République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mohamed Khider, Biskra

Théorie des Probabilités

Deuxième Année Licence Mathématiques

Présenté par :

Dr. Sayah Abdallah

2021/2022

Sommaire

Introduction

La théorie des probabilités occupe aujourd'hui une place importante dans la plupart des sciences. Tout d'abord, de par ses applications pratiques : en tant que base des statistiques, elle permet l'analyse des données
recueillies lors d'une expérience, lors d'un sondage, etc . . . , et elle possède en outre de nombreuses applications directes, par exemple en fiabilité, ou dans les assurances et la finance. D'un côté plus théorique, elle permet la modélisation de nombreux phénomènes, comme par exemple en économie, climatologie, informatique, hydraulique, en sociologie etc..
Pour la théorie des probabilités, on présente les deux résultats fondamentaux suivants : Les lois faible et forte des grands nombres qui assure la convergence en probabilité et près que sûre de la moyenne empirique des variables aléatoires vers la moyenne théorique quand le nombre d'observations indépendantes augmente, et le théorème central limite (TCL) qui précise la vitesse de cette convergence.
Le contenu de ce cours est le suivant:
Le premier chapitre est consacré à l'anlayse combinatoire qui est une branche des mathématiques qui étudie comment compter les objets. Elle fournit des méthodes de dénombrements particulièrement utiles en théorie des probabilités, et elle sert d'outil dans plusieurs problèmes élémentaires en théorie des probabilités, domaine des mathématiques qui trouve son origine dans l'étude des jeux de hasard.
Le second chapitre est consacré à une introduction sur la théorie des probabilités où Il aborde quelques notions de l'espace de probabilités et Il présente également les notions de l'expérience aléatoire et la définition des évènements, l'algèbre et la σ algèbre ainsi que les probabilités conditionnelles et indépendance.
Le troisième chapitre est consacré à l'étude des variables aléatoires, ainsi que leurs types (discrètes, continues) et leurs caractéristiques par exemple les tendances centrales et les paramètres de dispersion.
Dans le quatrième chapitre nous étudions les couples aléatoires, sa fonction de distribution ainsi que la loi de probabilité d'un couple discret ou continue.
Le cinquième chapitre est consacré à une large définition et étude de lois de variables aléatoires discrétes et continues
Le dernier chapitre est consacré à une présentation de différents modes de convergence de suites de variables aléatoires, comme la convergence en probabilité, en loi, présque sûre, en moyenne d'ordre p, et avec en particulier les lois faible et forte des grands nombres, le théorème central limite.

Analyse Combinatoire

Exemples
1) On achète une valise à code de 4 chiffres
Combien on a de possibilités de choisir un code.
Solution
Soit m le nombre de chiffres dans un code, on a: $m = 4$.
On a 10 possibilités de choisir le 1^{er} chiffre, le second, le $3^{ème}$ et le $4^{ème}$.
donc le nombre total de codes possibles est : $10 \times 10 \times 10 \times 10 = 10^4$.
2) On veut garer 2 voitures sur un parking de 4 places.
Combien de possibilités de garer ces 3 voitures.
Solution
Le nombre de possibilités de garer ces 3 voitures est: $4 \times 3 = 12$
3) Les plaques d'immatriculation de voitures dans un pays sont formées
de 3 lettres et 3 chiffres.
Quel est le nombre de plaques d'immatriculation possibles
a) Si que le nombre de lettres dans l'alphabet est 26 ?
b) Si les chiffres et les lettres sont tous différents.
Solution
a) Le nombre est : $26^3 \times 10^3$.

b) Le nombre est : $26 \times 25 \times 24 \times 10 \times 9 \times 8$.
Définition:
L'analyse combinatoire comprend un ensemble de méthodes qui permettent de déterminer le nombre de tous les résultats possibles d'une expérience particulière.
On distingue dans l'analyse combinatoire 3 types de dénombrements:
Les arrangements, les permutations et les combinaisons.

Les arrangements

Définition:
Étant donné un ensemble de n objets, on appelle arrangement de p objets parmi n le nombre de choix de ses p objets parmi n
On a 2 types d'arrangements
1er type est l'arrangement sans répétition

Arrangement sans répétition

Définition:
Un arrangement sans répétition est un arrangement de p objets choisis parmi n, peut être obtenu en tirant d'abord un objet parmi les n, puis un deuxième parmi les $(n - 1)$ *restants, ect...*

Le nombre d'arrangements sans répétition de p éléments pris dans un ensemble à n éléments est :

$$A_n^p = \frac{n!}{(n-p)!} = n \times (n-1) \times \ldots (n-p+1),$$

où pour tout $n \in \mathbb{N}$, on a: $n! = n\,(n-1)\,(n-2)\,(n-3)\ldots \times 2 \times 1$,
avec $0! = 1$,

Exemples

1) Les arrangements sans répétition à 2 éléments de l'ensemble$(1, 2, 3)$ sont:
$(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)$.
Le nombre d'arrangements est $A_3^2 = \frac{3!}{(3-2)!} = 3 \times 2 \times 1 = 6$
2) Combien y'a t'il de façons d'asseoir 4 personnes sur un banc qui ne porte que 3 places
On a 4 possibilités d'asseoir la 1ere personne, et 3 pour la seconde, 2 pour la troisième , donc on a
$A_4^3 = \frac{4!}{(4-3)!} = 4 \times 3 \times 2 = 24$
2éme type est l'arrangement avec répétition

Arrangement avec répétition

Définition:
Un arrangement avec répétition est un arrangement où chaque élément peut-être répété jusqu'à p fois.
Le nombre total de tels arrangements est donc :

$$A_n^p = n^p,\ 1 \leq p \leq n.$$

Exemple:
Le nombre total d'arrangements avec répétition d'ordre 2 de l'ensemble
$(1, 2, 3)$ sont:$(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)$.
Le nombre total d'arrangements est $A_3^2 = 3^2 = 9$.

Les Permutations

On a 2 types de permutations
1er type est la permutation sans répétition

Permutation sans répétition:

Définition:
Une permutation sans répétition ou permutation de n éléments distincts est un arrangement de $p = n$ éléments pris parmi n.
Le nombre de permutations de n éléments distincts est

$$A_n^n = n! = n\,(n-1)\ldots 1.$$

Exemples:

1) Combien de nombres de 3 chiffres peut-on former avec $1, 3$ et 5.
Le nombre des nombres de 3 chiffres formés avec $1, 3$ et 5 est $A_3^3 = 3! = 6$
$135, 153, 315, 513, 351, 531$.
2) De combien de façons différentes peut-on distribuer 4 cadeaux à
4 personnes.
Chaque personne ne peut recevoir q'un seul cadeau.
Le 1ér cadeau peut être donné à un des 4 personnes et le 2ème à un des 2 personnes etc..., donc le nombre total est $A_4^4 = 4! = 24$.
2éme type est la permutation avec répétition

Permutation avec répétition:

Définition:
Considérons un ensemble de n éléments divisés en p groupes d'éléments identiques, de taille $n_1, n_2, \ldots$,*et* n_p *éléments respectivement (avec* $n_1 + n_2 + \ldots + n_p = n$*).*
Alors le nombre de permutations avec répétition de cet ensemble est :

$$P_n = \frac{n!}{n_1!n_2!\ldots n_p!}.$$

Exemples:
1) Combien de mots (sans tenir compte du sens) peut- on former avec le mot papa
Les lettres du mot papa constituent un ensemble de 4 lettres ($n = 4$), cet ensemble est divisé en 2 groupes de lettres le 1er et le
second groupe contient chacun 2
Le nombre de mots possibles est $P_5 = \dfrac{4!}{2!2!} = 6$.

Combinaison

Définition:
Étant donné un ensemble de n objets (distincts), alors
le nombre de combinaison de p éléments choisis parmi n éléments est donné par:

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

Exemples :
1) Les combinaisons de 2 éléments pris dans $\{1, 2, 3, 4\}$ sont $\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}$.
$C_4^2 = \dfrac{4!}{2!2!} = \dfrac{A_4^2}{p!} = 6$
Le nombre de $p!$ combinaisons de p éléments parmi n éléments correspond à l'arrangement de p objets choisis parmi n

2) De combien de manières peut-on choisir une délégation de 3 hommes et 2 femmes pris parmi un groupe de 7 hommes et 5 femmes .

Il y a $C_7^3 = \dfrac{7!}{3!\,(7-3)!} = 35$ manières de choisir 3 hommes et $C_5^2 = \dfrac{5!}{2!\,(5-2)!} = 10$ manières de choisir 2 femmes.

En appliquant le principe fondamental de dénombrement on obtient $C_7^3 \times C_5^2 = 35 \times 10 = 350$ manières de choisir une délégation de 3 hommes et 2 femmes.

3) Une urne contient 10 boules dont 4 blanches et 6 noires. On en tire 4 boules.

a) Quel est le nombre de tirages possibles.

b) De combien de façons peut-on tirer :

4 boules blanches.

2 boules blanches et 2 noires.

Le nombre de tirages possibles est $C_{10}^4 = \dfrac{10!}{4!\,(10-4)!} = \dfrac{10 \times 9 \times 8 \times 7}{24} = 210.$

a) Le nombre de façons de tirer 4 boules blanches est $C_4^4 \times C_6^0 = 1$

b) Le nombre de façons de tirer 2 boules blanches et 2 noires est $C_4^2 \times C_6^2 = \dfrac{4 \times 3}{2} \times \dfrac{6 \times 5}{2} = 90.$

4) On considère un groupe de 10 personnes. Si chaque personne serre la main de toutes les autres, combien y a-t-il de poignées de main ?

La première peut serrer la main à 9 personnes, la deuxième à 8, la troisième à 7 etc.donc $9 + 8 + 7 + \ldots + 3 + 2 + 1 = \frac{10 \times 9}{2} = 45$ poignées de mains.

On a une combinaison de 2 éléments dans un ensemble de 10 éléments.

Le nombre de poignées de mains est $C_{10}^2 = \dfrac{10 \times 9}{2} = 45.$

Espace de probabilités

Historiquement, la théorie des probabilités s'est développée à partir du XVIIe siècle autour des problèmes de jeux dans des situations où le nombre de cas possibles est fini dont elle fournit des modèles mathématiques permettant l'étude d'expériences dont le résultat ne peut être prévu avec une totale certitude.

Expérience aléatoire et événements

Expérience aléatoire

Toute expérience qu'on ne peut pas connaitre son résultat par avance lorsqu'on répéte l'expérience dans les mˆemes conditions s'appelle expérience aléatoire.
Exemple :
Le jet d'un dé, ainsi que l'extraction d'une carte d'un jeu sont des expériences aléatoires.

Espace des évènements

Dans une expérience aléatoire, une proposition relative au résultat de cette expérience s'appelle un évènement par exemple dans le jet d'une pièce de monnaie on a deux résultats possibles pile où face. L'ensemble de tous les résultats possibles d'une expérience aléatoire s'appelle espace des évènements ou encore ensemble fondamental, et on le notera par Ω.
Exemple :
Lorsqu'on jette un dé alors l'ensemble fondamental est $\Omega = \{1; 2; 3; 4; 5; 6\}$, et quand on jette 2 dés alors: $\Omega = \{(1; 2) , (1; 3) \ldots\text{etc}\}$.
A tout évènement A, on l'associe son opposé noté $\bar{A}$ tel que la réalisation de A exclue la réalisation de l'autre et réciproquement. L'évènement $\bar{A}$ représente dans Ω la partie complémentaire de A.

Evènement élémentaire et composé

Définition :
Un évènement élémentaire est un sous ensemble de Ω qui a un seul élément, et un évènement compose´ est un ensemble d'´el´ements ´el´ementaires.
Exemple :
Dans le jet d'un dé, $\{2\}$ est un évènement élémentaire, avoir un chiffre pair est un évènement composé $\{2; 4; 6\}$

Algébre et σ-algébre

L'ensemble fondamental Ω ou ensemble des résultats possibles d'une expérience aléatoire soit il est:
1) Fini: contient un nombre fini d'éléments.
2) Infini dénombrable.

3) Continu (tout intervalle de $\mathbb{R}$).
On d´efinie l'alg`ebre dans le cas fini et la $\sigma$-alg`ebre dans les deux autres cas.
Définition 1 **:**
Soit $\mathcal{A}$ une classe de parties de Ω, on dit que $\mathcal{A}$ est une algèbre d'évènements sur Ω si elle vérifie

$$\begin{array}{ll} 1) & \Omega \in \mathcal{A} \\ 2) & \forall A \in \mathcal{A} \Longrightarrow \bar{A} \in \mathcal{A} \\ 3) & \forall A, B \in \mathcal{A} \Longrightarrow A \cup B \in \mathcal{A}. \\ 4) & \forall A, B \in \mathcal{A} \Longrightarrow A \cap B \in \mathcal{A}. \end{array}$$

De cette définition on peut déduire que:

$$\begin{array}{ll} 1) & \bar{\Omega} \in \mathcal{A}. \\ 2) & \forall A_i \ (i \in I \text{ fini}) : \bigcap\limits_{i \in I} A_i \in \mathcal{A}. \\ 3) & \forall A_i \ (i \in I \text{ fini}) : \bigcup\limits_{i \in I} A_i \in \mathcal{A}. \end{array}$$

On définit également l'évènement certain qui est l'ensemble fondamental Ω, et l'évènement impossible repr´esente´ par $\emptyset$.
Exemple 1:
On cite quelques exemples d'algèbres:
- La famille $\{\Omega; \emptyset\}$ est une algèbre (algèbre triviale).
- Soit Ω un espace des évènements, et soit A un évènement, alors l'ensemble $\{\Omega; A; \bar{A}; \emptyset\}$ est une algèbre sur Ω appelée algèbre de Bernoulli.
- L'ensemble de parties de Ω note´ $\mathcal{P}(\Omega)$ est une alg`ebre sur Ω.
D´efinition 2:
On appelle σ-algèbre ou tribu sur un espace Ω toute famille $\mathcal{T}$ qui vérifie:

$$\begin{array}{ll} 1) & \Omega \in \mathcal{T} \\ 2) & \forall A \in \mathcal{T} \Longrightarrow \bar{A} \in \mathcal{T} \\ 3) & \forall A_i \in \mathcal{T} \Longrightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{T} \end{array}$$

On peut d´eduire de cette d´efinition qu'une σ-alg`ebre est une algébre.
Exemple 2:
$\mathcal{P}(\Omega)$ est une σ-algébre.

Espace de probabilités

Soit Ω un espace d'évènements et $\mathcal{T}$ une tribu sur Ω, alors le couple $(\Omega; \mathcal{T})$ s'appelle espace probabilisable.

Définition 01:

Les évènements A, B sont dits incompatibles si la réalisation de l'un exclu la réalisation de l'autre, autrement dit $A \cap B = \varnothing$

D´efinition 02:

Soient les évènements non vides $A_1, A_2, ...A_n$, alors on dit que ces évènements forment un système complet si

$$1) \quad A_i \cap A_j = \varnothing \ \ \forall i \neq j$$

$$2) \quad \bigcup_{i=1}^{n} A_i = \Omega.$$

D´efinition 03:

On appelle probabilité sur $(\Omega; \mathcal{T})$ une application P de $\mathcal{T}$ dans $[0, 1]$ vérifiant

$$P(\Omega) = 1,$$

et pour tout ensemble dénombrable d'évènements incompatibles $A_1, A_2, ...A_n...$ on a:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \quad \text{(axiome } \sigma \text{ additivité).}$$

Dans le cas o`u Ω est fini l'axiome est dit axiome d'additivité.

Définition 04:

Soit $(\Omega; \mathcal{T})$ un espace probabilisable, alors si on définit sur cet espace une probabilité P, l'espace $(\Omega, \mathcal{T}, P)$ est dit espace de probabilité.

