemble des valeurs atteintes et le nombre de fois que chacune d'elle est atteinte sont utiles.

Soit une série statistique donnée.

1. L'effectif de la modalité x_i est le nombre n_i de fois que la valeur x_i est prise.
2. L'effectif cumulé croissant (resp. décroissant) associé à une valeur x_i de la variable X est le nombre d'individus dont la mesure est inférieure ou égale (resp. supérieure ou égale) à x_i .
3. La fréquence (relative), associée à la modalité x_i de la variable statistique

X est le nombre

$$f_i = \frac{n_i}{N}.$$

4. La fréquence cumulée croissante (resp. décroissante) permet de calculer le pourcentage d'individus de la population pour lesquelles la variable statistique X prend une valeur inférieure ou égale (resp. supérieure ou égale) à la modalité x_i

1.2 Représentations graphiques

1.2.1 Diagramme en bâtons

Le diagramme en bâtons des effectifs (resp. des fréquences) d'une distribution statistique discrète est constitué d'une suite de segments verticaux d'abscisses x_i dont la longueur est proportionnelle à l'effectif (resp. la fréquence) de x_i .

1.2.2 Histogramme

L'histogramme est la représentation graphique d'une distribution statistique groupée. Un histogramme est constitué de rectangles juxtaposés dont la base correspond à l'amplitude de chaque classe et dont la surface est proportionnelle à la fréquence ou à l'effectif de cette classe.

L'histogramme est un outil statistique facile à utiliser, donnant rapidement une image du comportement d'un procédé industriel et l'allure globale de la distribution; il montre l'étalement des données et apporte ainsi des renseignements sur la dispersion et sur les valeurs extrêmes; il permet de déceler, éventuellement, des valeurs aberrantes.

1.2.3 Polygone des effectifs ou des fréquences

Le polygone des effectifs (resp. des fréquences) d'une distribution statistique discrète est obtenu à partir du diagramme en bâtons des effectifs (resp. des fréquences) en joignant par un segment les sommets des bâtons. Dans le d'une distribution statistique groupée, il est obtenu en joignant, par des

segments de droite, les milieux des côtés supérieurs de chaque rectangle de l'histogramme. Pour fermer ce polygone, on ajoute à chaque extrémité une classe de fréquence nulle. De cette façon, l'aire totale sous le polygone est égale à l'aire totale de l'histogramme, de façon à respecter la proportionnalité des aires aux effectifs ou aux fréquences qui caractérise les histogrammes. Pour obtenir cette conservation de l'aire, il est impératif que toutes les classes soient de largeur égale (si ce n'est pas le cas, il faut subdiviser artificiellement les classes les plus larges pour obtenir l'égalité de toutes les largeurs).

1.2.4 Courbes de fréquences cumulées

Courbe cumulative croissante

On joint les points ayant pour abscisses la limite supérieure des classes et pour ordonnées les fréquences cumulées croissantes correspondant à la classe considérée (pour le premier point, on porte la valeur 0). Elle donne la fréquence des observations inférieures à une valeur quelconque de la série.

Courbe cumulative décroissante

La construction de cette courbe est analogue à la précédente. Les points ont pour abscisses, les limites inférieures des classes et pour ordonnées, les fréquences cumulées décroissantes (pour le dernier point, la valeur est 0). Elle donne la fréquence des observations supérieures à une valeur quelconque de la série.

1.3 Caractéristiques de tendance centrale (ou de position)

Elles donnent une idée de l'ordre de grandeur des valeurs constituant la série ainsi que la position où semblent se concentrer les valeurs de cette série. Les principales caractéristiques de tendance centrale sont la moyenne arithmétique, la médiane, le mode et les quantiles.

1.3.1 Le mode

Le mode est la valeur la plus fréquente dans l'échantillon. Il est souvent noté M_0 et parfois appelé valeur dominante. Dans une série statistique, on peut avoir plusieurs modes. Soit une série statistique présentant un regroupement en classes. On appelle classe modale de cette série toute classe d'effectif (ou de fréquence) maximal(e). On convient qu'un mode d'une série statistique groupée en classes est le centre d'une classe modale.

1.3.2 La médiane

La médiane est la valeur M_e , observée ou possible, dans la série des données classées par ordre croissant (ou décroissant) qui partage cette série en deux parties comprenant exactement le même nombre de données de part et d'autre de M_e .

La médiane d'une série statistique est la valeur, souvent notée M_e , qui dans la série ordonnée (croissante ou décroissante), partage en deux parties égales l'effectif total de la série. Soit une série statistique discrète (ordonnée) x_1, \dots, x_N d'effectif total N . En pratique, on distingue deux cas.

- Si l'effectif N est impair, $M_e = x_{\frac{N+1}{2}}$
- Si l'effectif N est pair, $M_e = \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2}$

Dans le cas d'une distribution en classes, la médiane M_e est l'abscisse du point de la courbe de fréquences cumulées (croissantes ou décroissantes) d'ordonnée 0,5 (ou 50%). Toutefois, si on identifie la classe médiane (celle correspondant à l'effectif cumulé $\frac{N}{2}$), on peut procéder à une interpolation à l'intérieur de cette classe médiane.

1.3.3 La moyenne

La moyenne, \bar{x} , d'une distribution statistique discrète est le nombre réel défini par

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N n_i x_i$$

Pour une série statistique groupée en classes $[a_i, b_i]$, la moyenne se calcule par

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i c_i$$

où k est le nombre de classes et c_i le centre de la classe $[a_i, b_i]$.

Remarque 1.2. Soient x une variable statistique, a et b des constantes réelles. Si on définit une nouvelle variable $z = ax + b$, alors la nouvelle variable z a pour moyenne $\bar{z} = a\bar{x} + b$.

1.3.4 Les quantiles

Les quantiles sont des caractéristiques de position partageant la série statistique ordonnée en k parties égales. Pour $k = 4$, les quantiles, appelés quartiles, sont trois nombres Q_1, Q_2, Q_3 tels que :

- 25% des valeurs prises par la série sont inférieures à Q_1 ,
 - 25% des valeurs prises par la série sont supérieures à Q_3 ,
 - Q_2 est la médiane M_e ,
 - $[Q_1, Q_3]$ est l'intervalle interquartile, il contient 50% des valeurs de la série.
- Pour $k = 10$, les quantiles sont appelés déciles, il y a neuf déciles D_1, D_2, \dots, D_9 . 10% des valeurs de la série sont inférieures à D_1 .
- Pour $k = 100$, les quantiles sont appelés centiles, il y a 99 centiles, chacun correspondant à 1% de la population.

Application

Le diagramme en boîte à moustaches ou box-plot (Tukey) permet de représenter schématiquement les principales caractéristiques d'une distribution en utilisant les quartiles. La partie centrale de la distribution est représentée par une boîte de largeur arbitraire et de longueur la distance interquartile, la médiane est tracée à l'intérieur. La boîte rectangle est complétée par des moustaches correspondant aux valeurs suivantes :

- valeur supérieure : $Q_3 + 1,5(Q_3 - Q_1)$,
- valeur inférieure : $Q_1 - 1,5(Q_3 - Q_1)$.

Les valeurs extérieures «aux moustaches» sont représentées par des étoiles.

1.4 Paramètres de dispersion

Ces caractéristiques quantifient les fluctuations des valeurs observées autour de la valeur centrale et permettent d'apprécier l'étalement de la série. Les principales sont : l'écart-type ou son carré appelé variance, le coefficient de variation et l'étendue.

1.4.1 Étendue

L'étendue, notée E , est la différence entre la plus grande et la plus petite des valeurs prises, donc $E = x_{max} - x_{min}$.

1.4.2 Variance et écart-type

La variance d'une série statistique quantitative est la moyenne arithmétique des carrés des écarts à la moyenne arithmétique. On la note souvent σ^2 ou $V(x)$.

La variance d'une distribution statistique discrète est définie par

$$V(x) = \frac{1}{N} \sum_{i=1}^N n_i (x_i - \bar{x})^2 = \left(\frac{1}{N} \sum_{i=1}^N n_i x_i^2 \right) - \bar{x}^2$$

Pour une série statistique groupée en classes $[a_i, b_i]$, la variance se calcule par

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k n_i (c_i - \bar{x})^2 = \left(\frac{1}{N} \sum_{i=1}^k n_i c_i^2 \right) - \bar{x}^2.$$

où k est le nombre de classes et c_i le centre de la classe $[a_i, b_i]$.