Propriétés:

Soient $(\Omega, \mathcal{T}, P)$ un espace de probabilité et $A, B \in \mathcal{T}$ alors on a:

$$\begin{array}{llcl}
1) & P(\bar{A}) & = & 1 - P(A) \\
2) & 1 = P(\bar{A} \cup A) & = & P(\bar{A}) + P(A) \\
3) & \text{Si } A \subset B & \Longrightarrow & P(A) \leq P(B) \ \text{(inégalité de Boole)} \\
4) & P(A \cup B) & = & P(A) + P(B) - P(A \cap B). \\
5) & \forall A \in \mathcal{T} & \Longrightarrow & P(A) = \sum_{i=1}^{n} P(A \cap B_i),
\end{array}$$

où $B_1, B_{2,}..., B_n$ est un système complet d'évènements.

Probabilité conditionnelle et indépendance

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilit´e, on consid´ere deux ´ev`enements $A, B \in \mathcal{T}$ tel que $P(B) > 0$.

Définition :

On appelle probabilité conditionnelle de A sachant l'évènement B notée $P(A/B)$ ou $P_B(A)$ la quantité

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

On vérifie que $P_B(A)$ est une probabilité sur $(\Omega, \mathcal{T})$
1)
$$P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = 1.$$
2) On a $P(A \cap B) \geq 0$ et $P(B) > 0$ alors $P_B(A) \geq 0$, et on a de plus $A \cap B \subset B$ ce qui implique que $P(A \cap B) \leq P(B)$ d'où $P_B(A) \leq 1$.
3) Si les les évènements $A_i,\ i \in \mathbb{N}^*$ sont incompatibles, donc les évènements $A_i \cap B$ les sont aussi et on a:

$$\begin{aligned} P_B\left(\bigcup_{i=1}^{\infty} A_i\right) &= P\left(\bigcup_{i=1}^{\infty} A_i / B\right) = \frac{P\left(\bigcup_{i=1}^{\infty} (A_i \cap B)\right)}{P(B)} \\ &= \frac{\sum_{i=1}^{\infty} P((A_i \cap B))}{P(B)} = \sum_{i=1}^{n} P(A_i/B). \end{aligned}$$

Conséquences:
On déduit d'une manière analogue que
1)
$$P(B/A) = \frac{P(A \cap B)}{P(A)},$$
d'où

$$P(A \cap B) = P(A/B)\,P(B) = P(B/A)\,P(A).$$

2) D'une façon générale si on a n évènements quelconques $A_1, A_2, ...A_n$, alors

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)\,P(A_2/A_1)\,P(A_3/A_2 \cap A_1)\,...P(A_n/A_1 \cap A_2 \cap ... \cap A_{n-1}).$$

Exemple :
Dans un jeu de 32 cartes, on tire successivement 2 cartes. Quelle est la probabilité d'avoir un as au second tirage sachant qu'on a obtenu un as au premier tirage.
Soient A et B les évènements: A avoir un as au premier tirage et B avoir un as au second tirage, et on cherche $P(B/A)$.
On a $P(A) = \frac{4}{32} = \frac{1}{8}$ et $P(A \cap B) = \frac{4}{32}.\frac{3}{31}$ alors $P(B/A) = \frac{3}{31}$.

Formule des probabilités totales

Théorème 1:
Soient $(\Omega, \mathcal{T}, P)$ *un espace de probabilité et* $A_1, A_2..., A_n$ *un système complet d'évènements de* Ω *tel que* $\forall 1 \leq i \leq n,\ P(A_i) > 0$, *et soit* A *un évènement de* $\mathcal{T}$ *alors on a:*

$$P(A) = \sum_{i=1}^{n} P(A_i)\, P(A/A_i)$$

Preuve:
On a

$$A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} (A \cap A_i).$$

Puisque les évènements $A_1, A_2..., A_n$ sont incompatibles ce qui entraine que

$$P(A) = P\left(\bigcup_{i=1}^{n} (A \cap A_i) \right) = \sum_{i=1}^{n} P(A \cap A_i) = P(A_i)\, P(A/A_i).$$

Cette formule est appel´ee formule des probabilit´es totales.
Exemple :
Dans une usine quatre machines M_1, M_2, M_3, M_4 produisent respectivement 20%, 30%, 15% et 35% de pièces avec 3%, 2%, 4% et 5% de produits défectueux.
Quelle est la probabilité qu'une pièce tirée au hasard soit défectueuse.
Soient les évènements M_i (la pièce tirée provient de la machine M_i $(1 \leq i \leq 4)$) est D (la pièce est défectueuse) et on cherche $P(D)$. On a

$$D = \bigcup_{i=1}^{4} (D \cap M_i),$$

d'où

$$\begin{aligned} P(D) &= P(D/M_1)\, P(M_1) + P(D/M_2)\, P(M_2) + P(D/M_3)\, P(M_3) + P(D/M_4)\, P(M_4) \\ &= 0,2 \times 0,03 + 0,3 \times 0,02 + 0,15 \times 0,04 + 0,35 \times 0,05 \\ &= 0,0355 \end{aligned}$$

Formule de Bayes

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilité, et on considére deux évènements $A,\ B \in \mathcal{T}$ tel que $P(A) > 0$ et $P(B) > 0$.
Comme

$$P(B/A) = \frac{P(A \cap B)}{P(A)},$$

alors on a

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

La formule de Bayes est
$$P(A/B) = \frac{P(B/A)\,P(A)}{P(B)}.$$
On remarque que
$$P(B) = P(A \cap B) + P\left(B \cap \bar{A}\right) = P(B/A)\,P(A) + P\left(B/\bar{A}\right)P\left(\bar{A}\right),$$
ainsi
$$P(A/B) = \frac{P(B/A)\,P(A)}{P(B/A)\,P(A) + P\left(B/\bar{A}\right)P\left(\bar{A}\right)}.$$
Dans le cas général on énonce le théorème suivant

Théorème 2:

Soient $A_1, A_2..., A_n$ un système complet d'évènements relatifs à un évènement quelconque A qui s'est réalisé alors

$$\begin{aligned} P(A_i/A) &= \frac{P(A_i)\,P(A/A_i)}{P(A)} \\ &= \frac{P(A_i)\,P(A/A_i)}{\sum_{i=1}^{n} P(A_i)\,P(A/A_i)}. \end{aligned}$$

Preuve:

On a:
$$A = \bigcup_{i=1}^{n} (A \cap A_i)$$
ce qui implique que
$$\begin{aligned} P(A) &= P\left(\bigcup_{i=1}^{n} (A \cap A_i)\right) \\ &= \sum_{i=1}^{n} P(A \cap A_i), \end{aligned}$$
et d'après une conséquence de la probabilité conditionnelle on a
$$P(A \cap A_i) = P(A)\,P(A_i/A) = P(A_i)\,P(A/A_i),$$
d'où le résultat.

Ce th´eor`eme permet d'´evaluer les probabilit´es des diff´erents ´ev`enements $A_1, A_2..., A_n$ qui peuvent causer la réalisation de l'´ev`enement A.

Exemple :

Si on reprend l'exemple 7 en supposant que la pièce tirée est défectueuse, alors la question qui se pose est de quelle machine provient -elle?

On cherche par exemple $P(M_2/D)$.

D'aprés la formule de Bayes

$$P(M_2/D) = \frac{P(M_2)\,P(D/M_2)}{P(D)} = \frac{0,3 \times 0,02}{0,0355} = 0,169$$

Evénements indépendants

Définition :
Soient A et B deux évènements d'un espace de probabilité $(\Omega, \mathcal{T}, P)$ tel que $P(A) > 0$ et $P(B) > 0$, alors on dit que A est indépendant de B si

$$\begin{aligned} P(A/B) &= P(A) \\ P(B/A) &= P(B). \end{aligned}$$

On dit encore que A et B sont indépendants si et seulement si la probabilité de la réalisation simultanée de ces événements est égale au produit de leurs probabilités individuelles

$$P(A \cap B) = P(A)\,P(B)$$

Ce qui signifie que la probabilité de la réalisation de l'évènement A n'est pas influée par la réalisation de B et inversement.
Exemple :
On lance une pièce de monnaie trois fois de suite, donc l'ensemle des résultats possibles est $\Omega = \{ppp, ppf, pff, fff, ffp, fpp, pfp, fpf\}$, où f signifie que le résultat du lancer est face et p est pile.
Soint les évènements: A avoir face au premier jet, B avoir face deux fois de suite et C avoir face au second jet.
Etudier l'indépendance de ces évenements.
On a:

$$A = \{fff, ffp, fpf, fpp\}, \; B = \{pff, fff, ffp\} \text{ et } C = \{pff, fff, ffp, pfp\},$$

d'où

$$P(A) = 1/2, \; P(B) = 3/8 \text{ et } P(C) = 1/2,$$

et

$$\begin{aligned} P(A \cap C) &= 1/4 = P(A)\,P(C) \\ P(A \cap B) &= 1/4 \neq P(A)\,P(B) \\ P(B \cap C) &= 3/8 \neq P(B)\,P(C). \end{aligned}$$

Ce qui r´esulte que seulement les ´ev`enements A et C qui sont ind´ependants.
Théoréme :
Soient A et B deux évènements d'un espace de probabilité $(\Omega, \mathcal{T}, P)$ tel que $P(A)$ et $P(B) > 0$, alors les propriétés suivantes sont équivalentes

1) Les évènements A et B sont indépendants.
2) Les évènements $\bar{A}$ et B sont indépendants.
3) Les évènements A et $\bar{B}$ sont indépendants.
4) Les évènements $\bar{A}$ et $\bar{B}$ sont indépendants.

Indépendance de plusieurs événements

D´efinition :
Soient A_i $(1 \leq i \leq n)$ n évènements d'un espace de probabilité $(\Omega, \mathcal{T}, P)$, alors on dit que les évènements sont totalement indépendants si

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i),$$

et le nombre de relations à vérifier pour l'indépendance totale de n évènements est $(2^n - n - 1)$.

Formule de Poincarré

Soient A, B et C trois évènements liés à une expérience aléatoire, alors la probabilité de la réunion de deux et trois évènements sont respectivements

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

et

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) \\ &- P(A \cap C) - P(C \cap B) + P(A \cap B \cap C). \end{aligned}$$

Généralisation à n évènements
Soient $A_1, A_2, ..., A_n$, n évènements quelconques d'un espace de probabilité $(\Omega, \mathcal{T}, P)$ liés à une expérience aléatoire, alors:

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} (-1)^{k-1} S_k,$$

où

$$S_k = \sum_{1 \leq j_1 \leq j_2 \leq ... \leq j_k \leq n} \sum ... \sum P\left(A_{j_1} \cap A_{j_2} \cap ... \cap A_{j_k}\right).$$

La quantité S_k est la somme des probabilités de toutes les intersections possibles de n évènements. On utilise la formule de Poincarré pour calculer la probabilité de la réalisation d'au moins un évènement parmi les n.

Variables aléatoires

Considérons le lancer de deux dés, cette expérience nous donne l'ensemble des résultats $\Omega = \{(1,1),(1,2),...,(6,6))\}$, on remarque que $\forall\omega \in \Omega$ on a $P(\omega) = \frac{1}{36}$.
On s'intérèsse maintenant à la somme marquée par les deux dés, donc on a défini une application: S de Ω vers l'ensemble $E = \{2,3,...,12\}$, par exemple $P(S=5) = \frac{1}{9}$.
En général on a: $P(S=s) = P\left(S^{-1}(s)\right)$.
On remarque que cette somme dépend de valeurs aléatoires, donc il s'agit d'une variable aléatoire.
Une variable aléatoire X est une application de $(\Omega, \mathcal{T}, P)$ dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, où $\mathcal{B}_{\mathbb{R}}$ est une tribu ou la σ-alg`ebre engendr´ee par les intervalles de $\mathbb{R}$.
Définition :
Soient $(\Omega, \mathcal{T}, P)$ *un espace de probabilité, une variable aléatoire reèlle* X *est une application:*

$$X : (\Omega, \mathcal{T}, P) \longrightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}}),$$

ceci signifie que $\left(\forall B \in \mathcal{B}_{\mathbb{R}} : \ X^{-1}(B) \in \mathcal{T}\right)$, et la variable X est appelée variable aléatoire reèlle.
Soit l'ensemble $A_x = \{\omega \in \Omega : X(\omega) \leq x\}$, alors X est une variable aléatoire réelle si et seulement si $A_x \in \mathcal{T}$.

Exemple :
Soit

$$\begin{array}{rccc} X : & (\Omega, \mathcal{T}, P) & \rightarrow & (\mathbb{R}, \mathcal{B}_{\mathbb{R}}) \\ & \omega & \rightarrow & X(\omega), \end{array}$$

où

$$X(\omega) = \begin{cases} 1 & \text{si} \quad \omega \in A \\ 0 & \text{si} \quad \omega \in \bar{A}. \end{cases}$$

Si

$$x < 0 \Longrightarrow A_x = \varnothing \in \mathcal{T},$$

et si

$$0 \leq x < 1 \Longrightarrow A_x = \bar{A} \in \mathcal{T}, \text{ et si } x \geq 1 \Longrightarrow A_x = A \cup \bar{A} \in \mathcal{T},$$

donc l'application est une variable aléatoire.

Fonction de répartition

Considérons l'application
$$\begin{array}{rccc} P: & \mathbb{R} & \rightarrow & [0,1] \\ & x & \rightarrow & P(A_x), \end{array}$$

où

$$P(A_x) = P(\omega \in \Omega, X(\omega) \leq x),$$

cette application est appelée fonction de répartition ou fonction de distribution de la variable X et on la notera $F(x) = P(X \leq x)$.

Exemple :
Si on prend l'exemple 10 on a si

$$x < 0 \Longrightarrow A_x \Longrightarrow F(x) = P(X \leq x) = P(\varnothing) = 0,$$

et si

$$0 \leq x < 1 \Longrightarrow F(x) = P(\bar{A}) = 1 - p, \text{ où } p = P(A),$$

enfin si

$$x \geq 1 \Longrightarrow F(x) = P(A \cup \bar{A}) = P(\Omega) = 1.$$

Propriétés de la fonction de répartition

Soit X une variable aléatoire réelle et $F(x)$ sa fonction de répartition alors:
1) Si $x < x' \Longrightarrow F(x) \leq F(x')$, la fonction F est une fonction monotone croissante.
2) $\lim\limits_{x \to b_0^-} F(x) = F(b_0)$, la fonction F est une fonction continue à gauche en tout point.
3) $\lim\limits_{x \to -\infty} F(x) = 0$ et $\lim\limits_{x \to +\infty} F(x) = 1$.
4) On a: $F(x) = P(]-\infty, x]) = P\left(X^{-1}\,]-\infty, x]\right)$, d'où $P(a < x \leq b) = F(b) - F(a)$.

Type de variables al´eatoires

Variable discréte

Définition :
Si la variable al´eatoire X prend ses valeurs dans un ensemble fini ou infini d´enombrable (son ensemble de d´efinition est inclus dans $\mathbb{N}$), on dit que X est une variable al´eatoire discréte.

Exemple :
Citons quelques exemples de variables aléatoires discrètes
1) Nombre de "pile" dans un lancer de 2 pièces : $X(\omega)$ prend les valeurs de 0 à 2.
2) Nombre de lancers de deux dés avant d'obtenir la paire $(5,5)$: $X(\omega)$ prend les valeurs de 0 à l'infini.
3) Nombre d'appels arrivants à un standard téléphonique en une minute: $X(\omega)$ prend les valeurs de 0 à 5 par exemple.