L'écart-type noté σ_x est défini comme étant la racine carrée de la variance,

$$\sigma_x = \sqrt{V(x)}.$$

L'avantage de l'écart-type par rapport à la variance est que c'est un nombre qui s'exprime dans la même unité que les valeurs observées.

1.4.3 Moments et moments centrés.

Soit k un entier naturel non nul. On appelle moment d'ordre k de la variable statistique x , le nombre réel noté $m_k(x)$ défini par :

$$m_k(x) = \frac{1}{N} \sum_{i=1}^N n_i(x_i)^k.$$

De même on définit le moment centré d'ordre k de x , noté $\mu_k(x)$ défini par :

$$\mu_k(x) = \frac{1}{N} \sum_{i=1}^N n_i(x_i - \bar{x})^k$$

1.4.4 Coefficient de dispersion ou coefficient de variation

Il s'exprime, sous la forme d'un pourcentage, par l'expression suivante :

$$CV = \frac{\sigma_x}{\bar{x}}$$

Propriétés

- Le coefficient de variation ne dépend pas des unités choisies.
- Il permet d'apprécier la représentativité de la moyenne arithmétique \bar{x} par rapport à l'ensemble des données.
- Il permet d'apprécier l'homogénéité de la distribution, une valeur du coefficient de variation inférieure à 15% traduit une bonne homogénéité de la distribution.
- Il permet de comparer deux distributions, même si les données ne sont pas exprimées avec la même unité ou si les moyennes arithmétiques des deux séries sont très différentes. Pour deux distributions différentes, la plus homogène est celle qui a le plus petit coefficient de variation.

1.5 Paramètres de forme

1.5.1 Coefficient d'asymétrie de Fisher

Une distribution est symétrique lorsque ses valeurs sont réparties de manière égale de part et d'autre de la valeur centrale (le polygone de fréquence

présente un axe de symétrie vertical). Si la distribution est asymétrique il peut être intéressant de savoir si la courbe s'étale à gauche ou à droite. Le coefficient d'asymétrie de Fisher est le nombre réel noté γ et défini par :

$$\gamma = \frac{\mu_3}{\sigma^3}.$$

1. Si $\gamma < 0$, la distribution est dite étalée vers la gauche ou oblique à droite (biais négatif).
2. Si $\gamma > 0$, la distribution est dite étalée vers la droite ou oblique à gauche (biais positif).
3. $\gamma = 0$, la distribution est symétrique.

1.5.2 Coefficient asymétrie de Pearson

Il est défini par

$$A_p = \frac{\bar{x} - M_0}{\sigma_x}$$

où M_0 est le mode.

1. Si $A_p < 0$, la distribution est dite étalée vers la gauche ou oblique à droite (biais négatif).
2. Si $A_p > 0$, la distribution est dite étalée vers la droite ou oblique à gauche (biais positif).
3. Si $A_p = 0$, la distribution est symétrique.

1.5.3 Coefficient d'asymétrie de Yule

Il est défini par

$$A_Y = \frac{Q_3 + Q_1 - 2M_e}{Q_3 - Q_1}$$

1. Si $A_Y < 0$, la distribution est dite étalée vers la gauche ou oblique à droite (biais négatif).
2. Si $A_Y > 0$, la distribution est dite étalée vers la droite ou oblique à gauche (biais positif).
3. Si $A_Y = 0$, la distribution est symétrique.

1.5.4 Coefficient d'aplatissement de Pearson

Il est défini par

$$\beta = \frac{\mu_4}{\sigma^4}.$$

1. Si $\beta = 3$, la distribution est normale (courbe en cloche de Gauss ou dite standard).
2. Si $\beta < 3$, la distribution est plus aplatie que la distribution normale. Le sommet de la courbe est moins pointue que celle de la cloche de Gauss (courbe standard).
3. Si $\beta > 3$, la distribution est moins aplatie que la distribution normale. Le sommet de la courbe est plus pointue que celle de la cloche de Gauss (courbe standard)

INTRODUCTION À LA THÉORIE DES PROBABILITÉS

2.1 Introduction

Dans des domaines très différents comme le domaine scientifique, sociologique, médical, les sciences humaines..., on s'intéresse à de nombreux phénomènes dans lesquels apparaît souvent l'effet du hasard. Ces phénomènes sont caractérisés par le fait que les résultats des observations varient d'une expérience à l'autre. Une expérience est appelée aléatoire s'il est impossible de prévoir son résultat et si, répétée dans des conditions identiques, elle peut donner, ou aurait pu donner, si l'expérience est unique, des résultats différents. En général, les résultats obtenus varient dans un certain domaine, certains résultats apparaissant plus fréquemment que d'autres. Ils peuvent être visualisés par des diagrammes, des histogrammes, des courbes cumulatives de fréquences, etc., et être caractérisés par quelques valeurs numériques telles que la moyenne arithmétique, la médiane, le mode, la variance.... Le mot probabilité est passé rapidement dans le langage courant bien que la théorie des probabilités soit une branche relativement récente des théories mathématiques. Le concept des probabilités semblait être connu des Grecs et des Égyptiens. Cependant, ce n'est que vers le milieu du *XVII^e* siècle que l'on peut situer le début de cette théorie. D'abord limitée à l'étude des jeux de hasard (jeux de pile ou face, roulettes, jeux de cartes...), elle s'est rapidement étendue à tous les domaines de la Science, en Physique, en Informatique, en Économie, en Génétique,...

Les premiers résultats mathématiques furent introduits par Pascal et Fermat au milieu du *XVII^e* siècle. Puis, apparaissent, à la fin du *XVII^e* siècle,

le nom de Huyghens et surtout au $XVIII^e$ siècle, les noms de Bernoulli, De Moivre, Bayes, Laplace, (le tome VII de ses œuvres s'intitule Calcul des Probabilités), Gauss et au XX^e siècle, Poincaré, Borel, Fréchet, Lévy, Kolmogorov,...

Au début du XX^e siècle, tous les concepts probabilistes fondamentaux dérivent des concepts de théorie de la mesure (tribu, espace probabilisable, mesure de probabilité...). Bien souvent (sans que cela soit absolument nécessaire) le calcul élémentaire de probabilité fait appel au langage des ensembles et à leur dénombrement.

2.2 Analyse combinatoire

L'analyse combinatoire fournit des méthodes de dénombrement particulièrement utiles en théorie des probabilités.

2.2.1 Le langage des ensembles

- Définition 2.1.**
1. *Un ensemble E est fini s'il est vide ou s'il existe un entier naturel n et une bijection de $\{1, \dots, n\}$ dans E . Si $E \neq \emptyset$, n est appelé cardinal de E , noté $\text{Card}(E)$ ou $|E|$. Par convention $\text{Card}(\emptyset) = 0$.*
 2. *Un ensemble E est dénombrable s'il existe une application bijective de E dans l'ensemble des entiers naturels \mathbb{N} .
Un ensemble dénombrable est donc infini.*
 3. *Un ensemble E est au plus dénombrable s'il existe une application bijective de E dans une partie de \mathbb{N} .*

Opérations sur les ensembles

1. **Réunion.** La réunion des deux ensembles A et B est notée $A \cup B$ et est définie par :

$$x \in A \cup B \Leftrightarrow (x \in A \text{ ou } x \in B).$$

2. **Intersection.** L'intersection des deux ensembles A et B est notée $A \cap B$

et est définie par :

$$x \in A \cap B \Leftrightarrow (x \in A \text{ et } x \in B).$$

Deux ensembles A et B sont disjoints si $A \cap B = \emptyset$.

3. **Partition.** Une famille $(A_i)_{i \in I}$ de parties d'un ensemble Ω est une partition de Ω si

$$\text{i) } \bigcup_{i \in I} A_i = \Omega$$

$$\text{ii) } \forall (i, j) \in I^2, (i \neq j \Rightarrow A_i \cap A_j = \emptyset).$$

4. **Complémentaire.** Soit A une partie d'un ensemble E , le complémentaire de A dans E est noté A^c et est défini par :

$$x \in A^c \Leftrightarrow (x \in E \text{ et } x \notin A).$$

5. **Différence.** Soit A et B deux parties de E , nous notons $A \setminus B$ l'ensemble défini par :

$$x \in A \setminus B \Leftrightarrow (x \in A \text{ et } x \notin B).$$

Par conséquent, nous avons l'égalité $A \setminus B = A \cap B^c$.