Pour ce type de variables on définit en chaque point $x_i \in I = \{x_0, x_1, ...x_k...\}$ la fonction poids ou la fonction de masse par:

$$P(X = x_i) = P(x_i) \ \forall x_i \in I,$$

vérifiant

$$0 \leq P(x_i) \leq 1 \text{ et } \sum_{x_i \in I} P(x_i) = 1.$$

Variable aléatoire absolument continue

Définition 01:
Une variable al´eatoire est dite absolument continue ou continue si elle peut prendre toutes les valeurs d'un intervalle de $\mathbb{R}$.
Dans ce cas, il ne s'agira plus de calculer une probabilite´ d'une valeur donn´ee mais d'un intervalle.
Exemple 1:
Citons quelques exemples de variables al´eatoires continues
1) Temps d'attente pour prendre un avion :$X(\omega) \in [0, 20 \text{ min}]$
2) Moyenne de poids de 25 étudiants pris aléatoirement : $X(\omega) \in [a, b]$.
Autre d´efinition d'une variable al´eatoire absolument continue.
D´efinition 02:
On dira qu'une variable aléatoire reélle est absolument continue s'il existe une fonction positive f définie pour tout reél x et toute partie $\mathcal{B}$ de $\mathbb{R}$ vérifiant:

$$\begin{aligned}
1) \quad P(X \in \mathcal{B}) &= \int_B f(x)\, dx. \\
2) \quad P(X \in]-\infty, \infty[) &= \int_{-\infty}^{\infty} f(t)\, dt. \\
3) \quad P(a \leq X \leq b) &= \int_a^b f(t)\, dt. \\
4) \quad P(X = a) &= \int_a^a f(t)\, dt = 0.
\end{aligned}$$

Fonction de répartition

Définition:
La fonction de répartition d'une variable aléatoire discrète X est définie par:

$$F(x) = \sum_{x_i \leq x} P(x_i).$$

En paticulier pour $x_1 < x_2 < ... < x_n$ on a:

$$F(x) = \begin{cases} 0 & \text{si} \quad x < x_1 \\ P(x_1) & \text{si} \quad x_1 \leq x < x_2 \\ P(x_1) + P(x_2) & \text{si} \quad x_2 \leq x < x_3 \\ . & \\ 1 & \text{si} \quad x \geq x_n. \end{cases}$$

(Dans ce cas, F est une fonction en escaliers, continue à gauche, ayant pour limite 0 en $-\infty$ et 1 en $+\infty$).

Propriétés:

1) F est non décroissante.
2) F est continue à gauche.
3) $\forall x_0 \in \mathbb{R},\ P(X = x_0) = \lim\limits_{x \to x_0} F(x) - F(x_0)$.
4) $P(a \leq X < b) = F(b) - F(a)$.
5) F est continue à droite dans le cas des variables aléatoires continues.

Définition :

On appelle densité de probabilité f d'une variable aléatoire absolument continue la dérivée de la fonction F tel que

$$f(x) = F'(x) \ \ et\ F(x) = \int_{-\infty}^{x} f(t)\, dt.$$

On peut déduire de la définition 16 que:

$$P(a \leq X \leq b) = P(a < X < b) = F(b) - F(a),$$

et la fonction F est dérivable et admet pour dérivée une fonction f appelée densité.

On remarque que la fonction densité f satisfait les deux conditions suivantes:

$$\begin{array}{ll} 1) & f(x) \geq 0\ \forall x \in \mathbb{R}. \\ 2) & \int_{\mathbb{R}} f(x)\, dx = 1. \end{array}$$

Exemple :

Soit X la variable aléatoire absolument continue de fonction de répartition $F(x)$ définie par:

$$F(x) = \begin{cases} 0 & \text{si} \quad x < a \\ \frac{x-a}{b-a} & \text{si} \quad a \leq x < b \\ 1 & \text{si} \quad x \geq b. \end{cases}$$

La fonction F est continue sur $\mathbb{R}$, et la fonction densité associée à cette fonction de répartition est définie par:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si} \quad a < x < b \\ 0 & \text{sinon.} \end{cases}$$

On remarque que la fonction densité est discontinue en a et b.
Remarque1:
La fonction densité peut être discontinue en certains points.

Caractéristiques d'une variable aléatoire

Tendance centrale

Quantile

Définition :
On appelle quantile d'ordre α $(0 \leq \alpha \leq 1)$ d'une variable aléatoire X de fonction de répartition $F(x)$ la ou les valeurs x_α tel que $F(x_\alpha) = \alpha$, ce qui équivalent à $P(X \leq x_\alpha) = \alpha$.
Remarque 2:
Si la variable aléatoire X est discrète, l'équation $F(x_\alpha) = \alpha$ admet une infinité de solutions ou aucune, et si la variable aléatoire X est absolument continue et si la fonction $F(x)$ est strictement monotone alors les quantiles de tout ordre existent et sont solutions de l'´equation $F(x_\alpha) = \alpha$.
Exemple :
Soit X la variable aléatoire absolument continue de fonction de répartition $F(x)$ définie par:

$$F(x) = \begin{cases} 1 - \exp(-x) \text{ si } x > 0 \\ 0 \qquad\qquad \text{sinon.} \end{cases}$$

L'équation $F(x_\alpha) = 0,2$ admet une solution $x_\alpha = F^{-1}(0,2)$.
$1 - \exp(-x_\alpha) = 0,2 \Longrightarrow x_\alpha = -\ln 0,8$.
Nous énoncons quelques quantiles particuliers.

Médiane

Définition :
La médiane est la valeur de x pour laquelle $F(x) = 1/2$.

Dans le cas continu la médiane est la valeur x vérifiant $\int_{-\infty}^{x} f(t)\,dt = \int_{x}^{\infty} f(t)\,dt = 1/2$, et dans le cas discrèt, la médiane n'existe pas toujours.
La médiane divise la population en deux parties égales, c'est une caractéristique de tendance centrale.

Les quartiles

Définition :

Les quartiles notés Q_i pour $i = 1; 2; 3$ correspondent au quantile d'ordre α égal à $1/4;\ 1/2;\ 3/4$ respectivement.

Les déciles

Définition : *Le k ^{ime} d´ecile ($k = 1$ à 9) est le quantile d'ordre $k/10$.* En particulier, le 5^{ime} d'ecile correspond à la médiane.

Le Mode:

Définition :

On appelle mode (valeur la plus probable)) d'une variable al´eatoire X la valeur x_m tel que la onction de masse ou la fonction densité soit maximale.

Parfois dans des distributions il y' a plusieurs modes, dans ce cas on dit que la distribution est multimodale.

Espérance mathématique

L'espérance mathématique d'une variable aléatoire est une notion trés importante en probabilités et statistique, cette esp´erance ou moyenne est une valeur unique qui joue le rˆole de repr´esentation de la moyenne des valeurs de X, pour cela elle est appelée souvent mesure de tendance centrale.

Définition :

Soit X une variable aléatoire, alors l'espérance ou la moyenne notée $E(X)$ ou μ est définie par:

$$E(X) = \sum_i x_i P(X = x_i),$$

si X est une variable aléatoire discrète de fonction de masse $P(X = x)$, cette quantité n'est définie que si la série de terme général $(x_i P(X = x_i))$ converge, et

$$E(X) = \int_{\mathbb{R}} x f(x)\, dx,$$

si X est une variable aléatoire absolument continue de fonction densité $f(x)$, et *cette quantité n'existe que si $\int_{\mathbb{R}} x f(x)\, dx$ est absolument convergente.*

$E(X)$ est la moyenne arithm´etique des diff´erentes valeurs de X pond´er´ees par leurs probabilités.

Remarque :

L'espérance math´ematique n'existe pas toujours.

Exemple :

Soit X une variable al´eatoire absolument continue de fonction densite´ $f(x)$ d´efinie par:

$$f(x) = \frac{1}{\pi\left(1+x^2\right)},\ x \in \mathbb{R},$$

alors $E(X) = \int\limits_{-\infty}^{\infty} xf(x)\,dx$ qui est une intégrale divergente, donc l'espérance n'existe pas.

Propriétés de l'espérance mathématique

1)

$$E\left(\sum_{i=1}^{n}\alpha_i X_i\right) = \sum_{i=1}^{n}\alpha_i E(X_i)\ \ \forall\alpha_i \in \mathbb{R}.$$

2) Si X est une variable aléatoire discrète.de fonction de masse $P(X=x)$, alors:

$$E(g(X)) = \sum_i g(x_i)\,P(X=x_i),$$

et si X est une variable aléatoire absolument continue de fonction densité $f(x)$, alors:

$$E(g(X)) = \int\limits_{\mathbb{R}} g(t)\,f(t)\,dt.$$

3)

$$E(c) = c \text{ et } E(aX+b) = aE(X)+b.$$

Paramétres de dispersion

Les moments

Les moments centrés

Définition :

Les moments centrés d'ordre k d'une variable aléatoire X sont les moments centrés par rapport à $E(X)$, appelés aussi moments centrés d'ordre k qui sont définis pour $k \in \mathbb{N}$ par:

$$\mu_k = E\left[(X-E(X))^k\right].$$

Pour une variable aléatoire discrète de fonction de masse $P(X=x)$ on a:

$$\mu_k = \sum_i (x_i - E(x))^k\,P(X=x_i),$$

et pour une variable aléatoire continue de fonction densité $f(x)$ on a:

$$\mu_k = \int\limits_{\mathbb{R}} (x-E(x))^k\,f(x)\,dx.$$

Par exemple on a $\mu_0 = 1,\ \mu_1 = 0$ et $\mu_2 = \sigma^2$.

Les moments non centrés

Définition :
Les moments non centrés d'ordre k d'une variable aléatoire X sont définis par:

$$\mu_k^{'} = E\left(X^k\right).$$

Pour une variable aléatoire discrète de fonction de masse $P(X = x)$ on a:

$$\mu_k^{'} = \sum_i x_i^k P(X = x_i),$$

et pour une variable aléatoire continue de fonction densité $f(x)$ on a:

$$\mu_k^{'} = \int_{\mathbb{R}} x^k f(x)\, dx,$$

Par exemple on a $\mu_2 = \mu_2^{'} - \left(\mu_1^{'}\right)^2$.

Moments factoriels

Définition :

Les moments factoriels d'ordre k d'une variable aléatoire X sont définis par:

$$\mu_k^{*} = E\left(X(X-1)\ldots(X-k+1)\right).$$

Par exemple on a $\mu_1^{*} = \mu_1^{'}$.

Variance et écart type

Une autre grandeur importante est la variance dont la d´efinition est:
Définition :
On appelle variance d'une variable aléatoire X notée $Var(X)$ ou σ^2, le moment centré d'ordre 2 de X (s'il existe) la quantité définie par:

$$\sigma^2 = Var(X) = E\left((X - E(X))^2\right),$$

où σ s'appelle l'écart type.
La variance ou l'écart type est une mesure de la dispersion des valeurs de la variable aléatoire autour de sa moyenne.
Remarque 4:
Si les valeurs de la variable aléatoire X tendent à se concentrer au voisinage de la moyenne $E(X)$, la variance de X est faible, par contre si les valeurs tendent à se disperser plus loin de la moyenne, la variance est grande.

Propriétés de la variance

1)

$$Var(c) = E\left((c - E(c))^2\right) = E(c - c) = 0,$$

la variance d'une constante est nulle.

2)

$$\sigma^2 = E\left((X - E(X))^2\right) = E\left(X^2\right) - E^2(X).$$

3)

$$Var(aX + b) = a^2 Var(X).$$

Caractéristique de la dispersion

Le coefficient d'asymétrie (skewness)

Dans la plupart des cas, la distribution n'est pas sym´etrique par rapport à un maximum, l'une de ses courbes s'étale à droite ou à gauche.

Définition :

La mesure de cette dissymétrie est donnée par le coefficient de la dissymétrie qui est définie par:

$$\gamma_1 = \frac{E\left((X - E(X))^3\right)}{\sigma^3} = \frac{\mu_3}{\sigma^3}.$$

Le coefficient γ_1 est une quantité positive si la courbe est étalée à droite et négative dans le cas contraire. Dans le cas d'une distribution symétrique $\gamma_1 = 0$.

Le coefficient d'aplatissement (kurtosis)

Dans des cas, la distribution est réprésentée par un pic aigu, la majorité de ses valeurs sont distribuées au voisinage de la moyenne, dans d'autres cas les distributions peuvent ˆetre plates.

Définition :

Le coefficient d'aplatissement (kurtosis) de Pearson donne une évaluation de l'importance du pic, il est défini par:

$$\gamma_2 = \frac{E\left((X - E(X))^4\right)}{\sigma^4} = \frac{\mu_4}{\sigma^4}.$$

Exemple :

Par exemple calculons le coefficient d'aplatissement de la loi normale.

Soit $X \to \mathcal{N}(0, 1)$, alors $X^2 \to \mathcal{X}_1^2$ (ces deux lois on va les voir par la suite)

On a:

$$E\left(X^{2p}\right) = 2^p \frac{\Gamma\left(\frac{1}{2}+p\right)}{\Gamma\left(\frac{1}{2}\right)},$$

où $\Gamma(p)$ est la fonction gamma (voir définition 50), et

$$\mu^4 = E\left((X - E(X))^4\right) = 4\frac{\Gamma\left(\frac{5}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} = 3,$$

donc $\gamma_2 = 3$.
Pour une variable aléatoire suivant une loi $\mathcal{N}(0,1)$, ce coefficient d'aplatissement vaut 3. C'est pour cela qu'on normalise la valeur pour mesurer l'excès d'aplatissement pour obtenir le coefficient d'aplatissement de Fisher. Le coefficient d'aplatissement de Fisher γ_3 est la valeur obtenue par le calcul suivant :

$$\gamma_3 = \frac{\mu_4}{\sigma^4} - 3.$$

Fonctions génératrice des moments

Définition :
Soit X une variable al´eatoire, alors si $E(\exp\ (tX))$ existe pour tout t appartient a` un intervalle ouvert $]t_1, t_2[$ contenant l'origine, alors la fonction

$$\begin{gathered} M_X :]t_1, t_2[\longrightarrow \mathbb{R} \\ t \longrightarrow M_X(t) = E(\exp(tX)), \end{gathered}$$

est appelée fonction génératrice des moments de la variable aléatoire X.
Si X est une variable aléatoire discrète de fonction de masse $P(x_i)$, alors:

$$M_X(t) = \sum_{x_i} P(x_i) \exp\ (tx_i),$$

et

$$M_X(t) = \int f(x) \exp\ (tx)\, dx,$$

si X une variable aléatoire absolument continue de fonction densité $f(x)$.
Exemple :
On cite quelques exemples
1) Soit X une variable aléatoire discrète de fonction de masse:

$$P(x) = \begin{cases} C_n^x p^x (1-p)^{n-x} & \text{si } x = 0, 1, \ldots, n \\ 0 & \text{sinon,} \end{cases}$$

où $C_n^x = \dfrac{n!}{x!\,(n-x)!}$, alors la fonction génératrice des moments est

$$\begin{aligned} M_X(t) &= \sum_{x=0}^{n} \exp\ (tx)\, P(x) &&= \sum_{x=0}^{n} \exp\ (tx)\, C_n^{\ x}\ \ p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} C_n^{\ x} (p \exp(t))^x (1-p)^{n-x} = && (1-p+p\exp(t))^n . \end{aligned}$$

2) Soit X une variable aléatoire absolument continue de fonction densité $f(x)$ définie par:

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

alors la fonction génératrice des moments est

$$\begin{aligned} M_X(t) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp\ (tx) \exp \quad \left(\right) (- \left(\frac{x-\mu}{\sqrt{2}\sigma} \right)^2 \Bigg) dt \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp\left(\ (-1/2\sigma^2\ \right) \left(\left(x - \left(\mu + t\sigma^2\right)\right)^2 - \left(\mu + t\sigma^2\right)^2 + \mu^2 \right) dt \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp(-1/2) \quad \left(x - \left(\mu + t\sigma^2\right)/\sigma\right)^2 \exp \quad (-1/2\sigma^2\ ) \left(\mu^2 - \left(\mu + t\sigma^2\right)^2 \right) dt \\ &= \quad \exp \quad (-1/2\sigma^2\ ) \left(\mu^2 - \left(\mu + t\sigma^2\right)^2 \right) = \exp\left(\ (\mu t + \quad \frac{t^2\sigma^2}{2} \right), \end{aligned}$$

car $\dfrac{1}{\sqrt{2\pi}\sigma} \displaystyle\int_{\mathbb{R}} \exp\ (-1/2) \quad \left(x - \left(\mu + t\sigma^2\right)/\sigma\right)^2 dt = 1.$

Quelques théorémes relatifs aux fonctions génératrices

Théoréme 1: *Soit X une variable aléatoire de fonction génératrice des moments $M_X(t)$ o̓u $t \in]t_1, t_2[$ avec $t_1 < 0 < t_2$, alors:*

1) Tous les moments de X existent.

2) $\forall t \in]-s, s[$ *tel que* $0 < s < \min(-t_1, t_2)$ *la fonction* $M_X(t)$ *admet un developpement en série entière*

$$M_X(t) = 1 + tE(X) + \frac{t^2}{2!} E\left(X^2\right) + \ldots + \frac{t^n}{n!} E(X^n) + \ldots$$

3) *Pour tout entier positif k on a*:

$$E\left(X^k\right) = M_X^{(k)}(0) .$$

Théoréme 2:

Si deux variables aléatoires X et Y admettent deux fonctions génératrices des moments $M_X(t)$ et $M_Y(t)$ respectivement avec $M_X(t) = M_Y(t)$ $\forall t \in]t_1, t_2[$ tel que $t_1 < 0 < t_2$, alors: Les variables aléatoires X et Y sont de même loi de probabilité.

Exemple :
Soit X une variable aléatoire de fonction génératrice des moments $M_X(t)$ définie par:

$$M_X(t) = (1 - p + p\exp(t))^n,$$

alors

$$M_X^{'}(t) = np(1 - p + p\exp(t))^{n-1}\exp(t),$$

d'où

$$E(X) = M_X^{'}(0) = np.$$

et

$$M_X^{''}(t) = n(n-1)p^2(1 - p + p\exp(t))^{n-2}\exp(t) + np(1 - p + p\exp(t))^{n-1}\exp(t),$$

d'où

$$E\left(X^2\right) = M_X^{''}(0) = n(n-1)p^2 + np,$$

et enfin

$$Var(X) = E\left(X^2\right) - E^2(X) = np(1-p).$$

Fonction génératrice des moments factoriels

Définition :
On appelle fonction génératrice des moments factoriels d'une variable aléatoire discrète X à valeurs entières positives et de fonction de masse $P(x)$ la quantité notée $m_X(t)$ définie pour $t > 0$ par:

$$m_X(t) = E\left(t^X\right) = \sum_x t^x P(X = x).$$

La série $\sum\limits_x t^x P(X = x)$ converge dans le domaine $t \leq 1$.

Fonction caractéristique

Définition :
Soit X une variable aléatoire réelle, on appelle fonction caractéristique de X la fonction de la variable réelle t notée $\varphi_X(t)$ définie par:

$$\varphi_X(t) = E(\exp()(itX)) = \int_{\mathbb{R}} f(x)\exp(itx)\,dx,$$

$\left(\text{avec } i^2 = -1\right)$ *si X une variable aléatoire absolument continue de fonction densité* $f(x)$, *et*

$$\varphi_X(t) = E(\exp\ (itX)) = \quad \sum_{x_j} P(x_j)\exp\quad (itx_j\ ),$$

si X est une variable aléatoire discrète de fonction de masse $P(x_j)$.

Exemple :

Soit X une variable aléatoire absolument continue de fonction densité $f(x)$ définie par:

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

alors

$$\varphi_X(t) = \exp\left(i\mu t - \frac{\sigma^2 t^2}{2}\right).$$

Remarque : La fonction caract´eristque existe pour tout t et elle est d´efinie pour toute variable aléatoire X et on a si X est une variable absolument continue de fonction densité $f(x)$

$$|\varphi_X(t)| = \left|\int_{\mathbb{R}} \exp()\,(itx)\, f(x)\,dx\right| \leq \int_{\mathbb{R}} f(x)\,dx = 1.$$

Une raison d'introduire la fonction caractéristique est qu'elle existe toujours, qui n'est pas le cas pour la fonction génératrice des moments.

On ´enonce maintenant quelques th´eor`emes importants consernants la fonction caract´eristique

Theoréme 1:

Soient X *une variable aléatoire et* $n \in \mathbb{N}$ *tel que* $\forall k \leq n,\ E\left(X^k\right) < \infty$, *alors la fonction caractéristique* $\varphi_X(t)$ *de* X *est dérivable au moins jusqu'à l'ordre* n *et on a:*

$$\varphi_X^k(0) = i^k E\left(X^k\right).$$

Theoréme 2:

Soient X *et* Y *deux variables aléatoires ayant pour fonctions caractéristiques* $\varphi_X(t)$ *et* $\varphi_Y(t)$ *respectivement alors les variables aléatoires* X *et* Y *suivent la même loi de probabilité si et seulement si*

$$\varphi_X(t) = \varphi_Y(t).$$

Loi d'une fonction d'une variable al´eatoire $Y = \Psi(X)$

Soient X une variable al´eatoire et Ψ une fonction d´efinie par $\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{\Psi} \mathbb{R}$.

On posant $Y = \Psi(X)$ et on suppose que Y est une variable aléatoire

Cas discrét

Théoréme :
Soit X une variable aléatoire discrète de support S_X et de fonction de masse P_X, alors la variable aléatoire $Y = \Psi(X)$ est discrète de support $S_Y = \Psi(S_X)$ et de fonction de masse

$$P_Y(y) = \begin{cases} \displaystyle\sum_{x \in \{\Psi(x)=y\}} P(X = x) & \text{si } y \in S_Y \\ 0 & \text{sinon.} \end{cases}$$

Exemple :
On jette un dé et on note X la variable qui indique le numéro de la face du dé jeté. Trouver la fonction de masse de la variable aléatoire $Y = X^2$.
Le support S_X de la variable aléatoire X est $S_X = \{1, 2, 3, 4, 5, 6\}$ et

$$P(X = x) = \begin{cases} 1/6 & \text{si } x \in S_X \\ 0 & \text{sinon.} \end{cases}$$

La variable aléatoire $Y = X^2$ a pour support $S_Y = \{1, 4, 9, 16, 25, 36\}$ et de fonction de masse

$$P_Y(y) = \begin{cases} P_Y(1) = \displaystyle\sum_{x \in \{x^2=1\}} P(X = x) = 1/6 \\ . \\ . \\ . \\ P_Y(36) = \displaystyle\sum_{x \in \{x^2=36\}} P(X = x) = 1/6. \end{cases}$$

Cas continu

Théoréme :
Soient X une variable aléatoire absolument continue de support S_X et de fonction de densité $f(x)$ et Ψ une fonction de $\mathbb{R}$ dans $\mathbb{R}$ tel que:
$\forall x \in S_X$ la fonction Ψ est dérivable avec $\Psi'(x) \neq 0$ sauf pour un nombre fini de points, et pour chaque $y \in \mathbb{R}$: $\exists$ exactement m points $x_1, x_2, ..., x_m$ avec $(m \geq 1)$ tel que $\Psi(x_i) = y$, alors:

$$f_Y(y) = \begin{cases} \displaystyle\sum_{i=1}^{m} f(x_i) \left|\Psi'(x_i)\right|^{-1} & \text{si } m \neq 0 \\ 0 & \text{sinon.} \end{cases}$$

Exemple :
Soit X une variable aléatoire absolument continue de fonction de densité $f(x)$ définie par:

$$f(x) = \begin{cases} \exp(-x) & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases}$$

Posons $Y = X^2$, trouver la fonction densité de la variable Y.
La fonction $y = \Psi(x) = x^2$ est une fonction continue monotone de support $\Psi(S_X) =]0, \infty[$, et elle admet une fonction inverse définie par $x = \Psi^{-1}(y) = \sqrt{y}$, et on a de plus $\left(\Psi^{-1}(y)\right)' = \dfrac{1}{2\sqrt{y}}$, d'où la fonction densité de Y est définie par

$$f_Y(y) = \begin{cases} \exp\ \left(-\ \sqrt{y}\right) \dfrac{1}{2\sqrt{y}} & \text{si } y > 0 \\ 0 & \text{sinon.} \end{cases}$$

Exemple :
Soit X une variable aléatoire absolument continue de fonction densité $f(x)$ définie par:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\ \frac{x^2}{2}\right),\ x \in \mathbb{R},$$

et soit la variable aléatoire $Y = X^2$. Cherchons la fonction densité de la variable Y.
La fonction $y = \Psi(x) = x^2$ est une fonction continue non monotone sur $\mathbb{R}$, dérivable sur $\mathbb{R}$ de dérivée $\Psi'(x) = 2x$ qui s'annule au point $x = 0$, et de plus la fonction $y = \Psi(x) = x^2$ admet deux solutions $x_1 = \sqrt{y}$ et $x_2 = -\sqrt{y}$, ce qui entraine que

$$f_Y(y) = \begin{cases} f_X(x_1)\left|\dfrac{1}{2\sqrt{y}}\right| + f_X(x_2)\left|\dfrac{1}{2\sqrt{y}}\right| & \text{si } y > 0 \\ 0 & \text{sinon} \end{cases}$$

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} \exp\left(-\ \dfrac{y}{2}\right)\left|\dfrac{1}{2\sqrt{y}}\right| + \dfrac{1}{\sqrt{2\pi}} \exp\left(-\ \dfrac{y}{2}\right)\left|\dfrac{1}{2\sqrt{y}}\right| & \text{si } y > 0 \\ 0 & \text{sinon,} \end{cases}$$

et enfin

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi y}} exp\left(-\dfrac{y}{2}\right) & \text{si } y > 0 \\ 0 & \text{sinon.} \end{cases}$$

Couple aléatoire

Définition :
Soit $(\Omega, \mathcal{T}, P)$ *un espace de probabilité, alors un couple aléatoire ou vecteur aléatoire de dimension deux est une application:* $Z = (X, Y) : (\Omega, \mathcal{T}) \to \mathbb{R}^2$, *tel que* $\forall z = (x, y) \in \mathbb{R}^2$, *l'ensemble* $\{\omega \in \Omega;\ Z(\omega) \leq z\} \in \mathcal{T}$, *où* $Z(\omega) = (X(\omega), Y(\omega))$, *et* $X(\omega), Y(\omega)$ *sont des variables aléatoires sur* $(\Omega, \mathcal{T})$.

Fonction de répartition d'un couple

Définition :
Soit (X, Y) *un couple de variables aléatoires. La fonction de répartition* F *de* (X, Y) *est une application de* $\mathbb{R}^2 \to [0, 1]$ *définie par:*

$$F(x, y) = P(X \leq x, Y \leq y).$$

Propriétés de la fonction de répartition:

1) $0 \leq F(z) \leq 1 \ \forall z \in \mathbb{R}^2$.
2) $\lim F(z) = 0$ si $x \to -\infty$ ou $y \to -\infty$ et $\lim F(x) = 1$ si $x \to \infty$ et $y \to \infty$.

Loi de probabilité d'un couple aléatoire discréte

Fonction de masse jointe
Définition :
Soient X *et* Y *deux variables aléatoires discrètes, alors la fonction de masse jointe de* X *et* Y *est définie par:*

$$P(X = x, Y = y).$$

Exemple :
Une urne contient 4 boules blanches, 2 boules rouges et 4 boules noires. On extraît de cette urne 3 boules sans remise.
On note X le nombre de boules blanches et Y le nombre de boules rouges qui sont parmi les 3 boules tirées. Déterminer la fonction de masse jointe des variables aléatoires X, et Y.
L'ensemble des valeurs possibles pour $Z = (X, Y)$ est

$$\{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2), (2,0), (2,1), (3,0)\}.$$

Les résultats sont donnés comme suit

$$P(x, y) = \frac{C_4^{x_1} C_2^{x_2} C_4^{3-x_1-x_2}}{C_{10}^3}.$$

Par exemple $P(x,y) = 1/30$ si $(x,y) \in \{(0,0),(0,2),(1,2),(3,0)\}$, $P(x,y) = 3/30$ si $(x,y) \in \{(0,1),(2,1)\}$, $P(x,y) = 6/30$ si $(x,y) \in \{(1,0),(2,0)\}$ et enfin $P(x,y) = 8/30$ si $(x,y) \in \{(1,1)\}$.

Fonction de répartition jointe

Définition :
Soit (X,Y) un couple de variables aléatoires.dicrètes, alors la fonction de répartition jointe de (X,Y) est définie par:

$$F(x,y) = \sum_{x_i \leq x} \sum_{y_k \leq y} P(X = x_i, Y = y_k)$$

Fonction de masse marginale

Définition :
Soit (X,Y) un couple de variables aléatoires discrètes, et supposons que X prend m valeurs $x_1, x_2, ..., x_m$ et Y prend n valeurs $y_1, y_2, ..., y_n$, alors la probabilité de l'évènement $X = x_j, Y = y_k$ est

$$P\left(X = x_j, Y = y_k\right) = P\left(x_j, y_k\right),$$

et les fonctions de masse marginales de X et Y sont définies respectivement par:

$$P\left(X = x_j\right) = \sum_{k=1}^{n} P\left(x_j, y_k\right) \text{ et } P(Y = y_k) = \sum_{j=1}^{m} P\left(x_j, y_k\right).$$

Les variables X et Y sont appel´ees variables marginales du couple (X, Y) et leurs fonctions de masse, appelées fonction de masse marginale de X (resp. de Y).

Remarque :
On remarque que

$$\sum_{j=1}^{m} P\left(X = x_j\right) = 1, \sum_{k=1}^{n} P(Y = y_k) = 1$$

et

$$\sum_{j=1}^{m} \sum_{k=1}^{n} P\left(x_j, y_k\right) = 1.$$

Fonction de répartition marginale

Définition :
Les fonctions de répartitions marginales des variables aléatoires X et Y sont définies respectivement par:

$$F_X(x) = P(X \leq x) = P(X \leq x, Y < \infty)$$

et

$$F_Y(y) = P(Y \leq y) = P(X < \infty, Y \leq y)$$

Loi de probabilité d'un couple aléatoire continue

Fonction de répartition jointe

Définition :
Soit (X, Y) un couple aléatoire absolument continue, alors la fonction de répartition jointe des variables aléatoires X et Y est définie par:

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dxdy,$$

où $f(x, y)$ est la fonction densité jointe du couple (X, Y) qu'on va définir ci-dessous.

Fonction de densite´ jointe

Définition :
Soit (X, Y) un couple aléatoire absolument continue, de fonction de répartition jointe $F(x, y)$ alors la fonction densité jointe des variables aléatoires X et Y est définie par:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y},$$

et cette fonction vérifie les conditions suivantes

$$\begin{array}{ll} 1) & f(x, y) \geq 0\ \forall (x, y) \in \mathbb{R}^2 \\ 2) & \displaystyle\iint_{\mathbb{R}^2} f(x, y)\, dxdy = 1. \end{array}$$

Fonction densité marginale

Définition :
Soit (X, Y) un couple aléatoire absolument continue, de fonction densité jointe $f(x, y)$, alors les fonctions densités marginales de X et Y sont définies respectivement par:

$$f_X(x) = \int_{\mathbb{R}} f(x, y)\, dy,$$

et

$$f_Y(y) = \int_{\mathbb{R}} f(x, y)\, dx.$$

Fonction de répartition marginale

Définition :
Soit (X, Y) un couple aléatoire absolument continue, de fonction densité jointe $f(x, y)$, alors les fonctions de répartitions marginales de X et Y sont définies respectivement par:

$$F_X(x) = \int_{-\infty}^{x} \int_{\mathbb{R}} f(x, y)\, dxdy$$

et

$$F_Y(y) = \int_{\mathbb{R}} \int_{-\infty}^{y} f(x, y)\, dxdy.$$

Variables aléatoires indépendantes

Définition :
Soient X et Y deux variables aléatoires, si les évènements $X = x$, et $Y = y$ sont indépendantes pour tout x et y alors les variables aléatoires X et Y sont dites indépendantes et on a:

$$P(X = x, Y = y) = P(X = x) . P(Y = y),$$

si les variables aléatoires X et Y sont discrétes et

$$f(x, y) = f_X(x) . f_Y(y),$$

si les variables aléatoires X et Y sont absoluments continues.
De plus X et Y sont deux variables aléatoires indépendantes si

$$P(X \leq x, Y \leq y) = P(X \leq x) . P(Y \leq y),$$

ou

$$F(x, y) = F_X(x) . F_Y(y).$$

Exemple :
Soit (X, Y) un couple aléatoire absolument continue, de fonction densité jointe $f(x, y)$ définie par

$$f(x, y) = \begin{cases} \exp(-)(x + y) & \text{si } (x, y) \in \mathbb{R}_+^{*2} \\ 0 & \text{sinon.} \end{cases}$$

On a

$$f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy = \int_{\mathbb{R}} \exp(-(x+y))\,dy = \exp(-x) \text{ pour } x > 0,$$

et

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx = \int_{\mathbb{R}} \exp(-(x+y))\,dx = \exp\ (-y) \text{ pour } y > 0,$$

donc

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = f_X(x)\,.\,f_Y(y)\,,$$

et les variables al´eatoires X et Y sont ind´ependantes.