6. **Différence symétrique.** Soit A et B deux parties de E , nous notons $A \Delta B$ l'ensemble défini par :

$$x \in A \Delta B \Leftrightarrow (x \in A \cup B \text{ et } x \notin A \cap B).$$

Par conséquent, nous avons l'égalité $A \Delta B = (A \cup B) \setminus (A \cap B)$.

7. **Produit cartésien.** Le produit cartésien des deux ensembles E et F est noté $E \times F$. Il est défini par :

$$E \times F = \{(x, y), x \in E \text{ et } y \in F\}.$$

8. **Ensemble des parties.** L'ensemble des parties d'un ensemble E , noté $\mathcal{P}(E)$, est l'ensemble de tous les sous-ensembles de E . Ainsi,

$$A \in \mathcal{P}(E) \Leftrightarrow (A \subset E).$$

2.2.2 Dénombrement de listes

Une p -liste (ou p -uplet) d'éléments d'un ensemble fini non vide E est un élément de l'ensemble $E \times E \times \dots \times E =: E^p$.

Si $\text{card}(E) = n$ alors $\text{card}(E^p) = n^p$.

Dans un ensemble à n éléments, on peut donc former au total n^p p -listes distinctes.

Proposition 2.1. *Le nombre d'applications d'un ensemble A à p éléments vers un B à n est égal à n^p .*

2.2.3 Dénombrement d'arrangements

Définition 2.2. *Soit E un ensemble à n éléments et p un nombre entier non nul tel que $n \geq p$. On appelle arrangement de p éléments de E (ou p -arrangement d'éléments de E), toute p -liste d'éléments deux à deux distincts de E .*

Le nombre d'arrangements de p éléments d'un ensemble à n éléments, noté A_n^p , est donné par :

$$A_n^p = n(n-1)(n-2)\dots(n-p+1) = \frac{n!}{(n-p)!}.$$

Proposition 2.2. *Soit A et B deux ensembles finis tels que $\text{Card}(A) = p \leq n = \text{Card}(B)$. Le nombre d'applications injectives de A dans B est égal à A_n^p .*

Définition 2.3. *E étant un ensemble non vide tel que $\text{card}(E) = n$, on appelle permutation des n éléments de E tout arrangement des n éléments de E .*

Le nombre de permutations de n éléments est $A_n^n = n!$.

2.2.4 Combinaisons

Soit E un ensemble non vide tel que $\text{card}(E) = n$. Soit $p \in \mathbb{N}$ tel que $p \leq n$. On appelle combinaison de p éléments de E toute partie de E ayant p éléments.

Le nombre de combinaisons de p éléments d'un ensemble à n éléments, noté C_n^p est :

$$C_n^p = \frac{n!}{p!(n-p)!}$$

Proposition 2.3. (Formule du binôme de Newton) Soit a et b deux nombres réels et $n \in \mathbb{N}^*$, on a le résultat suivant :

$$(a + b)^n = \sum_{k=0}^n C_n^k a^{n-k} b^k$$

2.3 Mesure de probabilité et espace probablisé

2.3.1 Notion de tribu

Définition 2.4. Soit Ω un ensemble non vide. On appelle tribu sur Ω , toute famille \mathcal{F} de parties de Ω vérifiant :

- $\emptyset \in \mathcal{F}$
- $\forall A \in \mathcal{F}, A^C \in \mathcal{F}$
- $(\forall n \in I \subset \mathbb{N}, A_n \in \mathcal{F}) \Rightarrow \bigcup_{n \in I} A_n \in \mathcal{F}$

Exemple 2.1. Quelques exemples de tribus.

- $\mathcal{F}_0 = \{\emptyset, \Omega\}$
- $\mathcal{F}_1 = \mathcal{P}(\Omega)$
- $\mathcal{F}_2 = \{\emptyset, \Omega, A, A^C\}$

Proposition 2.4. Soit Ω un ensemble non vide et A un sous-ensemble non vide de Ω . L'ensemble des tribus de parties de Ω contenant A est une tribu sur Ω . C'est la plus petite tribu sur Ω contenant A encore appelée tribu engendrée par A .

Proposition 2.5. Soit \mathcal{F} une tribu sur un ensemble Ω . On a :

- \mathcal{F} est stable par réunion dénombrable ou finie.
- \mathcal{F} est stable par intersection dénombrable ou finie.
- Si A et B sont des éléments de \mathcal{F} , $A \setminus B$ et $A \Delta B$ sont aussi des éléments de \mathcal{F} .

Définition 2.5. On appelle espace mesurable ou espace probablisable, un couple (Ω, \mathcal{F}) , où Ω est un ensemble non vide et \mathcal{F} une tribu de parties de Ω .

2.3.2 Mesures

Définition 2.6. Soit \mathcal{F} une tribu sur Ω . On appelle mesure positive sur l'espace mesurable (Ω, \mathcal{F}) , une application $\mu : \mathcal{F} \rightarrow [0, +\infty[$ telle que :

- $\mu(\emptyset) = 0$
- μ est σ -additive ; c'est-à-dire : pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de \mathcal{F} deux à deux disjoints, on a :

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n)$$

Définition 2.7. -Si μ est une mesure sur (Ω, \mathcal{F}) , le triplet $(\Omega, \mathcal{F}, \mu)$ est appelé espace mesuré.

- Si de plus $\mu(\Omega) = 1$, la mesure μ est appelée mesure de probabilité sur l'espace mesurable Ω (ou simplement probabilité) et le triplet $(\Omega, \mathcal{F}, \mu)$ est appelé espace probabilisé.

2.4 Calcul de probabilités

2.4.1 Vocabulaire des probabilités

Définition 2.8. Une expérience aléatoire (ou épreuve) est une expérience dont l'issue n'est pas prévisible avec une certitude totale, mais dont on connaît l'ensemble de tous les résultats possibles appelé espace fondamental.

Eventualité

On appelle éventualité ou cas possible tout résultat susceptible de se produire.

Univers

Nous appelons univers (des possibles) associé à une expérience aléatoire \mathcal{E} , l'ensemble de tous les résultats possibles de \mathcal{E} . L'univers est donc l'ensemble de toutes les éventualités. Généralement, l'univers est noté Ω .

Événement

Lorsqu'on exécute une expérience aléatoire c'est souvent dans le but d'examiner si une propriété ou si une proposition A donnée est vérifiée ou non par le résultat obtenu. Par exemple, dans l'expérience aléatoire consistant à lancer un dé parfait, un joueur peut vérifier si le résultat obtenu est un chiffre pair ou multiple de 3. En théorie de probabilités, on parlera d'événement, et on dira que l'événement A est réalisé si le résultat de l'épreuve possède la propriété A , et qu'il n'est pas réalisé si le résultat ne possède pas cette propriété. Un événement A associé à une épreuve se définit donc comme un sous-ensemble de l'univers Ω constitué des résultats qui possèdent la propriété A . C'est tout membre de la tribu $\mathcal{P}(\Omega)$.

Terminologie des événements

- L'ensemble vide, \emptyset , est appelé événement impossible.
 - Ω est appelé événement certain.
 - Tout singleton $\{\omega\}$, où $\omega \in \Omega$, est appelé événement élémentaire.
- À tout événement A est associé son contraire, non A ou \bar{A} ou A^c qui est réalisé si et seulement si A ne l'est pas.
- Pour tout couple d'événements A et B , l'événement « A et B » est réalisé si A et B sont réalisés. On le note $A \cap B$.
 - Pour tout couple d'événements A et B , l'événement « A ou B » est réalisé si l'un des deux ou si les deux sont réalisés. On le note $A \cup B$.
 - Deux événements A et B sont incompatibles si la réalisation de l'un implique la non réalisation de l'autre.
 - Les événements A_1, A_2, \dots, A_n forment un système complet d'événements ou système exhaustif si les ensembles qui leur sont associés forment une partition de l'univers Ω .

2.4.2 Définition de la probabilité et propriétés

Définition 2.9. Une probabilité sur un espace probabilisable (Ω, \mathcal{F}) est une application p de \mathcal{F} dans $[0, 1]$ vérifiant les deux propriétés suivantes (axiomes

de Kolmogorov) :

1. $p(\emptyset) = 1$;
2. pour tout suite $(A_n)_{n \in I}$ au plus dénombrable d'événements de \mathcal{F} deux à deux incompatibles, nous avons

$$p\left(\bigcup_{n \in I} A_n\right) = \sum_{n \in I} p(A_n)$$

c'est-à-dire que la probabilité d'un événement qui est la réunion disjointe d'événements est égale à la somme des probabilités de ces événements (σ -additivité). Le réel $p(A)$ est appelé la probabilité de l'événement A . Le triplet (Ω, \mathcal{F}, p) est appelé espace probabilisé.

Remarque 2.2. Pour la suite on prendra comme espace probabilisable le couple $(\Omega, \mathcal{P}(\Omega))$.

Proposition 2.6. Soit $(\Omega, \mathcal{P}(\Omega), p)$ un espace probabilisé; et $A \in \mathcal{P}(\Omega)$, $B \in \mathcal{P}(\Omega)$. On a :

1. $p(\emptyset) = 0$
2. $p(\bar{A}) = 1 - p(A)$
3. $p(A) \leq p(B)$ si $A \subset B$
4. $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

2.4.3 Equiprobabilité

Soit $(\Omega, \mathcal{P}(\Omega), p)$ un espace probabilisé fini c'est-à-dire l'univers Ω est fini. Les événements A et B sont dits équiprobables s'ils ont même probabilité (d'être réalisés), c'est-à-dire $p(A) = p(B)$.

On dit qu'il y a équiprobabilité lorsque les probabilités de tous les événements élémentaires sont égales. Dans ce cas, p est la probabilité uniforme sur $(\Omega, \mathcal{P}(\Omega))$.

Conséquence : S'il y a équiprobabilité, pour tout événement A , on a :

$$p(A) = \frac{\text{Card}A}{\text{Card}\Omega}.$$

Exemple 2.3. Une urne contient 8 boules blanches et 6 boules noires, chaque boule ayant la même probabilité d'être tirée.

1. On tire simultanément 5 boules. Quelle est la probabilité d'obtenir :
 - a) 3 boules blanches et 2 boules noires ?
 - b) des boules de couleurs différentes ?
2. On tire successivement 5 boules sans remise. Quelle est la probabilité d'avoir :
 - a) exactement 3 boules blanches ?
 - b) moins de 2 boules noires ?

2.5 Probabilité conditionnelle- Indépendance

2.5.1 Probabilité conditionnelle

Définition 2.10. Soit $(\Omega, \mathcal{P}(\Omega), p)$ un espace probabilisé. Soit B un événement de probabilité non nulle. Pour tout événement A , posons

$$p(A/B) = \frac{p(A \cap B)}{p(B)}.$$

L'application qui à un événement A associe $p(A/B)$ est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, appelée probabilité conditionnelle relative à B ou probabilité sachant B .

Remarque 2.4. $p(A/B)$ est aussi noté $p_B(A)$.

Formule des probabilités composées

La formule des probabilités composées fournit la règle de calcul de la probabilité de la réalisation simultanée de deux événements. de la définition axiomatique de la probabilité conditionnelle, il résulte que si A et B sont deux événements de probabilités non nulles, on a :

$$p(A \cap B) = p(A) \times p_A(B) = p(B) \times p_B(A).$$

Cette relation porte le nom de formule des probabilités composées.

Formule des probabilités totales

Soit $(A_i)_{i \in I}$ un système complet d'événements de probabilités toutes non nulles. Pour tout événement B , on a :

$$p(B) = \sum_{i \in I} p(B \cap A_i) = \sum_{i \in I} p(B/A_i) \times p(A_i)$$

Formule de Bayes

Pour tout système complet d'événements $(A_i)_{i \in I}$ de probabilités non nulles et pour tout événement B de probabilité non nulle, on a :

$$p(A_i/B) = \frac{p(B/A_i) \times p(A_i)}{\sum_{j \in I} p(B/A_j) \times p(A_j)}.$$

Exemple 2.5. *Trois machines automatiques produisent des pièces de voitures. La machine M_1 produit 40% du total des pièces, la machine M_2 25% et la machine M_3 produit 35%. En moyenne, les pourcentages des pièces non conformes aux critères imposés sont de 10% pour la machine M_1 , de 5% pour la machine M_2 et de 1% pour la machine M_3 .*

Une pièce est choisie au hasard dans la production totale des trois machines. On constate qu'elle n'est pas conforme aux critères imposés. Quelle est la probabilité qu'elle ait été produite par la machine M_1 ?

2.5.2 Indépendance

Définition 2.11. *L'événement A est indépendant de l'événement B si la probabilité de réalisation de l'événement A n'est pas modifiée par une information concernant la réalisation de l'événement B , c'est-à-dire si :*

$$p(A/B) = p(A).$$

Il est équivalent de dire que A et B sont indépendants si et seulement si

$$p(A \cap B) = p(A) \times p(B).$$

Proposition 2.7. *Si A et B sont indépendants alors \bar{A} et B sont également indépendants.*

Indépendance mutuelle

Les événements $A_i, i \in \{1, \dots, n\}$, sont mutuellement indépendants si, pour toute partie I de l'ensemble des indices, on a :

$$p\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} p(A_i)$$

L'indépendance mutuelle implique l'indépendance deux à deux mais c'est une condition plus forte.

Exemple 2.6. *On lance deux dés et on considère les événements suivants :*

- A : le premier dé donne une face impaire,
- B : le deuxième dé donne une face impaire,
- C : la somme des points apparaissant sur les deux faces est impaire.

Les événements A, B et C sont deux à deux indépendants. En effet :

$$P(A) = P(B) = P(C) = \frac{1}{2}$$

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}.$$

Les événements A, B et C ne sont pas mutuellement indépendants :

$$P(A \cap B \cap C) = 0.$$

VARIABLES ALÉATOIRES RÉELLES

3.1 Tribu des boréliens

3.1.1 Ouverts et fermés de \mathbb{R} , tribu des boréliens

Nous rappelons ici quelques éléments de la topologie de l'ensemble des nombres réels \mathbb{R} .

Définition 3.1. - Soit x_0 un élément de \mathbb{R} , un sous-ensemble \mathcal{V} de \mathbb{R} est un voisinage de x_0 s'il existe un intervalle ouvert contenant x_0 et contenu dans \mathcal{V} .

- Un sous-ensemble de \mathbb{R} est dit ouvert s'il est vide ou bien s'il est voisinage de chacun de ses points.
- Un sous-ensemble de \mathbb{R} est dit fermé si son complémentaire est ouvert.

Les ouverts et les fermés de \mathbb{R} vrifient les propriétés suivantes.

- Proposition 3.1.**
1. Une réunion quelconque d'ouverts est un ouvert ;
 2. Une intersection finie d'ouverts est un ouvert ;
 3. \mathbb{R} et \emptyset sont à la fois des ouverts et des fermés.
 4. Une intersection quelconque de fermés est un fermé ;
 5. Une réunion finie de fermés est un fermé.

Définition 3.2. On appelle tribu des boréliens, la tribu de parties de \mathbb{R} engendrée par la famille des ouverts de \mathbb{R} et notée $\mathcal{B}(\mathbb{R})$. Chaque élément de \mathcal{B} est appelé un borélien.

3.2 Notion de Variable aléatoire réelle

Définition 3.3. *une variable aléatoire réelle définie sur un espace probabilisable (Ω, \mathcal{F}) est toute application $X : \Omega \rightarrow \mathbb{R}$ telle que pour tout borélien B de \mathbb{R} , $X^{-1}(B) \in \mathcal{F}$. L'ensemble $X(\Omega)$ est appelé ensemble fondamental de la variable aléatoire réelle X ou encore univers-image.*

3.2.1 Loi de probabilité d'une variable aléatoire

Soit X une variable aléatoire réelle définie sur l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$. La loi de probabilité de la variable aléatoire X , on dit aussi la distribution de X , est le mode de répartition de la probabilité p sur l'ensemble des observables. Il en résulte que connaître la loi de probabilité de X , c'est être capable d'attribuer une probabilité à tout événement " $X \in I$ ".