Théoréme 01:

Si X, et Y sont deux variables aléatoires indépendantes de fonctions génératrices de moments $M_X(t)$ et $M_Y(t)$ respectivement alors

$$M_{X+Y}(t) = M_X(t)\,.M_Y(t)\,.$$

En général la fonction génératrice des moments de la somme de variables aléatoires indépendantes est égale au produit des fonctions génératrices des moments de chacune des variables aléatoires ind´ependantes de la somme.

Théorème 02:

Soient X et Y deux variables aléatoires indépendantes de fonctions caractéristiques $\varphi_X(t)$ et $\varphi_Y(t)$ respectivement alors:

$$\varphi_{X+Y}(t) = \varphi_X(t)\,.\varphi_Y(t)\,.$$

En général la fonction caractéristique de la somme de variables aléatoires indépendantes est égale au produit de leurs fonctions caractéristiques.

Changement de variables

Théorème 01:

Soient X et Y deux variables aléatoires absolument continues de fonction densité jointe $f(x,y)$ et soient les variables aléatoires $U = \varphi_1(X,Y)$ et $V = \varphi_2(X,Y)$ où $X = \psi_1(U,V)$ et $Y = \psi_2(U,V)$ alors la fonction densité jointe $g(u,v)$ de U et V est définie par:

$$g(u,v) = f(\psi_1(u,v)\,,\psi_2(u,v))\,.\,|J(u,v)|$$

avec

$$J(u,v) = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/du & \partial y/\partial v \end{vmatrix}.$$

Exemple 0 1:

Soit (X,Y) un couple aléatoire absolument continue, de fonction densité jointe $f(x,y)$ définie par:

$$f(x,y)=\begin{cases}\exp(-)(x+y) & \text{si } (x,y)\in\mathbb{R}_{+}^{*2}\\ 0 & \text{sinon.}\end{cases}$$

Posons $U=X+Y$ et $V=X-Y$.
Quelle est la fonction densité jointe des variables aléatoires U et V.
On a $X=\dfrac{U+V}{2}$ et $Y=\dfrac{U-V}{2}$, donc la fonction densité jointe $g(u,v)$ des variables aléatoires U et V est

$$g(u,v)=f\left(\frac{u+v}{2},\frac{u-v}{2}\right)|J(u,v)|=\ \exp()(-u)\quad \text{si } |v|<u \text{ avec } u>0$$

Théorème 02:

Soient X et Y deux variables aléatoires absoluments continues de fonction densité jointe $f(x,y)$ et soient les variables aléatoires $U=\varphi(X,Y)$ et $V=X$ ou Y, où $X=\psi_1(U,V)$ et $Y=\psi_2(U,V)$, alors la fonction densité jointe $g(u,v)$ de U et V est définie par:

$$g(u,v)=f(\psi_1(u,v),\psi_2(u,v))\,.\,|J(u,v)|,$$

et la fonction densité de la variable U est la fonction densité marginale de U définie par:

$$g_U(u)=\int_{\mathbb{R}} g(u,v)\,dv.$$

Exemple 02:

Soient X, et Y deux variables aléatoires absoluments continues de fonction densité jointe $f(x,y)$ et soit la variable aléatoire $U=\varphi(X,Y)=X.Y$, quelle est la fonction densité de la variable U.
Soit le couple $(U,V)=(XY,Y)$, donc la fonction densité jointe $g(u,v)$ des variables aléatoires U et V est

$$g(u,v)=f(x,y)\,.\,|J(u,v)|=f(u/v,v)\,.\left|\frac{1}{v}\right|,$$

d'où la fonction densité de la variable U est

$$g_U(u)=\int_{\mathbb{R}} g(u,v)\,dv=\int_{\mathbb{R}} f(u/v,v)\,.\left|\frac{1}{v}\right|dv.$$

Si les deux variables aléatoires X et Y sont indépendantes alors:

$$g_U(u)=\int_{\mathbb{R}} f_X(u/v)\,.f_Y(v)\,.\left|\frac{1}{v}\right|dv.$$

Exemple 03:

Soient X et Y deux variables aléatoires absoluments continues de fonction densité jointe $f(x,y)$ et soit la variable aléatoire $U=\varphi(X,Y)=X+Y$, alors la fonction densité de la variable U est

$$g_U(u) = \int_{\mathbb{R}} f(u - v, v)\, dv.$$

Soient les variables aléatoires U et V tel que $U = X + Y$ et $V = Y$, ce qui correspond à $x = u - v$ et $y = v$, d'où $J(u, v) = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$, et la fonction densité jointe $g(u, v)$ des variables aléatoires U et V est

$$g(u, v) = f(x, y) = f(u - v, v),$$

et la fonction densité de la variable U est

$$g_U(u) = \int_{\mathbb{R}} f(u - v, v)\, dv.$$

Si X, et Y sont deux variables aléatoires indépendantes alors la fonction de densité de la variable U est

$$g_U(u) = \int_{\mathbb{R}} f_X(u - v) . f_Y(v)\, dv = f_X * f_Y \text{ (formule de convolution).}$$

D'aprés la formule de convolution on a

$$f_X * f_Y = f_Y * f_X,$$

d'où

$$g_U(u) = \int_{\mathbb{R}} f_X(u - v) . f_Y(v)\, dv = \int_{\mathbb{R}} f_Y(u - v) . f_X(v)\, dv.$$

Exemple 04:
Soient X, et Y deux variables aléatoires indépendantes et absolument continues de fonction densité jointe $f(x, y) = f_X(x) . f_Y(y)$ avec

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\Big()\,(-x^2/2\Big),\ x \in \mathbb{R}$$

et

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\Big()\,(-y^2/2\Big),\ y \in \mathbb{R}.$$

Trouver la fonction densité $g_U(u)$ de la variable $U = X + Y$.
La fonction densité $g_U(u)$ est

$$\begin{aligned}
g_U(u) &= \int_{\mathbb{R}} \frac{1}{2\pi} \exp\left(\right)\left(-(u-v)^2/2\right) . \exp\left(\right)\left(-v^2/2\right) dv \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \exp\left(\right)\left(-\frac{u^2}{2} + uv - \frac{v^2}{2} - \frac{v^2}{2}\right) dv \\
&= \frac{1}{2\pi} \exp\left(\right)\left(-\frac{u^2}{2}\right) \int_{\mathbb{R}} \exp\left(\right)\left(-\frac{1}{2}\left(2v^2 - 2uv\right)\right) dv \\
&= \frac{1}{2\pi} \exp\left(\right)\left(-\frac{u^2}{4}\right) \int_{\mathbb{R}} \exp\left(\right)\left(-\frac{1}{2}\left(\sqrt{2}v - \frac{u}{\sqrt{2}}\right)^2\right) dv.
\end{aligned}$$

En posant $t = \sqrt{2}v - \dfrac{u}{\sqrt{2}}$ on trouve que $\dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \exp\left(\right)\left(-\frac{1}{2}t^2\right) \frac{dt}{\sqrt{2}} = 1$, et

$$\begin{aligned}
g_U(u) &= \frac{1}{\sqrt{2\pi}} \exp\left(\right)\left(-\frac{u^2}{4}\right) \int_{\mathbb{R}} \exp\left(\right)\left(-\frac{1}{2}t^2\right) \frac{dt}{\sqrt{2}} \\
&= \frac{1}{\sqrt{2\pi}} \exp(-) \frac{1}{2}\left(\frac{u^2}{2}\right), \; u \in \mathbb{R}.
\end{aligned}$$

Variance des distributions jointes et covariance

Définition :

Soient X et Y deux variables aléatoires discrètes, de fonction de masse jointe $P(X = x, Y = y)$, alors les espérances et les variances des variables aléatoires X et Y sont définies respectivement par:

$$E(X) = \mu_X = \sum_x \sum_y x P(x, y), \; E(Y) = \mu_Y = \sum_x \sum_y y P(x, y),$$

$$Var(X) = \sigma_X^2 = \sum_x \sum_y (x - \mu_X)^2 P(x, y), \; Var(Y) = \sigma_Y^2 = \sum_x \sum_y (y - \mu_Y)^2 P(x, y),$$

et de plus on a:

$$Cov(X, Y) = \sigma_{XY} = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) P(x, y),$$

et si les variables aléatoires X et Y sont absolument continues de fonction densité jointe $f(x, y)$ alors les espérances et les variances des variables aléatoires X et Y sont définies respectivement par:

$$\mu_X = \int\int x f(x, y)\, dxdy, \; \mu_Y = \int\int y f(x, y)\, dxdy,$$

$$\sigma_X^2 = \int\int (x - \mu_X)^2 f(x, y)\, dxdy, \; \sigma_Y^2 = \int\int (y - \mu_Y)^2 f(x, y)\, dxdy,$$

et de plus on a:

$$Cov\left(X,Y\right)=\sigma_{XY}=\int\int\left(x-\mu_X\right)\left(y-\mu_Y\right)f\left(x,y\right)dxdy=E\left(\left(X-\mu_X\right).\left(Y-\mu_Y\right)\right).$$

Les théorèmes suivants ont une importance relative à la covariance.

Théorème 01**:**

Soient X et Y deux variables aléatoires alors:

$$E\left(X+Y\right)=E\left(X\right)+E\left(Y\right).$$

De plus si les variables aléatoires X et Y sont indépendantes alors:

$$E\left(XY\right)=E\left(X\right).E\left(Y\right).$$

Preuve

Supposons que les variables aléatoires X et Y sont discrètes, de fonction de masse jointe $P\left(x,y\right)$ alors on a:

$$\begin{aligned} E\left(X+Y\right) &= \sum_x\sum_y\left(x+y\right)P\left(x,y\right) \\ &= \sum_x xP\left(x,y\right)+\sum_y yP\left(x,y\right) \\ &= E\left(X\right)+E\left(Y\right). \end{aligned}$$

et

$$\begin{aligned} E\left(XY\right) &= \sum_x\sum_y xyP\left(x,y\right) &&= \sum_x\sum_y xyP_X\left(x\right)P_Y\left(y\right) \\ &= \sum_x x\left(P_X\left(x\right)\sum_y yP_Y\left(y\right)\right) &&= \sum_x xP_X\left(x\right)E\left(Y\right) \quad = E\left(X\right)E\left(Y\right). \end{aligned}$$

Et si les variables aléatoires X et Y sont absolument continues de fonction densité jointe $f\left(x,y\right)$ alors:

$$\begin{aligned} E\left(XY\right) &= \int\int xyf\left(x,y\right)dxdy &&= \int\int xyf_X\left(x\right)f_Y\left(y\right)dxdy \\ &= \int xf_X\left(x\right)dx\int yf_Y\left(y\right)dy &&= E\left(X\right)E\left(Y\right). \end{aligned}$$

Théorème 02:

Soient X et Y deux variables aléatoires alors:

$$\begin{aligned} &1)\quad Cov\left(X,Y\right) &&= E\left(XY\right)-E\left(X\right).E\left(Y\right). \\ &2)\quad Var\left(X\pm Y\right) &&= Var\left(X\right)+V\left(Y\right)\pm Cov\left(X,Y\right). \\ &3)\quad Cov^2\left(X,Y\right) &&\leq Var\left(X\right)Var\left(Y\right). \end{aligned}$$

Preuve de 3):
On a

$$\forall a \in \mathbb{R} : E\left((X - aY)^2\right) \geq 0,$$

d'où

$$a^2 E\left(Y^2\right) - 2aE\,(XY) + E\left(X^2\right) \geq 0 \text{ simplement si } E^2\,(XY) - E^2\,(X)\,E^2\,(Y) \leq 0.$$

Remplaçons X et Y par $(x - \mu_X)$ et $(y - \mu_Y)$ respectivement pour conclure.
Si les variables aléatoires X et Y sont indépendantes alors: $Cov\,(X, Y) = 0$, et le contraire n'est pas vrai. Par ailleurs, si les variables sont dépendantes par exemple $X = Y$, alors $Cov\,(X, Y) = \sigma_{XY} = \sigma_X \sigma_Y$. Cela nous ramène à évaluer la dépendance des variables aléatoires X et Y, cette évaluation s'exprime par $\rho = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$, qui est une grandeur sans dimension appelée coefficient de corrélation linéaire entre X et Y.
D'aprés le théorème précédent on remarque que $-1 \leq \rho \leq 1$.
Dans le cas où $\rho = 0$ on dit que les variables X et Y sont non correlés, elles peuvent être indépendantes ou non.

Distribution conditionnelle

Soient A et B deux évènements avec $P\,(A) > 0$. On a vu précédement que

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

Définition
Si X et Y sont deux variables aléatoires discrètes, et si nous avons les évènements $(A : X = x)\,,\ (B : Y = y)\,,$ *alors* $P\,(B/A)$ *devient*

$$P\,(B/A) = P\,(Y = y/X = x) = \frac{P\,(X = x,\ Y = y)}{P_X\,(x)},$$

cette fonction est appelée fonction de masse conditionnelle de Y *sachant* X.
De même la fonction densité conditionnelle de Y *sachant* X *où* X *et* Y *sont des variables aléatoires absolument continues est donnée par:*

$$f\,(y/x) = \frac{f\,(x, y)}{f_X\,(x)},$$

où $f\,(x, y)$ *est la densité jointe de* X *et* Y, *et* $f_X\,(x)$ *est la densité marginale de* X.

Espérance conditionnelle

Définition :
Soient X, et Y deux variables aléatoires discrètes de fonction de masse jointe $P(x,y)$, alors on peut définir l'espérance conditionnelle de $Y = y$ sachant $X = x$ par:

$$E(Y/X=x) = \sum_y yP(Y=y/X=x).$$

Si X, et Y sont deux variables aléatoires continues de fonction densité jointe $f(x,y)$, l'espérance conditionnelle de Y sachant X s'exprime par:

$$E(Y/X=x) = \int_{\mathbb{R}} yf(y/x)\,dy,$$

où $f(y/x)$ est la fonction densité conditionnelle de Y sachant X.
L'espérance conditionnelle de $Y = y$ sachant $X = x$ est l'espérance de Y prise par rapport à sa loi conditionnelle.

Propriétés
Si X et Y sont deux variables aléatoires indépendantes alors:

$$E(Y/X=x) = E(Y).$$

et

$$E(Y) = \int E(Y/X=x)\, f_X(x)\,dx.$$

On remarque que $E(Y/X=x)$ est une fonction de x qui est égale à une fonction $\theta(x)$ que l'on appelle fonction de régréssion de Y en X.
La quantité $E(Y/X=x)$ dépend des valeurs prises par X ce qui nous conduit à définir la variable aléatoire espérance conditionnelle qui prend pour valeur $E(Y/X=x)$ avec les probabilités $P(X=x)$, alors

$$E(Y/X=x) = \theta(X).$$

Théorème de l'espérance totale
Théorème :
Soient X, et Y deux variables aléatoires de fonction de masse jointe $P(x,y)$, alors:

$$E(\theta(X)) = E(Y/X) = E(Y).$$

Preuve:
Si X, et Y deux variables aléatoires de fonction de masse jointe $P(x,y)$, alors:

$$\begin{aligned}
E(E(Y/X=x)) &= \sum_x E(Y/X=x)P(X=x) \\
&= \sum_x \left(\sum_y yP(Y=y/X=x)\right)P(X=x) \\
&= \sum_y y\left(\sum_x \frac{P(Y=y,X=x)}{P(X=x)}P(X=x)\right) \\
&= \sum_y y(P(Y=y)) = E(Y),
\end{aligned}$$

et dans le cas où les variables X et Y sont absolument continues de fonction densité jointe $f(x,y)$, alors:

$$\begin{aligned} E(E(Y/X=x)) &= \int_{\mathbb{R}} E(Y/X=x)\, f_X(x)\, dx \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} y f(y/x)\, dy \right) f_X(x)\, dx \\ &= \int_{\mathbb{R}} y \left(\int_{\mathbb{R}} \frac{f(x,y)}{f_X(x)} f_X(x)\, dx \right) dy \\ &= \int_{\mathbb{R}} y f_Y(y)\, dy = E(y). \end{aligned}$$

Variance conditionnelle

De la même façon on peut définir la variance conditionnelle de Y sachant $X = x$ de la manière suivante:

Définition :
On appelle variance conditionnelle de Y sachant $X = x$ notée $Var(Y/X = x)$ la quantité

$$Var(Y/X=x) = E\left((Y - E(Y/X=x))^2 / X = x\right).$$

Si X, et Y deux variables aléatoires discrètes de fonction de masse jointe $P(x,y)$, la variance conditionnelle de Y sachant $X = x$ s'écrit

$$Var(Y/X=x) = \sum_x \left(\left(Y - \left(\sum_y yP(Y=y/X=x) \right) \right)^2 / X = x \right) P(X=x),$$

et si X, et Y sont deux variables aléatoires continues de fonction densité jointe $f(x,y)$, alors

$$Var(Y/X=x) = \int_{\mathbb{R}} (y - E(Y/X))^2 f(y/x)\, dy.$$

La variance conditionnelle de Y sachant $X = x$ est l'espérance conditionnelle du carré de l'écart à l'espérance conditionnelle.
Comme pour l'espérance on définit une variable aléatoire appelée variance conditionnelle par $Var(Y/X) = x) = \zeta(X)$.