Plus précisément, la loi de probabilité de X est une mesure de probabilité définie sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ par : $\forall B \in \mathcal{B}(\mathbb{R}), P_X(B) = p(X^{-1}(B))$. Une variable aléatoire réelle traduit donc l'idée de résultat numérique associé à un phénomène aléatoire.

3.2.2 Fonction de répartition d'une variable aléatoire

Une loi de probabilité est une mesure abstraite sur \mathbb{R} , elle est donc en général peu maniable et peu utilisée dans les applications concrètes. Or, la tribu de Borel $\mathcal{B}(\mathbb{R})$ contient les intervalles du type $] -\infty, x[$, on en déduit la notion de fonction de répartition.

Soit X une variable aléatoire réelle définie sur l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$. La fonction de répartition de X est définie par :

$$F : \mathbb{R} \longrightarrow [0, 1]$$

$$x \longmapsto F(x) = p(X < x)$$

Remarque 3.1. *Soit X une variable aléatoire réelle et F sa fonction de répartition.*

- La fonction de répartition F est croissante,
- la limite de $F(x)$ quand x tend vers $-\infty$ est égale à 0,

- la limite de $F(x)$ quand x tend vers $+\infty$ est égale à 1,
- La fonction de répartition est continue à gauche.
- $x_0 \in \mathbb{R}$ est un point de discontinuité, on a $P(X = x_0) = F(x_0^+) - F(x_0)$

Réciproquement on a le théorème suivant.

Théorème 3.2. Soit F une fonction définie sur \mathbb{R} à valeurs dans $[0, 1]$ vérifiant les propriétés suivantes :

1. F est croissante ;
2. $\lim_{-\infty} F = 0$ et $\lim_{+\infty} F = 1$
3. F est continue à gauche en chacun de ses points de discontinuité.

Alors F est la fonction de répartition d'une variable aléatoire réelle.

3.3 Variable aléatoire réelle discrète

Définition 3.4. On dit qu'une variable aléatoire réelle X est discrète si son ensemble fondamental (univers-image) est un sous-ensemble fini ou dénombrable de \mathbb{R} . Dans ce cas, on notera $x_n, n \in \mathbb{N}^*$, les éléments de l'ensemble fondamental rangés par ordre de valeurs croissantes.

3.3.1 Moments d'une variable aléatoire

Moments d'ordre k d'une variable aléatoire

Soit X une variable aléatoire réelle prenant un ensemble fini ou dénombrable de valeurs sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$ et $\{x_i\}_{(i \in I)}$ l'ensemble des valeurs prises par X . Dans la suite on supposera que l'ensemble de ces valeurs est fini. Les formules sont cependant également valables dans le cas dénombrable sous l'hypothèse d'existence des quantités définies.

Soit k un entier naturel non nul. On appelle moment d'ordre k de X , le nombre réel noté $E(X^k)$ ou $m_k(X)$ défini par :

$$m_k(X) = E(X^k) = \sum_{i \in I} (x_i)^k p(X = x_i)$$

Espérance mathématique

L'espérance mathématique (ou valeur moyenne) d'une v.a.r. discrète X est égale au moment d'ordre 1 de X . On le note $E(X)$ et on a :

$$E(X) = \sum_{i \in I} x_i p(X = x_i)$$

Proposition 3.2. Soient X et Y deux v.a.r. discrète, λ et μ deux nombres réels.

On a :

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y),$$

$$E(\lambda X + \mu) = \lambda E(X) + \mu.$$

Espérance mathématique d'une fonction de variable aléatoire

De façon générale, pour une distribution de probabilité d'une variable aléatoire discrète X sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$, l'espérance mathématique (la moyenne) d'une fonction $g(X)$ de X est donnée par :

$$E(g(X)) = \sum_{i \in I} g(x_i) p(X = x_i)$$

Moments centrés d'ordre k d'une variable aléatoire

De même on définit le moment centré d'ordre k de X , noté $\mu_k(X)$ défini par :

$$\mu_k(X) = \sum_{i \in I} (x_i - E(X))^k p(X = x_i)$$

Variance et écart-type d'une variable aléatoire

On appelle variance d'une v.a.r. X , son moment centré d'ordre 2. On le note $V(X)$ (ou $\sigma^2(X)$) et on a :

$$V(X) = \sum_{i \in I} (x_i - E(X))^2 p(X = x_i)$$

L'écart-type de X est le réel positif $\sigma(X)$ défini par

$$\sigma(X) = \sqrt{V(X)}$$

Proposition 3.3. Soit X v.a.r. discrète sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$, λ et μ deux nombres réels. On a :

1. $V(X) = E(X^2) - (E(X))^2$,
2. $V(\lambda X + \mu) = \lambda^2 V(X)$,
3. $\sigma(\lambda X + \mu) = |\lambda| \sigma(X)$.

3.4 Variable aléatoire réelle continue

Un deuxième type de distribution, pour une variable aléatoire à valeur dans \mathbb{R} , est caractérisé par sa fonction de répartition continue sur \mathbb{R} . On dit alors que la variable aléatoire est de distribution continue.

Définition 3.5. Une v.a.r. X est dite continue si sa fonction de répartition est continue à droite, donc continue (parce que déjà continue à gauche par définition).

Une variable aléatoire continue prend ses valeurs sur un ensemble infini non dénombrable de points, elle décrit par exemple la durée de vie d'une batterie de voiture, l'heure d'arrivée des voitures à un péage donné d'autoroute...

Remarque 3.3. Si X est une variable aléatoire continue de loi de Probabilité P_X , alors on a :

- Pour tout $t \in \mathbb{R}$, $P_X(\{t\}) = 0$. Pour une variable aléatoire de distribution continue, la probabilité d'observer une valeur réelle t donnée est donc nulle quelle que soit t .

- Il résulte de ce qui précède que, si X est de distribution continue, et si $a < b$, on a : $P_X([a, b]) = P_X(]a, b]) = P_X([a, b[) = P_X(]a, b[)$

Exemple 3.4. On considère la fonction $F : \mathbb{R} \rightarrow \mathbb{R}$ définie par :

$$\begin{cases} F(x) = 0 & \text{si } x \leq 0 \\ F(x) = \frac{1}{2}x & \text{si } 0 < x \leq 2 \\ F(x) = 1 & \text{si } x > 2 \end{cases}$$

Justifier que F est la fonction de répartition d'une variable aléatoire continue X .

3.4.1 Densité de probabilité et fonction de répartition

Définition 3.6. Une fonction réelle f est une densité de probabilité si :

- f est positive ;
- f est continue sur \mathbb{R} sauf peut-être en un nombre fini de points ;
- $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Définition 3.7. On dit qu'une v.a.r. X admet une densité de probabilité f lorsque sa fonction de répartition F peut s'écrire sous la forme

$$F(x) = \int_{-\infty}^x f(t)dt,$$

où f est une fonction densité de probabilité. On dit d'une telle variable aléatoire réelle qu'elle est absolument continue de densité f .

Remarque 3.5. Si une v.a.r. X admet une densité de probabilité f , sa fonction de répartition F admet pour dérivée f . Il en résulte que toute distribution absolument continue est continue. La réciproque n'est en général pas vraie.

Exemple 3.6. Soit X une variable aléatoire réelle, de densité de probabilité f définie par :

$$\begin{cases} f(x) = 0 & \text{si } x < 0 \\ f(x) = e^{-x} & \text{si } x \geq 0 \end{cases}$$

1. Justifier que f est effectivement une densité de probabilité.
2. Déterminer la fonction de répartition de X .
3. Calculer $P(-1 \leq X \leq 3)$

3.4.2 Moments d'une variable aléatoire

Soit X une variable aléatoire réelle, de densité de probabilité f .

- Le moment d'ordre k ($k \in \mathbb{N}^*$) de X (s'il existe), est le nombre réel noté $E(X^k)$ ou $m_k(X)$ défini par :

$$m_k(X) = E(X^k) = \int_{-\infty}^{+\infty} x^k f(x)dx$$

- Pour $k = 1$, on obtient l'espérance mathématique de X définie par :

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

- Le moment centré d'ordre k ($k \in \mathbb{N}^*$) de X (s'il existe), est le nombre réel noté $\mu_k(X)$ défini par :

$$\mu_k(X) = \int_{-\infty}^{+\infty} (x - E(X))^k f(x) dx$$

- Pour $k = 2$, on obtient la variance de X définie par :

$$V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx$$

- Soit X une v.a.r. de densité f , admettant une espérance mathématique $E(X)$. Alors, $V(X)$ existe si et seulement si $E(X^2)$ existe et on a :

$$V(X) = E(X^2) - [E(X)]^2$$

- Si X admet une variance, on appelle écart-type de X le nombre réel

$$\sigma(X) = \sqrt{V(X)}$$

- Soit X une v.a.r. de densité f , et ϕ une fonction réelle mesurable, l'espérance mathématique de $Y = \phi(X)$ est donnée par :

$$E(Y) = E(\phi(X)) = \int_{-\infty}^{+\infty} \phi(x) f(x) dx$$

- Une v.a.r. X est dite centrée si son espérance mathématique est nulle.

- Une v.a.r. X est dite réduite si son écart-type est égal à 1.

- Une v.a.r. X est dite centrée réduite si son espérance mathématique est nulle et son écart-type est égal 1.

- Soit X une variable aléatoire d'espérance $E(X)$ et d'écart type $\sigma(X) \neq 0$. Alors la v.a.r. $U = \frac{X - E(X)}{\sigma(X)}$ est une variable aléatoire centrée réduite.

Exemple 3.7. On considère la variable aléatoire X de densité f définie par

$$\begin{cases} f(x) = 2x & \text{si } x \in]0, 1[\\ f(x) = 0 & \text{sinon} \end{cases}$$

1. Déterminer la fonction de répartition de X .

2. Calculer l'espérance mathématique et l'écart-type de X .

3.5 Fonction génératrice des moments et fonction caractéristique

3.5.1 Fonction génératrice des moments

L'utilité de la fonction génératrice des moments réside dans le fait qu'elle facilite bien souvent le calcul des moments (simples) de divers ordres d'une loi donnée.

Définition 3.8. On appelle fonction génératrice des moments de la variable aléatoire réelle X , si elle existe, la fonction définie par :

$$M_X(t) = E(e^{tX}).$$

Théorème 3.8. Soit X une variable aléatoire dont la fonction génératrice des moments M_X est définie pour tout $t \in]t_1, t_2[$ avec $t_1 < 0 < t_2$. Alors :

1. Tous les moments de X existent.
2. Pour tout $t \in]-s, s[$ avec $0 < s < t_0 = \min(-t_1, t_2)$, $M_X(t)$ admet un développement en série entière :

$$M_X(t) = 1 + E(X)t + \frac{E(X^2)}{2!}t^2 + \frac{E(X^3)}{3!}t^3 + \dots + \frac{E(X^k)}{k!}t^k + \dots$$

3. Pour tout entier naturel non nul k , on a :

$$E(X^k) = \frac{d^k}{dt^k} M_X(0).$$

Exemple 3.9. On considère la variable aléatoire réelle X , de densité de probabilité f définie par :

$$\begin{cases} f(x) = 0 & \text{si } x < 0 \\ f(x) = e^{-x} & \text{si } x \geq 0 \end{cases}$$

1. Déterminer sa fonction génératrice des moments.
2. Déterminer $E(X^k)$ pour tout $k \in \mathbb{N}^*$.

3.5.2 Fonction caractéristique

C'est une fonction complexe de variable réelle que l'on peut associer (bijectivement) à toute variable aléatoire réelle et qui est un outil privilégié d'étude des lois de probabilité.

Définition 3.9. Soit X une variable aléatoire réelle. On appelle fonction caractéristique de X la fonction φ_X de la variable réelle t définie par

$$\varphi_X(t) = E(e^{itX})$$

où i est le nombre imaginaire vérifiant $i^2 = -1$.

Proposition 3.4. Soit X une variable aléatoire réelle de fonction génératrice des moments M_X et de fonction caractéristique φ_X .

1. $\varphi_X(0) = 1$.
2. $|\varphi_X(t)| \leq 1$ et $\varphi_X(-t) = \overline{\varphi_X(t)}$.
3. $\varphi_X(t) = M_X(it)$
4. Si $Y = aX + b$, on a : $\varphi_Y(t) = e^{ibt} \varphi_X(at)$

Remarque 3.10. Malgré la relation directe de la fonction génératrice avec les moments, les mathématiciens préfèrent travailler avec la fonction caractéristique, qui peut être définie pour toute variable aléatoire X , bijectivement, et qui existe pour toute valeur de t .

3.6 Inégalités intéressantes

3.6.1 Inégalité de Markov

Proposition 3.5. Soit X une variable aléatoire réelle sur un espace probabilisé (Ω, \mathcal{F}, p) , g une fonction borélienne définie de \mathbb{R} dans $]0; +\infty[$. Alors, si $E[g(X)]$ existe, pour tout $\epsilon > 0$, on a :

$$p(g(X) \geq \epsilon) \leq \frac{E[g(X)]}{\epsilon}$$

3.6.2 Inégalité de Bienaymé-Tchebychev

Proposition 3.6. Soit X une v.a.r. sur un espace probabilisé (Ω, \mathcal{F}, p) . Alors, si $V(X)$ existe, pour tout a strictement positif, on a :

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

Cette inégalité est souvent démontrée à partir de l'inégalité de Markov. Elle permet de démontrer la loi faible des grands nombres, et son intérêt théorique est donc grand.

Exemple 3.11. *On suppose que le nombre de pièces sortant d'une usine donnée en l'espace d'une semaine est une v.a.r d'espérance 50 et de variance 25. Peut-on estimer la probabilité que la production de la semaine prochaine soit comprise entre 40 et 60 pièces ?*

LOIS CLASSIQUES DE PROBABILITÉ

Bien souvent, pour décrire un ensemble de données ou faire des modélisations statistiques, il est nécessaire de connaître parfaitement les lois classiques d'usage courant. Le choix d'une loi est lié à plusieurs considérations, entre autres, la nature du phénomène étudié afin de choisir entre loi discrète ou loi continue, la forme de la distribution (histogramme), la connaissance et l'interprétation des principales caractéristiques de l'ensemble des données : espérance, médiane, variance, écart-type, coefficients d'asymétrie et d'aplatissement etc. De façon générique et sauf mention expresse, on notera X une variable aléatoire réelle (v.a.r.) qui suit la loi décrite.

4.1 Lois de probabilité discrètes

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

- à la nature du phénomène étudié afin de choisir entre loi discrète et loi continue,
- à la forme de la distribution (histogramme),
- 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.,
- au nombre de paramètres des lois, une loi dépendant de plusieurs paramètres peut s'adapter plus facilement à une distribution.

4.1.1 Loi de Dirac

La loi de Dirac au point a de \mathbb{R} est la loi de probabilité δ_a , définie par

$$\delta_a(X = a) = 1, \delta_a(X = x) = 0 \text{ si } x \neq a.$$

C'est la loi la plus simple, associée à un phénomène déterministe X dont le résultat de toute expérience est égale à a (i.e X est une variable certaine). On a :

$$E(X) = a, V(X) = 0.$$

4.1.2 Loi uniforme discrète

Soit n un entier naturel non nul ; on dit qu'une v.a.r. X suit la loi uniforme discrète sur $\{1, 2, \dots, n\}$ lorsque X prend les valeurs $k = 1, 2, \dots, n$ avec les probabilités

$$p(X = k) = \frac{1}{n}$$

On a

$$E(X) = \frac{n+1}{2}, V(X) = \frac{n^2-1}{12}$$

La loi uniforme discrète intervient dans de nombreux domaines comme les jeux de pile ou face, les jeux de dé (parfaitement équilibré), les jeux de carte, de loterie, les sondage etc.

4.1.3 Loi de Bernoulli

On considère une expérience dont le résultat ne peut prendre que deux valeurs appelées, par convention, succès ou échec : un candidat est reçu ou non à un examen, une pièce usinée est bonne ou défectueuse, une porte est ouverte ou fermée... À une expérience de ce type, est associée une variable aléatoire X prenant la valeur 1 pour le succès et la valeur 0 pour l'échec, avec les probabilités respectives p et $(1 - p) = q$. Cette variable est appelée variable de Bernoulli.