Théorème de la variance totale

Théorème :
Soient X, et Y deux variables aléatoires alors on a:

$$Var(Y) = E(Var(Y/X)) + Var(E(Y/X)).$$

Lois de probabilités usuelles

On va étudier dans ce chapitre des lois usuelles fréquement utilisées en pratique. On peut partager ces lois en deux catégories, les lois discrètes et les lois continues.

Lois disrétes

Loi discréte uniforme

Définition :
Soit X une variable aléatoire, on dit que X suit la loi discrète uniforme si X prend les valeurs naturelles de 1 à n avec $P(X = x) = \frac{1}{n}\ \forall x \in \{1, 2, ..., n\}$.
Caractéristiques

$$E(X) = \sum_{x=1}^{n} xP(X = x) = \frac{1}{n}\sum_{x=1}^{n} x = \frac{n(n+1)}{2n} = \frac{(n+1)}{2}.$$

$$E\left(X^2\right) = \frac{1}{n}\sum_{x=1}^{n} x^2 = \frac{n(n+1)(2n+1)}{6n} = \frac{(n+1)(2n+1)}{6},$$

d'où

$$Var(X) = E\left(X^2\right) - E^2(X) = \frac{n^2 - 1}{12},$$

et la fonction caractéristique est

$$\varphi_X(t) = \frac{\exp\left(12it(n+1) \right)}{n\frac{\sin\left(\frac{nt}{2}\right)}{\sin\left(\frac{t}{2}\right)}}$$

Loi de Bernoulli

Supposons qu'on va réaliser une expérience à deux résultats possibles, un résultat qui s'appelle succés avec une probabilité p et l'autre s'appelle échec avec une probabilité
$1 - p$, cette expérience s'appelle expérience de Bernoulli du nom du Jaques Bernoulli qui les étudia vers la fin du dix-septi`eme si`ecle.
Définition :
Soit X une variable aléatoire qui ne prend que les deux valeurs 1 ou 0 avec les probabilités p ou $1 - p$ avec $p \in \,]0, 1[$, alors on dit que X suit la loi de Bernoulli de paramètre p qu'on notera $\mathcal{B}(p)$, et la fonction de masse de la variable X est donnée par:

$$P(X=x)=\begin{cases} p & \text{si } x=1 \\ 1-p & \text{si } x=0 \\ 0 & \text{sinon.} \end{cases}$$

Caractéristiques

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

et la fonction caractéristique est

$$\varphi_X(t)=(1-p)+p\exp(it).$$

Loi Binomiale

Définition :
Considérons une expérience de Bernoulli de paramètre p qu'on répète n fois de façon indépendante. Soit X la variable aléatoire qui compte le nombre de succès parmi ces n expériences, alors on dira X suit la loi binomiale de paramètres n et p, et on note $X \to \mathcal{B}(n,p)$, et la fonction de masse de X est définie par:

$$P(X=x)= \begin{matrix} C_n^x p^x (1-p)^{n-x} & \text{si } x=0,1,\ldots,n \\ 0 & \text{sinon,} \end{matrix}$$

où $C_n^x=\dfrac{n!}{x!(n-x)!}$.
Exemple :
On lance une pièce de monnaie 5 fois. Quelle est la probabilité d'obtenir exactement 3 faces.
La probabilité d'obtenir exactement 3 faces est

$$P(X=3)=C_5^3(1/2)^3(1/2)^2=\frac{5!}{2!.3!}(1/2)^5=3/8.$$

Caractéristiques de la loi binomiale
Soit X une variable aléatoire qui suit la loi $\mathcal{B}(n,p)$ alors:

$$E(X)=np,\ Var(X)=np(1-p),$$

$$\gamma_1=\frac{1-2p}{\sqrt{np(1-p)}},\ \gamma_2=3+\frac{1-6p(1-p)}{\sqrt{np(1-p)}},$$

et la fonction caractéristique

$$\varphi_X(t)=(1-p+p\exp\ (it))^n.$$

Théorème 18:
Si les variables aléatoires X_1 et X_2 sont indépendantes et suivent $\mathcal{B}(n_1,p)$ et $\mathcal{B}(n_2,p)$ respectivement, la variable aléatoire $X=X_1+X_2$ suit la loi $\mathcal{B}(n_1+n_2,p)$.

Loi multinomiale

Dans une expérience aléatoire, supposons que les évènements $A_1, A_2, ..., A_k$ peuvent se réaliser avec les probabilités $p_1, p_2, ..., p_k$ respectivement où $p_1 + p_2 + ... + p_k = 1$. Soient les variables aléatoires $X_1, X_2, ..., X_k$ qui représentent le nombre de fois que $A_1, A_2, ..., A_k$ se réalisent sur n expériences de manière que $X_1 + X_2 + ... + X_k = n$, alors la fonction jointe des variables aléatoires $X_1, X_2, ..., X_k$ est définie par:

$$P(X_1 = n_1, X_2 = n_2, ..., X_k = n_k) = \frac{n!}{n_1!n_2!...n_k} p_1^{n_1} p_2^{n_2} ... p_k^{n_k}, \text{ où } n_1 + n_2 + ... + n_k = n.$$

Cette loi est une généralisation de la loi binômiale, et on a:

$$E(X_1) = np_1, \ E(X_2) = np_2, ..., \ E(X_k) = np_k$$

Loi hypergéométrique

Soit une population de N individus pour laquelle une proportion Np possède un certain caractère, on extrait de cette population un échantillon de taille n d'une façon exaustive (sans remise). Soit la variable aléatoire X qui représente le nombre d'individus de l'échantillon possédant ce caractère. Le nombre de cas possibles de tirer un échantillon de taille n parmi une population de taille N est C_N^n, et le nombre de cas favorables de l'évènement $X = x$ est $C_{Np}^x C_{N(1-p)}^{n-x}$, où C_{Np}^x est le nombre de groupes de x individus possédant le caractère, et $C_{N(1-p)}^{n-x}$ est le nombre de groupes de $n - x$ individus ne le possède pas, d'où la fonction de masse de la variable aléatoire X est

$$P(X = x) = \frac{C_{Np}^x C_{N(1-p)}^{n-x}}{C_N^n}, \text{ avec } \min X = \max(0, n - N(1-p)) \text{ et } \max X = \min(n, Np).$$

On dit que X suit la loi hypergéométrique de paramètres N, n, et p, et on notera
$X \rightarrow H(N, n, p)$

Exemple :

Une urne contient b boules blanches et r boules rouges, on extrait de cette urne n boules. Soit la variable aléatoire X qui représente le nombre de boules blanches parmi les n tirées, alors la fonction de masse de X est $P(X = x) = \dfrac{C_b^x C_r^{n-x}}{C_{(b+r)}^n}$.

Caractéristiques de la loi hypergéométrique

Soit X une variable aléatoire qui suit la loi $H(N, n, p)$ alors:

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

On peut considérer que la variable aléatoire X est une somme de n variables aléatoires de Bernoulli non indépendantes qui correspondent aux tirages successifs de n individus.
On a

$$E\left(X_1\right) = P\left(X_1 = 1\right) = p, \text{ et } E\left(X_2\right) = P\left(X_2 = 1\right).$$

Puisque les variables X_1 et X_2 sont dépendantes, donc

$$\begin{aligned} E\left(X_2\right) = P\left(X_2 = 1\right) &= P\left(X_2 = 1/X_1 = 1\right) P\left(X_1 = 1\right) + P\left(X_2 = 1/X_1 = 0\right) P\left(X_1 = 0\right) \\ &= \frac{Np-1}{N-1}p + \frac{Np}{N-1}\left(1-p\right) \\ &= p. \end{aligned}$$

Alors

$$E\left(X\right) = E\left(\sum_{i=1}^{n} X_i\right) = np$$

Pour la variance on a:

$$\begin{aligned} Var\left(X\right) &= Var\left(\sum_{i=1}^{n} X_i\right) \\ &= \sum_{i=1}^{n} Var\left(X_i\right) + \sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} Cov\left(X_i, X_j\right) \\ &= np\left(1-p\right) + \sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} Cov\left(X_i, X_j\right). \end{aligned}$$

Calculons $Cov\left(X_i, X_j\right)$

$$Cov\left(X_i, X_j\right) = E\left(X_i X_j\right) - p^2,$$

et

$$E\left(X_i X_j\right) = P\left(X_i X_j = 1\right) = P\left(X_j = 1/X_i = 1\right) P\left(X_i = 1\right) = \frac{Np-1}{N-1}p.$$

On remarque que $P\left(X_i X_j = 1\right)$ ne dépend pas des indices i et j, et puisque il y a $n\left(n-1\right)$ manières de prendre le couple $\left(X_i, X_j\right)$, d'où

$$\begin{aligned} Var\left(X\right) &= np\left(1-p\right) + n\left(n-1\right)\frac{Np-1}{N-1}p \\ &= \frac{np\left(1-p\right)\left(N-n\right)}{N-1}. \end{aligned}$$

Remarque :

Lorsque $N \to \infty$ (ou N est trés grand devant n), la loi hypergéométrique de paramètres N, n, et p est approximée par la loi $\mathcal{B}(n, p)$.
En pratique on peut faire cette approximtion dés que $\dfrac{n}{N} < 10\%$.

Loi binomiale négative

Considérons une expérience de Bernoulli $\mathcal{B}(p)$. On refait cette expérience de manière indépendante jusqu à l'obtention du k^{ieme} succés. Soit X la variable aléatoire qui représente le nombre d'essais nécéssaires pour avoir le k^{ieme} succés, alors l'évènement

$$\begin{aligned}\{X = x\} &= \{\text{Faire } X \text{ expériences pour obtenir } k \text{ succés}\} \\ &= \{\text{Avoir } (k-1) \text{ succés en } (x-1) \text{ expériences}\} \\ &\cap \left\{\text{Avoir le } k^{ieme} \text{ succés à la } x^{ime} \text{ expérience}\right\},\end{aligned}$$

a comme fonction de masse

$$P(X = x) = C_{x-1}^{k-1} p^k (1-p)^{x-k}\ , \text{ où } x = k, k+1, \ldots$$

On dit que X suit la loi binômiale négative et on notera $X \to \mathcal{BN}(k, p)$.
Caractéristiques de la loi binomiale négative
Soit X une variable aléatoire qui suit la loi $\mathcal{BN}(k, p)$ alors:

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

et

$$\varphi_X(t) = \frac{p\exp(it)}{(1-(1-p)\exp\ (it))^k}.$$

Loi géométrique

Le cas particulier $k = 1$ pour la loi binomiale négative donne la loi géométrique.

Loi de Poisson

Soit X une variable aléatoire discrète, alors on dit que X suit la loi de Poisson de paramètre $\lambda > 0$ (en référence au nom de S.D.Poisson qui l'a établi en 1838) si sa fonction de masse est définie par:

$$P(X = x) = \frac{\lambda^x \exp()(-\lambda)}{x!},\ x \in \mathbb{N},$$

et on notera $X \to \mathcal{P}(\lambda)$.
Caractéristiques de la loi de Poisson
Soit X une variable aléatoire qui suit la loi $\mathcal{P}(\lambda)$, alors:

$$E(X) = \lambda,\ Var(X) = \lambda,$$

$$\gamma_1 = 1/\sqrt{\lambda},\ \gamma_2 = 3 + (1/\lambda)\,.$$

et la fonction caractéristique est

$$\varphi_X(t) = \exp\left(\lambda\left(\exp\left(\right)(it) - 1\right)\right)$$

La loi de Poisson régi dans les cas suivants par exemple:
- Le nombre d'appels téléphoniques reçus pendant un intervalle de temps.
- Le nombre de voitures passant par une route dans un intervalle de temps.
- Le nombre d'arrivées à une station dans un intervalle de temps.

Remarque 6:
Lorsque $n \longrightarrow \infty$, et p voisine de 0, la loi binomiale de paramètres n, et p est approximée par la loi $\mathcal{P}(np)$.
En pratique on peut faire cette approximtion dés que $n > 50$ et $np < 5$.

Lois absolument continues

Loi Uniforme

Une variable aléatoire X est dite de distribution uniforme sur $[a, b]$ si sa densité de probabilité est définie par:

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

et on note $X \longrightarrow U([a, b])$

Caractéristiques de la loi uniforme
Soit X une variable aléatoire qui suit la loi $U([a, b])$ alors

$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{si } x < a \\ \dfrac{x-a}{b-a} & \text{si } a \leq x < b \\ 1 & \text{si } x \geq b. \end{cases}$$

$$E(X) = \frac{a+b}{2},\ Var(X) = (b-a)^2/12$$

et

$$\varphi_X(t) = \begin{cases} \dfrac{\exp\left(\right)(itb) - \exp\left(\right)(ita)}{it(b-a)} & \text{si } t \neq 0 \\ 1 & \text{si } t = 0. \end{cases}$$

Loi gamma

Commençons premièrement par la définition de la fonction gamma

Définition 50:

On appelle fonction gamma, l'intégrale reccurente définie sur $\mathbb{R}_+$ par:

$$\Gamma(p) = \int_0^\infty x^{p-1} \exp()(-x)\,dx, \text{ où } p > 0.$$

Propriétés de la fonction gamma

$$\begin{array}{lrcl} 1) & \Gamma(p+1) & = & p\Gamma(p) \text{ et } \Gamma(1) = 1. \\ 2) & \Gamma(p+1) & = & p! \text{ si } p \in \mathbb{N}^*. \\ 3) & \Gamma(1/2) & = & \sqrt{\pi}. \\ 4) & \Gamma(p) & \sim & (p/\exp()(1))^p \sqrt{2\pi p}\,(1 + (1/12p)) \text{ quand } p \to \infty. \end{array}$$

Soit X une variable aléatoire, alors on dit que X suit la loi gamma de paramètres $\lambda,\ p,\ (\lambda > 0,\ p > 0)$ notée $X \to \gamma(\lambda, p)$ si sa fonction densité est définie par

$$f(x) = \begin{cases} \dfrac{\lambda^p}{\Gamma(p)} x^{p-1} \exp()(-\lambda x) & \text{si } x \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

La fonction $f(x)$ est vraiment une fonction densité

$$\begin{aligned} \int_0^\infty f(x)\,dx &= \int_0^\infty \frac{\lambda^p}{\Gamma(p)} x^{p-1} \exp()(-\lambda x)\,dx \\ &= \frac{\lambda^p}{\lambda\Gamma(p)} \int_0^\infty \left(\frac{t}{\lambda}\right)^{p-1} \exp()(-t)\,dt \\ &= \frac{1}{\Gamma(p)} \int_0^\infty t^{p-1} \exp()(-t)\,dt = 1. \end{aligned}$$

Propriétés de loi gamma

Soit X une variable aléatoire qui suit la loi $\gamma(\lambda, p)$, alors:

$$E(X) = p/\lambda,\ Var(X) = p/\lambda^2,$$

et

$$\varphi_X(t) = \left(\frac{\lambda}{\lambda - it}\right)^p.$$

Cas particulier: Si $p = 1$, la loi gamma $\gamma(\lambda, p)$ est appelée la loi exponentielle de paramètre λ notée $\mathcal{E}(\lambda)$.

Théorème :
Si X et Y deux variables aléatoires indépendantes suivants respectivement les lois $\gamma(\lambda, p)$ et $\gamma(\lambda, q)$, alors $X+Y$ suit la loi $\gamma(\lambda, p+q)$.

Loi beta

Pour la loi bêta on distingue deux types de loi bêta

Loi béta de type 1 C'est la loi d'une variable aléatoire X, avec $0 \leq X \leq 1$, dépendant de $p,\ q,\ (p>0, q>0)$ dont la fonction densité est définie par:

$$f(x) = \begin{cases} \dfrac{1}{B(p,q)} x^{p-1}(1-x)^{q-1} & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon,} \end{cases}$$

où $B(p,q)$ est la fonction bêta qui est définie par:

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx.$$

Si X suit la loi bêta de type 1 de paramètres p, et q, on note $X \to B_1(p,q)$.
Caractéristiques de loi bêta de type 1
Soit X une variable aléatoire qui suit la loi $B_1(p,q)$, alors:

$$E(X) = \frac{p}{p+q},\ Var(X) = \frac{pq}{(p+q+1)(p+q)^2}.$$

Loi bêta de type 2 Soit X une variable aléatoire suivant la loi bêta de type 1 de paramètres p,et q, alors la variable aléatoire $Y = \frac{X}{1-X}$ suit la loi bêta de type 2 dont la fonction densité est définie par:

$$f(y) = \begin{cases} f_X\left(\dfrac{y}{y+1}\right).\dfrac{1}{(y+1)^2} = \dfrac{1}{B(p,q)}\dfrac{y^{p-1}}{(1+y)^{p+q}} & \text{si } 0 \leq y < \infty \\ 0 & \text{sinon} \end{cases}$$

Si Y suit la loi bêta de type 2 de paramètres p, et q, on note que $Y \to B_2(p,q)$.
Caractéristiques de loi bêta de type 2
Soit Y une variable aléatoire qui suit la loi $B_2(p,q)$, alors

$$E(Y) = \frac{p}{q-1} \text{ pour } q > 1,\ Var(Y) = \frac{p(p+q-1)}{(q-1)^2(q-2)} \text{ pour } q > 2.$$

Loi normale ou loi de Laplace-gauss

Une des distributions continues les plus importantes dans la théorie de probabilité est la distribution normale ou la distribution de Gauss.
Soit X une variable aléatoire absolument continue, alors on dit que X suit la loi normale $\mathcal{N}(m, \sigma)$ ou $LG(m, \sigma)$ si sa fonction densité est définie pour $x \in \mathbb{R}$ par:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\right) \left(-\frac{1}{2} \left(\frac{x-m}{\sigma} \right)^2 \right),$$

où m et σ sont respectivement, la moyenne et l'écart type. Dans ce cas on dit que X est normalement distribuée de moyenne m et de variance σ^2.

Caractéristiques de la loi normale

Soit X une variable aléatoire qui suit la loi normale $\mathcal{N}(m, \sigma)$, alors les fonctions de distribution et caractéristique correspondantes sont

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

et

$$\varphi_X(t) = \exp\left(\right) \left(imt - \left(\frac{\sigma t}{\sqrt{2}} \right)^2 \right).$$

Posons $U = \dfrac{X-m}{\sigma}$, alors la moyenne et la variance de U sont 0 et 1 respectivement, et la densité de probabilité de U est

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(\right) \left(-\frac{1}{2} u^2 \right),$$

et U sera appelée variable aléatoire normale centrée réduite, $U \to \mathcal{N}(0, 1)$.
La fonction de distribution de U notée $\Phi(u)$ est de la forme

$$\Phi(u) = P(U \leq u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} \exp\left(\right) \left(-\frac{1}{2} t^2 \right) dt = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_{0}^{u} \exp\left(\right) \left(-\frac{1}{2} t^2 \right) dt.$$

Exemple :

Soit U une variable aléatoire suivant la loi $\mathcal{N}(0, 1)$, alors
$P(-1 \leq U \leq 1) = \Phi(1) - \Phi(-1) = 0,6827$, l'aire à l'intérieur des écarts types à la moyenne égale à 68,27%.

Caractéristiques de la loi normale centrée réduite

Soit U une variable aléatoire suivant la loi $\mathcal{N}(0, 1)$, alors

$$E(U) = 0,\ Var(U) = 1,$$

$$\gamma_1 = 0,\ \gamma_2 = 3,$$

et la fonction caractéristique

$$\varphi_U(t) = \exp\left(\right)(-\frac{t^2}{2}\right),$$

et si X est une variable suivant la loi normale $\mathcal{N}(m, \sigma)$, alors

$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a-m}{\sigma} \leq \frac{X-m}{\sigma} \leq \frac{b-m}{\sigma}\right) \\ &= \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right). \end{aligned}$$

Approximation de la loi binomiale par la loi normale

Si n est grand , p et q ne sont pas voisines de zéro, alors la distribution binomiale peut être une trés bonne approximation de la distribution normale de la variable réduite.
Si $X \to \mathcal{B}(n, p)$, alors on a:

$$\lim_{n\to\infty} P\left(a \leq \frac{X - np}{\sqrt{np(1-p)}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(\right)(-\frac{1}{2}t^2\right) dt.$$

En d'autres termes, la variable aléatoire réduite $\dfrac{X - np}{\sqrt{np(1-p)}}$ est asymptotiquement normale.
En pratique, on donne la condition np et $n(1-p) > 5$ pour utiliser cette approximation.

Approximation de la loi de Poisson par la loi normale

Si X est une variable aléatoire de Poisson $\mathcal{P}(\lambda)$ alors la variable réduite correspondante $\dfrac{X-\lambda}{\sqrt{\lambda}}$ vérifie le résultat suivant:

$$\lim_{\lambda\to\infty} P\left(a \leq \frac{X-\lambda}{\sqrt{\lambda}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(\right)(-\frac{1}{2}t^2\right) dt.$$

En d'autres termes, la distribution de Poisson tend vers la distribution normale quand $\lambda \to \infty$, ou $\dfrac{X-\lambda}{\sqrt{\lambda}}$ est asymptotiquement normale.
En pratique on utilise cette approximation dés que $\lambda > 18$.

Loi de Cauchy

On dit que la variable aléatoire X suit la loi de cauchy si sa fonction densité est définie par:

$$f(x) = \frac{c}{\pi\left(x^2 + c^2\right)}, \quad c > 0, \ -\infty < x < \infty.$$

On remarque que cette distribution n'admet pas de moyenne, de variance et de moments d'ordres supérieurs.

Loi de khi-deux $\mathcal{X}^2$

Soient $X_1, X_2, ..., X_p$, p variables aléatoires indépendantes et normalement distribuées de moyenne 0 et de variance 1.

On appelle loi de khi-deux à p degrés de liberté la loi de la variable aléatoire $\sum_{1=1}^{p} X_i^2$.

Soit U une variable aléatoire qui suit la loi $\mathcal{N}(0,1)$, alors la variable aléatoire $Y = U^2$ suit la loi khi-deux à 1 degrés de liberté $\left(\mathcal{X}_1^2\right)$, et la fonction densité de la variable aléatoire Y est

$$f(y) = \frac{1}{\sqrt{2\pi}} y^{-1/2} \exp()(-t/2).$$

Puisque $\Gamma(1/2) = \sqrt{\pi}$, ce qui entraine que la variable al´eatoire $Y/2 \rightarrow \gamma(1, 1/2)$.

Théorème :

Soit X une variable aléatoire qui suit la loi $\gamma(1, p/2)$, alors la variable aléatoire $2X$ suit la loi X_p^2.
De ce théorème on peut montrer que la densité d'une variable aléatoire X qui suit la loi khi-deux à p degrés de liberté $\left(\mathcal{X}_p^2\right)$ est définie pour $x > 0$ par

$$f(x) = \frac{1}{2^{\frac{p}{2}}\Gamma(p/2)} x^{\frac{p}{2}-1} \exp()(-x/2)$$

Caractéristiques de la loi $\mathcal{X}_p^2$

Soit X une variable aléatoire qui suit la loi khi-deux à p degrés de liberté $\left(\mathcal{X}_p^2\right)$, alors

$$E(X) = p, \ Var(X) = 2p,$$

et

$$\varphi_X(t) = \frac{1}{(1-2it)^{p/2}}.$$

Approximation de la loi $\left(\mathcal{X}_p^2\right)$ par la loi normale

Soit X une variable aléatoire qui suit $\left(\mathcal{X}_p^2\right)$, alors on a les approximations suivantes:

La variable aléatoire $\dfrac{X-p}{\sqrt{2p}}$ est asymptotiquement normale. En d'autres termes, la distribution de $\left(\mathcal{X}_p^2\right)$ tend vers la distribution normale quand $p \to \infty$

Approximation de Ficher

Si $p \to \infty$, la variable aléatoire $\sqrt{2X} - \sqrt{2p-1}$ est normalement distribuée de moyenne 0 et de variance 1.Cette approximation est appelée approximation de Fisher.

Approximation de Wilson-Hilferty

Si $p \to \infty$, la variable aléatoire $\dfrac{\sqrt[3]{\frac{X}{p}} - \left(1 - \frac{2}{9p}\right)}{\sqrt{\frac{2}{9p}}}$, est normalement distribuée de moyenne 0 et de variance 1.

On peut établir que la meilleure approximation asymptotique est l'approximation de Wilson-Hilferty suivie par l' approximation de Ficher.

En pratique on applique ces deux dernières approximations dés que $p > 30$.

Théorème :

Si X et Y deux variables aléatoires indépendantes suivants respectivement les lois $\left(\mathcal{X}_p^2\right)$ et $\left(\mathcal{X}_q^2\right)$ alors $X+Y$ suit la loi $\left(\mathcal{X}_{p+q}^2\right)$.

Loi de student

Soient U et X deux variables aléatoires indépendantes suivant les lois $\mathcal{N}(0,1)$ et $\left(\mathcal{X}_p^2\right)$ respectivement, alors la variable aléatoire $T = \dfrac{U}{\sqrt{X/p}}$ suit la loi de Student à p degrés de liberté dont la densité de probabilité est définie pour $t \in \mathbb{R}$ par:

$$f(t) = \frac{\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{p\pi}\,\Gamma(p/2)} \left(1 + \frac{t^2}{p}\right)^{-(p+1)/2}.$$

Cas particulier, si $p = 1$ la loi de Student devient la loi de Cauchy

Caractéristiques de la loi Student à p degrés de liberté

Soit T une variable aléatoire qui suit la loi de Student à p degrés de liberté, alors:

$$E(T) = 0 \text{ si } p > 1, \text{ et } Var(X) = \frac{p}{p-2} \text{ si } p > 2.$$

Loi de Fisher-Snedecor

Soient X et Y deux variables aléatoires suivants $\mathcal{X}_p^2$ et $\mathcal{X}_q^2$ respectivement. On définit la variable aléatoire $F = \dfrac{X/p}{Y/q}$.
Pour trouver la fonction densit´e de F on ´enonce le th´eor`eme suivant
Théorème ::
Soient X et Y deux variables aléatoires suivants $\gamma(1, p)$ et $\gamma(1, q)$ respectivement, alors la variable aléatoire $Z = \dfrac{X}{Y}$ suit la bêta de type 2 de paramètres p,et q, d'où la densité de Z est

$$f_Z(z) = \frac{1}{B(p,q)} \frac{z^{p-1}}{(1+z)^{p+q}}.$$

Les variables aléatoires X et Y suivent les lois $\gamma(1, p/2)$ et $\gamma(1, q/2)$ respectivement, donc la variable aléatoire Z peut s'écrire sous la forme $Z = \dfrac{p}{q}F$, d′où la densité de F est:

$$g(f) = \begin{cases} \dfrac{1}{B(p,q)} \dfrac{\left(\dfrac{p}{q}\right)^{\dfrac{p}{2}}}{\left(1 + \dfrac{p}{q} f\right)} f^{\dfrac{p}{2}-1} & \text{si } f \geq 0 \\ 0 & \text{sinon,} \end{cases}$$

et on dit que la varaible F suit la loi de Fisher à p et q degrés de liberté.
Caractéristiques de la loi de Fisher-Snedecor à p et q degrés de liberté
Soit F une variable aléatoire qui suit la loi de Fisher-Snedecor à p et q degrés de liberté, alors

$$E(F) = \frac{q}{q-2} \ (q > 2), \ Var(F) = \frac{2q^2(p+q-2)}{p(q-2)^2(q-4)} \ (q > 4).$$

Loi exponentielle

Soit X une variable aléatoire qui suit la loi $\gamma(\lambda, p)$, alors si $p = 1$, on dit que X suit la loi exponentielle de paramètre $\lambda \ (\lambda > 0)$, et on note $X \longrightarrow \mathcal{E}(\lambda)$.
Caractéristiques de la loi exponentielle
Soit X une variable aléatoire qui suit la loi $\mathcal{E}(\lambda)$, alors

$$E(X) = 1/\lambda, \ Var(X) = 1/\lambda^2,$$

et

$$\varphi_X(t) = \frac{\lambda}{\lambda - it}.$$

Loi de Weibull

Posons $X = \dfrac{(Y)^{1/b}}{a}$ $(a > 0, b > 0)$. Si la variable aléatoire Y suit la loi $\mathcal{E}(1)$, alors la variable aléatoire X suit la loi de Weibull de paramètres a et b et de fonction densité:

$$f(x) = \begin{cases} abx^{b-1} \exp() \left(-ax^b\right) & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases}$$

Caractéristiques de la loi de Weibull

Soit X une variable aléatoire qui suit la loi de Weibull alors:

$$E(X) = a^{-1/b}\Gamma(1 + (1/b)), \ Var(X) = a^{-2/b}\left[\Gamma(1 + (2/b)) - \Gamma^2(1 + (1/b))\right].$$

Loi de Pareto

Une variable aléatoire absolument continue est dite suivre la loi de Pareto si sa fonction densité est définie pour $x \geq 1$ et $\alpha > 0$ par:

$$f(x) = \alpha x^{-(\alpha+1)}.$$

Caractéristiques de la loi de Pareto

Soit X une variable aléatoire qui suit la loi de Pareto, alors:

$$E(X) = \frac{\alpha}{\alpha - 1}, \ Var(X) = \frac{\alpha}{(\alpha - 1)^2 (\alpha - 2)}.$$

L'espérance et la variance existent simplement si $\alpha > 1$, et $\alpha > 2$ respectivement.

Convergence des suites de variables aléatoires

On va étudier dans ce chapitre le comportement asymptotique de suites de variables aléatoires, en énonçons quelques types de convergence de ces suites (convergence en probabilité, convergence prèsque sûre, convergence en moyenne d'ordre p, convergence en loi) ainsi les théorèmes limites (loi faible, loi forte des grands nombres).
On énonce deux inégalités qui sont trés importantes en théorie de probabilités surtout dans les cas de convergence.

Inégalité de Markov

Soit X une variable aléatoire de moyenne μ fini alors $\forall c > 1$ on a:

$$P(X \geq c) \leq \mu/c$$

Inégalité de Bienaymé Tchebycheff:

Soit X une variable aléatoire de moyenne μ et de variance σ^2 finies, alors on a pour tout réel strictement positif ε l'inégalité suivante:

$$P(|X-\mu| > \varepsilon) \leq \sigma^2/\varepsilon^2.$$

Convergence en probabilite´

Définition :

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires définies sur l'espace de probabilité $(\Omega, \mathcal{T}, P)$, alors on dit que cette suite converge en probabilité vers la constante a si

$$\forall\varepsilon > 0 \quad \lim_{n\to\infty} P(|X_n - a| < \varepsilon) = 0,$$

et on dit que cette suite converge en probabilité vers la variable X si

$$\forall\varepsilon > 0 \quad \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 0,$$

ou

$$\forall\varepsilon > 0 \quad \lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 1,$$

et on notera $X_n \xrightarrow{\mathcal{P}} X$.