La loi de probabilité d'une variable de Bernoulli est définie par :

$$p(X = 1) = p; p(X = 0) = 1 - p.$$

Ses moments sont :

$$E(X) = p \text{ et } Var(X) = p(1 - p) = pq.$$

4.1.4 Loi binomiale

On réalise n épreuves indépendantes de la même expérience telles que :

- chaque épreuve ne peut avoir que deux résultats, s'excluant mutuellement, soit le succès, soit l'échec,
- la probabilité p de succès est constante à chaque épreuve, la probabilité d'échec est également constante et égale à $1 - p = q$. Soit X la variable aléatoire qui « compte » le nombre de succès au cours de n épreuves.

La loi de la variable X est appelée loi binomiale de paramètres (n, p) , notée $\mathcal{B}(n; p)$. X peut prendre les valeurs $0, \dots, n$.

$$\forall k \in \{0, \dots, n\}, p(X = k) = C_n^k p^k (1 - p)^{n-k}.$$

Une variable binomiale est égale à la somme de n variables aléatoires indépendantes de Bernoulli, la loi binomiale est donc la loi des épreuves répétées.

Proposition 4.1. - *Si la variable aléatoire X suit la loi binomiale de paramètres (n, p) alors on a : $E(X) = np$ et $Var(X) = np(1 - p) = npq$.*

- *Si X_1 et X_2 sont deux variables aléatoires indépendantes suivant respectivement les lois binomiales $\mathcal{B}(n_1; p)$ et $\mathcal{B}(n_2; p)$, alors la variable aléatoire réelle $X_1 + X_2$ suit la loi binomiale $\mathcal{B}(n_1 + n_2; p)$.*

Exemple 4.1. *On veut réaliser une étude clinique sur des malades se présentant à une consultation hospitalière. Pour cette étude, seuls les malades répondant à un ensemble de critères C sont retenus. Des statistiques antérieures ont montré que 20% des consultants présentent les critères C . 10 malades viennent consulter le premier jour. Soit X la variable aléatoire « nombre de malades retenus » c'est-à-dire répondant à l'ensemble des critères C .*

1. *Calculer la probabilité qu'aucun malade ne soit retenu ce jour.*
2. *Calculer la probabilité qu'au moins un malade soit retenu ce jour.*

4.1.5 Loi géométrique

La loi géométrique est la loi de la variable X « 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.

$$P(X = k) = p(1 - p)^{k-1}, k \in \mathbb{N}^*.$$

On a :

$$E(X) = \frac{1}{p}, V(X) = \frac{1-p}{p^2}$$

Exemple 4.2. *Un certain matériel a une probabilité $p = 0,02$ constante de défaillance à chaque mise en service. On procède à l'expérience suivante, l'appareil est mis en marche, arrêté, remis en marche, arrêté, jusqu'à ce qu'il tombe en panne. Calculer la probabilité que ce matériel tombe en panne pour la première fois au dixième essai.*

4.1.6 Loi de Pascal

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

$$p(Z = z) = pC_{z-1}^{k-1}p^{k-1}(1-p)^{z-k}, z \geq k, z \in \mathbb{N}^*.$$

On a :

$$E(Z) = \frac{k}{p}, V(Z) = \frac{k(1-p)}{p^2}$$

4.1.7 Loi binomiale négative

À partir de la loi de Pascal, on définit la loi binomiale négative, ou loi de la variable aléatoire $T = Z - k$: c'est la loi d'une variable aléatoire discrète de compte du nombre d'échecs précédant le k -ième succès dans des épreuves répétées.

$$p(T = t) = C_{k+t-1}^{k-1}p^{k-1}(1-p)^t, z \geq k, z \in \mathbb{N}^*.$$

On a :

$$E(T) = \frac{k(1-p)}{p}, \quad V(T) = \frac{k(1-p)}{p^2}$$

4.1.8 Loi hypergéométrique

Soit une population de N individus parmi lesquels une proportion p (donc Np individus) possède un certain caractère. On effectue un tirage équiprobable et sans remise (tirage exhaustif) d'un échantillon de n individus parmi cette population. Soit X la variable aléatoire réelle égale au nombre d'individus de l'échantillon possédant la propriété envisagée. La variable aléatoire ainsi définie suit une loi hypergéométrique $\mathcal{H}(N; n; p)$. Cette loi dépend de trois paramètres N, n et p . La v.a.r. X prend les valeurs k comprises entre $\min X = \max(0; n - (1-p)N)$ et $\max X = \min(n; Np)$ avec les probabilités

$$p(X = x) = \frac{C_{Np}^x C_{N-Np}^{n-x}}{C_N^n}.$$

Le quotient $\frac{n}{N}$ est appelé taux de sondage et le réel $\frac{N-n}{N-1}$ est le facteur d'exhaustivité. On a :

$$E(X) = np, \quad Var(X) = np(1-p) \frac{N-n}{N-1}.$$

Remarque 4.3. *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 $\mathcal{B}(n; p)$. En pratique si le taux de sondage $\frac{n}{N}$ est inférieur à 0.1 et si X suit la loi hypergéométrique $\mathcal{H}(N; n; p)$, on peut considérer que X suit approximativement la loi binomiale $\mathcal{B}(n; p)$.*

4.1.9 Loi de Poisson

La loi de Poisson de paramètre λ est la loi d'une variable aléatoire discrète réelle X , prenant toutes les valeurs entières positives, avec les probabilités :

$$p(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \mathbb{N}^*.$$

On a :

$$E(X) = V(X) = \lambda.$$

Proposition 4.2. *La somme de deux variables aléatoires de Poisson, indépendantes, de paramètres λ_1 et λ_2 , est une variable aléatoire de Poisson, de paramètre $\lambda = \lambda_1 + \lambda_2$.*

La loi de Poisson est la loi des événements rares ou loi des petites probabilités. Un phénomène discret est rare si la probabilité qu'il apparaisse plus d'une fois est petite : on appelle signal une apparition du phénomène.

Exemple 4.4. *Selon les données recueillies depuis plusieurs années, le nombre de pannes hebdomadaires du système informatique d'une entreprise suit une loi de Poisson de paramètre $\lambda = 0,05$. Soit X la variable aléatoire « nombre de pannes hebdomadaires ». Calculer $p(X \geq 2)$.*

Approximation d'une loi binomiale par une loi de Poisson

On considère une variable aléatoire X suivant la loi binomiale $\mathcal{B}(n, p)$ telle que :

- le nombre n d'essais est assez grand (en pratique, $n \geq 30$)
- la probabilité de succès p est assez petite (en pratique $p < 0,1$)
- le produit np est fini et pas assez grand (en pratique $np \leq 15$).

Dans ce cas, on peut approximer la loi binomiale par une loi de Poisson de paramètre $\lambda = np$.

4.2 Lois de probabilité continues

4.2.1 Loi uniforme

Une variable aléatoire réelle X suit une loi uniforme (continue) sur l'intervalle $[a, b]$ ($a < b$), si sa densité de probabilité est définie par :

$$\begin{cases} f(x) = \frac{1}{b-a} & \text{si } x \in [a, b] \\ f(x) = 0 & \text{sinon} \end{cases}$$

La loi uniforme sur $[a, b]$ traduit l'hypothèse d'équirépartition, ou répartition indifférente, sur $[a, b]$. Son espérance mathématique et sa variance sont données par :

$$E(X) = \frac{a+b}{2}; \quad V(X) = \frac{(b-a)^2}{12}$$

4.2.2 La loi exponentielle

Une variable aléatoire réelle positive X suit une loi exponentielle, de paramètre $\lambda > 0$, si sa densité de probabilité est donnée par :

$$\begin{cases} f(x) = \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ f(x) = 0 & \text{sinon} \end{cases}$$

La loi exponentielle est la loi d'une variable aléatoire continue « temps d'attente » qui intervient notamment dans le processus de Poisson et la modélisation de la fiabilité. En théorie, la loi exponentielle est la loi du temps d'attente d'un processus poissonnien de taux λ : temps d'attente du premier évènement, ou intervalle entre deux évènements successifs (dans certaines situations concrètes, on parle de « durée de vie » plutôt que de temps d'attente).

Dans la pratique, la loi exponentielle est la loi des situations concrètes modélisées par un processus poissonnien, comme une succession de pannes à « taux de défaillance » constant, ou la désintégration d'un atome radioactif (dans ce cas le processus est un processus « de mort » qui s'interrompt après le premier évènement).

Son espérance mathématique et sa variance sont données par :

$$E(X) = \frac{1}{\lambda}; \quad V(X) = \frac{1}{\lambda^2}$$

Remarque 4.5. *Une erreur à ne pas commettre : la somme de deux variables aléatoires exponentielles indépendantes n'est jamais une variable aléatoire exponentielle. La loi de la somme de n variables aléatoires exponentielles indépendantes de même paramètre λ est une loi d'Erlang de paramètres n et λ .*

4.2.3 Loi normale

La loi normale est certainement la loi de probabilité la plus utilisée. Elle intervient souvent comme loi limite vers laquelle convergent certains modèles : on peut citer la moyenne de n mesures, indépendantes et de même loi, dont la distribution, lorsque n augmente indéfiniment, se rapproche de plus en plus d'une distribution normale. Introduite comme loi "normale des erreurs", la loi normale semble décrire assez bien la distribution de certains caractères biométriques, par exemple la taille d'un individu choisi au hasard dans une population donnée. La loi normale est aussi à l'origine du développement de modèles probabilistes, les modèles gaussiens, dont on abordera l'étude et l'utilisation plus loin : lois du chi-deux, lois de Student, lois de Fisher-Snedecor, etc. Dans un premier temps, on l'utilisera essentiellement comme approximation des lois, discrètes, binomiale et de Poisson.

Définition 4.1. *Une variable aléatoire réelle X , suit une loi de Laplace-Gauss ou loi normale, de paramètres m et σ , si sa densité de probabilité est donnée par :*

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

On dit indifféremment qu'une variable suivant une telle loi est une variable normale ou gaussienne. L'espérance mathématique (ou moyenne) d'une telle variable est $E(X) = m$ et son écart-type est $\sigma(X) = \sigma$.

Si X suit la loi normale $\mathcal{N}(m; \sigma)$ alors la variable aléatoire $U = \frac{X-m}{\sigma}$ satisfait $E(U) = 0$; $\sigma(U) = 1$. C'est la variable aléatoire centrée réduite associée à la variable aléatoire X . La variable aléatoire U suit la loi normale $\mathcal{N}(0; 1)$ appelée loi normale centrée réduite. Sa fonction densité est donnée par :

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

Remarque 4.6. *Les calculs des valeurs à partir de la fonction de densité de la v.a.r. X suivant la loi normale $\mathcal{N}(m; \sigma)$ ne sont pas aisés. On utilise le plus souvent les valeurs tabulées de la fonction de répartition F de la loi normale centrée réduite.*

La fonction F est tabulée le plus souvent pour des valeurs positives de u . Pour les valeurs négatives de u , on utilise l'égalité $F(-u) = 1 - F(u)$. Le calcul des probabilités se fonde sur les propriétés :

- $F(u) = p(U \leq u)$;
- $p(U = u) = 0$ pour tout réel u ;
- $p(u_1 < U < u_2) = F(u_2) - F(u_1)$ pour tous réels u_1 et u_2 .

Exemple 4.7. *La taille d'une population est distribuée normalement avec une moyenne de 165 cm et un écart-type égal à 15 cm. Déterminer la probabilité pour que la taille d'une personne soit :*

1. inférieure à 150 cm;
2. inférieures ou égale à 175 cm;
3. supérieure ou égale à 175 cm;

Proposition 4.3. *Soient X et Y deux variables indépendantes telles que X suit la loi normale $\mathcal{N}(m_1; \sigma_1)$ et Y suit la loi normale $\mathcal{N}(m_2; \sigma_2)$. Alors la variable aléatoire $X + Y$ suit la loi normale $\mathcal{N}(m_1 + m_2; \sqrt{\sigma_1^2 + \sigma_2^2})$.*

Ce résultat se généralise facilement à la somme de n variables aléatoires gaussiennes, indépendantes.

Approximation d'une loi binomiale par une loi normale

Soit X la variable aléatoire, nombre de succès lors de la réalisation de n épreuves indépendantes, la probabilité de succès pour chaque épreuve étant égale à p . La loi de la variable aléatoire :

$$U = \frac{X - np}{\sqrt{np(1-p)}}$$

tend vers la loi $\mathcal{N}(0; 1)$ quand n tend vers l'infini. En d'autres termes, la loi de la variable X tend vers la loi normale $\mathcal{N}(np; \sqrt{np(1-p)})$.

En pratique, cette approximation est valable dès que $n \geq 30, np \geq 15$ et $n(1-p) \geq 15$.

Remarque 4.8. *Deux remarques importantes doivent être faites :*

- *une variable aléatoire binomiale est une variable discrète prenant ses valeurs dans l'intervalle $[0, n]$ alors qu'une variable aléatoire gaussienne est une variable continue prenant ses valeurs dans \mathbb{R} ;*

- *dans le cas d'une loi binomiale, un point a une mesure ou une probabilité non nulle alors que dans le cas d'une loi normale, un point est un ensemble de mesure nulle.*

Pour ces deux raisons, on doit faire une correction de continuité quand on utilise l'approximation d'une loi binomiale par une loi normale (voir exercice).

Approximation d'une loi de Poisson par une loi normale

Soit X une variable aléatoire suivant la loi de Poisson de paramètre λ . La loi de la variable aléatoire :

$$U = \frac{X - \lambda}{\sqrt{\lambda}}$$

tend vers la loi $\mathcal{N}(0; 1)$ quand λ tend vers l'infini. En d'autres termes, la loi de la variable X tend vers la loi normale $\mathcal{N}(\lambda; \sqrt{\lambda})$.

En pratique, cette approximation est valable dès que $\lambda > 18$.

Comme pour la loi binomiale, on doit faire une correction de continuité.

4.2.4 La loi gamma

La loi exponentielle est un cas particulier de la famille des lois gamma. Une variable aléatoire réelle positive X suit une loi gamma $\Gamma(t; \lambda)$, de paramètres strictement positifs t et λ , si sa densité de probabilité est donnée par :

$$\begin{cases} f(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{t-1}}{\Gamma(t)} & \text{si } x \geq 0 \\ f(x) = 0 & \text{sinon} \end{cases}$$

où Γ est la fonction définie pour $t > 0$ par :

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

On a :

$$E(X) = \frac{t}{\lambda}; \quad V(X) = \frac{t}{\lambda^2}$$

Remarque 4.9. *La loi gamma dépendant de deux paramètres peut être utilisée pour représenter un grand nombre de distributions. Ainsi :*

- dans la théorie des files d'attente, la loi gamma représente la loi de probabilité d'occurrence de t événements (t étant un entier), dans un processus poissonnien. Si le temps T , entre les défaillances successives d'un système, suit une loi exponentielle, le temps cumulé d'apparitions de λ défaillances suit une loi gamma $\Gamma(t; \lambda)$,

- en fiabilité, la loi gamma peut être utilisée pour modéliser les temps de défaillance d'un matériel. Selon les valeurs des paramètres, la loi gamma s'identifie à d'autres lois :

- la loi exponentielle si $t = 1$,

- la loi d'Erlang si t est égal à un entier n supérieur à 1, sa densité est :

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

En effet, $\Gamma(n) = (n - 1)!$

Proposition 4.4. *la somme de deux variables aléatoires indépendantes, suivant des lois gamma $\Gamma(t; \lambda)$ et $\Gamma(u; \lambda)$, suit une loi gamma $\Gamma(t + u; \lambda)$.*

Bibliographie

- [1] C. ATINDOGBE, *Cours de statistique et probabilité*, 2016.
- [2] C. OGOUYANDJOU, *Statistique et probabilités, deuxième année*, 2011.
- [3] R. VEYSSEYRE, *Statistique et probabilités pour l'ingénieur*, Dunod, Paris, 2001, 2006.
- [4] F. DRESS, *Les probabilités et la statistique :500 définitions, formules et tests d'hypothèses*, Dunod.