On peut définir la convergence en probabilité de la suite $(X_n)_{n\in\mathbb{N}^*}$ vers une variable aléatoire X comme la convergence en probabilit´e de la suite $(X_n - X)_{n\in\mathbb{N}^*}$ vers 0.

Théorème :

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires, alors $X_n \xrightarrow{\mathcal{P}} X$ si et seulement si

$$E(X_n) \underset{n\to\infty}{\longrightarrow} X \text{ et } \sigma^2 = Var(X_n) \underset{n\to\infty}{\longrightarrow} 0.$$

Lorsque $E(X_n) \underset{n\to\infty}{\longrightarrow} X$, alors pour prouver que $X_n \xrightarrow{\mathcal{P}} X$, il suffit simplement de montrer que $Var(X_n) \underset{n\to\infty}{\longrightarrow} 0$.

Remarque:
On peut avoir la convergence en probabilité de la suite $(X_n)_{n\in\mathbb{N}^*}$ vers X sans que $E(X_n) \underset{n\to\infty}{\longrightarrow} X$ et $Var(X_n) \underset{n\to\infty}{\longrightarrow} 0$.

Loi faible des grands nombres

La loi faible des grands nombres est un r´esultat tr´es important du l'inégalité de Bienaymé- Tchebytchev
Théorème :
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables al´eatoires.ind´ependantes et identiquement distribu´ees (i-i-d) d'esp´erance μ et de variance σ^2.
Considérons la variable aléatoire $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, qui est appelée la moyenne empirique de $X_1, X_2, ..., X_n$, alors

$$\bar{X}_n \xrightarrow{\mathcal{P}} \mu,$$

Preuve

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu,$$

et

$$Var(\bar{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \sigma^2/n \underset{n\to\infty}{\longrightarrow} 0,$$

c'est à dire que quand $n \to \infty$,

$$P\left(\left|\bar{X}_n - \mu\right| > \varepsilon\right) \to 0.$$

Ce théorème énonce que la probabilité pour que la moyenne empirique $\bar{X}_n$ s'écarte de sa valeur espéreé μ de plus de ε, tend vers zéro quand n tend vers l'infini.

Convergence en moyenne d'ordre p

Définition :
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires.définies sur l'espace de probabilité $(\Omega, \mathcal{T}, P)$, avec la condition que $E(|X_n - X|^p)$ existe, alors on dit que cette suite converge en moyenne d'ordre p vers la variable aléatoire X si

$$E(|X_n - X|^p) \underset{n\to\infty}{\longrightarrow} 0.$$

Pour $p = 2$ on dit qu'on a une convergence en moyenne quadratique et on note $X_n \overset{m.q}{\to} X$.
On énonce quelques théorèmes consernants la convergence en moyenne quadratique

Théorème 01 :
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires.définies sur l'espace de probabilité $(\Omega, \mathcal{T}, P)$ tel que tous les moments d'ordre deux existent. Une condition nécéssaire et suffisante pour la convergence en moyenne quadratique de cette suite vers X est

$$\begin{array}{ccc} E(X_n) & \underset{n\to\infty}{\to} & E(X) \\ E\left(X_n^2\right) & \underset{n\to\infty}{\to} & E\left(X^2\right). \end{array}$$

Théorème 02 :
La convergence en moyenne quadratique entraine la convergence en probabilité, et la réciproque est fausse.

Convergence présque sure ou convergence forte

Définition :
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires.définies sur l'espace de probabilité $(\Omega, \mathcal{T}, P)$, alors on dit que $(X_n)_{n\in\mathbb{N}^*}$ converge prèsque sûrement vers X si

$$P\left(\omega \in \Omega / \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\right) = 0,$$

et on note $X_n \overset{ps}{\to} X$.
En d'autres termes l'ensemble des points de divergence est de probabilités nulle.
Théorème 01:
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires.définies sur l'espace de probabilité $(\Omega, \mathcal{T}, P)$. Si on pose $M_N = \sup_{n\geq N} |X_n - X|$, alors *une condition nécésaire et suffisante de la convergence prèsque sûre de la suite vers X est que*

$$M_N \overset{\mathcal{P}}{\to} 0.$$

C'est à dire que

$$X_n \overset{ps}{\to} X \Leftrightarrow M_N \overset{\mathcal{P}}{\to} 0.$$

Théorème 02:
La convergence prèsque sûre implique la convergence en probabilité, et la réciproque est fausse.

Loi forte des grands nombres

On améliorant le résultat de la loi faible des grands nombres: on a en fait, sous certaines hypothèses, la convergence presque sûre, et non pas seulement en probabilité, de la moyenne empirique $\bar{X}_n$ vers la moyenne μ. La loi forte des grands nombres nous donne une condition nécessaire et suffisante pour l'existence de la limite au sens de la convergence presque sˆure.

Théoréme:

Soit $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires indépendantes et identiquement distribuées (i-i-d) d'espérance* μ et intégrables *(*$\forall n \in \mathbb{N}^* : E(|X_n|) < \infty$), alors:

$$\bar{X}_n \overset{ps}{\rightarrow} \mu = E(X_1)$$

Convergence en loi

Cette convergence est la plus faible, mais elle est la plus utilisée en pratique car elle peut approximer la fonction de distribution de X_n par celle de X.

Définition :

On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ *converge en loi vers la variable aléatoire* X *de fonction de distribution* F *si en tout point de continuité de* F *la suite* (F_n) *des fonctions de distribution des* X_n *converge vers* F, *et on note* $X_n \overset{\mathcal{L}}{\rightarrow} X$.

Pour des variables discrètes la convergence en loi vers une variable discrète s'exprime par:

$$P(X_n = x) \underset{n\longrightarrow\infty}{\longrightarrow} P(X = x).$$

Un autre résultat montre également que si $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires.de fonctions densités f_n, et X est une variable aléatoire de fonction densité f alors

$$X_n \overset{\mathcal{L}}{\rightarrow} X \Longrightarrow f_n(x) \underset{n\longrightarrow\infty}{\longrightarrow} f(x) \quad \forall x.$$

La convergence en loi est tr´es li´ee a` la convergence des fonctions caract´eristiques comme le pr ´ecise le th´eor`eme fondamental suivant:

Théorème : de Paul Levy

Théorème 01:

Si une suite de variables al´eatoires $(X_n)_{n\in\mathbb{N}^*}$ *converge en loi vers la variable al´eatoire* X, *alors la fonction caract´eristique* $\varphi_n(t)$ *de* X_n *tend vers la fonction caract´eristique* $\varphi(t)$ *de* X. *R ´eciproquement si on a une suite de variables al´eatoires* $(X_n)_{n\in\mathbb{N}^*}$ *dont les fonctions caract ´eristiques* $\varphi_n(t)$ *tendent vers une fonction* $\varphi(t)$ *et si* $\varphi(t)$ *est continue au point* $t = 0$, *alors la suite* $(X_n)_{n\in\mathbb{N}^*}$ *converge en loi vers la variable al´eatoire* X *dont la fonction caract´eristique est* $\varphi(t)$.

Théorème 02 :

La convergence en probabilit´e entraine la convergence en loi. La r´eciproque est fausse en général.

Applications

Convergence de la loi binomiale vers la loi normale

Théoréme de De Moivre-Laplace

Théorème :

Soit $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires binomiales* $\mathcal{B}(n,p)$ *alors:*

$$\frac{X_n - np}{\sqrt{npq}} \overset{\mathcal{L}}{\rightarrow} \mathcal{N}(0,1).$$

Preuve:
La fonction caractéristique $\varphi_n(t)$ de X_n est égale à $(1-p+p\exp()(it))^n$, donc celle de la variable aléatoire $\frac{X_n-np}{\sqrt{npq}}$ est

$$\varphi(t)=\left(1-p+p\exp\left()(\frac{it}{\sqrt{np(1-p)}}\right)\right)^n\exp\left()(\frac{-inpt}{\sqrt{np(1-p)}}\right),$$

et

$$\ln\varphi(t)=n\ln\left(1-p\left(\exp\left()(\frac{it}{\sqrt{np(1-p)}}\right)-1\right)\right)-\frac{-inpt}{\sqrt{np(1-p)}}$$

En développant au second ordre l'exponentielle et puis le logarithme on trouve

$$\ln\varphi(t)\simeq n\ln\left(1-p\left(\frac{it}{\sqrt{np(1-p)}}-\frac{t^2}{2np(1-p)}\right)\right)-\frac{-inpt}{\sqrt{np(1-p)}},$$

$$\ln\varphi(t)\simeq n\left(\frac{pit}{\sqrt{np(1-p)}}-\frac{pt^2}{2np(1-p)}+\frac{p^2t^2}{2np(1-p)}\right)-\frac{-inpt}{\sqrt{np(1-p)}},$$

soit

$$\ln\varphi(t)\simeq-\frac{pt^2}{2p(1-p)}+\frac{p^2t^2}{2p(1-p)}=\frac{-t^2p(1-p)}{2p(1-p)}=\frac{-t^2}{2}$$

ce qui montre que $\varphi(t)\longrightarrow\exp\left()(-\frac{t^2}{2}\right)$ qui est la fonction caractéristique de la loi normale $\mathcal{N}(0,1)$.

Convergence de la loi binomiale vers la loi de Poisson.

Théorème :

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires binomiales $B(n,p)$ alors

$$X_n\xrightarrow{\mathcal{L}}\mathcal{P}(np)$$

Preuve

Supposons que $\lambda=np$, alors

$$\begin{aligned}\varphi_{X_n}(t) &= (1-p+p\exp()(it))^n &&= \left(1-\tfrac{\lambda}{n}+\tfrac{\lambda}{n}\exp()(it)\right)^n\\ &= \left(1-\tfrac{\lambda}{n}(1-\exp()(it))\right)^n &&= \left[\left(1-\tfrac{\lambda}{n}(1-\exp()(it))\right)^{\frac{-n}{\lambda(1-\exp()(it))}}\right]^{-\lambda(1-\exp()(it))}\\ \underset{n\to\infty}{\longrightarrow}\ & \exp(\lambda)(\exp()(it)-1) &&= \varphi_X(t).\end{aligned}$$

et on a de plus la fonction $\varphi_X(t)$ est continue en 0 d'où le résultat.

Convergence de la loi de Poisson vers la loi normale

Théorème :
Soit X une variable aléatoire qui suit la loi de Poisson $P(\lambda)$, alors quand $\lambda \to \infty$ la variable aléatoire

$$\frac{X-\lambda}{\sqrt{\lambda}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

Preuve
On a $\varphi_X(t) = \exp()\,(\lambda(\exp()\,(it)-1))$, donc la fonction caractéristique de la variable aléatoire $\dfrac{X-\lambda}{\sqrt{\lambda}}$ est

$$\begin{aligned}\varphi_{\frac{X-\lambda}{\sqrt{\lambda}}}(t) &= \exp\left(\right)\left(\lambda\left(\exp\left(\right)\left(\tfrac{it}{\sqrt{\lambda}}\right)-1\right)\right)\exp\left(\right)\left(\frac{-it\lambda}{\sqrt{\lambda}}\right)\\ &= \exp\left(\right)\left(\lambda\exp\left(\right)\left(\tfrac{it}{\sqrt{\lambda}}\right)-\lambda-it\sqrt{\lambda}\right),\end{aligned}$$

comme

$$\exp\left(\right)\left(\frac{it}{\sqrt{\lambda}}\right) \simeq 1+\frac{it}{\sqrt{\lambda}}-\frac{t^2}{2\lambda}$$

il vient que

$$\varphi_{\frac{X-\lambda}{\sqrt{\lambda}}}(t) \simeq \exp\left(\right)\left(\lambda\left(1+\frac{it}{\sqrt{\lambda}}-\frac{t^2}{2\lambda}\right)-\lambda-it\sqrt{\lambda}\right) = \exp\left(\right)\left(-\frac{t^2}{2}\right).$$

Theorème central limite

On a vu pécèdement qu'on peut approximer les lois binômiale et de Poisson par la loi normale, donc il est trés évident de poser la question de savoir s'il existe d'autres distributions de variance et de moyenne finies qui seront approxim´ees par la loi n ormale. La r´eponse a` cette question est le fameux théorème appelé théorème central limite

Théorème ::
Soient $X_1, X_2, ..., X_n$ des variables aléatoires indépendantes et identiquement distribuées (de méme loi) de moyenne m et de variance σ^2 existent et finies, alors si $S_n = X_1 + X_2 + \ldots + X_n$

$$\lim_{n\to\infty} P\left(a \leq \frac{S_n - nm}{\sigma\sqrt{n}} \leq b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b \exp\left(\right)\left(-\frac{1}{2}t^2\right)dt.$$

C'est à dire que la variable aléatoire réduite $Y_n = \dfrac{S_n - nm}{\sigma\sqrt{n}}$ est asymptotiquement normale.

Preuve
On pose φ_{Y_n}la fonction caractéristique de Y_n, et montre que $\varphi_{Y_n}(t) \underset{n\to\infty}{\longrightarrow} \exp\left(\right)\left(-\frac{t^2}{2}\right)$.

$$\begin{aligned}
\varphi_{Y_n}(t) &= E\left(\exp\left(\right)\left(\frac{it}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} X_i - nm\right)\right)\right) \\
&= \exp\left(\right)\left(-it\left(\frac{\sqrt{n}m}{\sigma}\right)\right) E\left(\exp\left(\right)\left(\frac{it}{\sigma\sqrt{n}}\sum_{i=1}^{n} X_i\right)\right) \\
&= \exp\left(\right)\left(-it\left(\frac{\sqrt{n}m}{\sigma}\right)\right)\left[\varphi_{X_i}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n,
\end{aligned}$$

où φ_{X_i} est la fonction caractéristique de X_i.

$$\varphi_{X_i}\left(\frac{t}{\sigma\sqrt{n}}\right) \simeq 1 + \frac{itm}{\sigma\sqrt{n}} - \frac{t^2\mu_2'^2}{2\sigma^2 n} + \ldots$$

$$\begin{aligned}
\ln\varphi_{Y_n}(t) &= -it\frac{\sqrt{n}m}{\sigma} + n\ln\varphi_{X_i}\left(\frac{t}{\sigma\sqrt{n}}\right) \\
&\simeq -it\frac{\sqrt{n}m}{\sigma} + n\ln\left(1 + \frac{itm}{\sigma\sqrt{n}} - \frac{t^2\mu_2'^2}{2\sigma^2 n} + \ldots\right) \\
&= -it\frac{\sqrt{n}m}{\sigma} + n\left(\frac{itm}{\sigma\sqrt{n}} - \frac{t^2\mu_2'^2}{2\sigma^2 n} + \frac{t^2 m^2}{2\sigma^2 n} + \ldots\right)
\end{aligned}$$

$$\ln\varphi_{Y_n}(t) \simeq -\frac{t^2\left(\mu_2'^2 - m^2\right)}{2\sigma^2} + \text{une infinité de termes négligeables,}$$

donc

$$\forall t \in \mathbb{R} : \ln\varphi_{Y_n}(t) \underset{n\to\infty}{\longrightarrow} \frac{-t^2}{2}$$

et

$$\varphi_{Y_n}(t) \underset{n\to\infty}{\longrightarrow} \exp\left(\right)\left(\frac{-t^2}{2}\right),$$

d'où le resultat.

References

[1] Philippe Tassi: Méthodes statistiques 2^e Eddition , Economica 1989..

[2] Kedjdal Kaci: Cours de probabilités

[3] G. Saporta: probabilités, Analyse des données et statistiques. Edition Technip. 1990.

[4] Murray R. Spiege: Probabilités et Statistique Cours et Problèmes Série Schaum.

[5] Sayah Abdallah: Cours de probabilités et statistiques 2. (PS2) 4^e DES Mathématiques.