Cours de probabilités

Ibrahim AMRANI

Préambule

L'objectif de ce cours est de présenter les principaux fondamentaux de l'aspect aléatoire de la statistique. C'est un cours de base dont la compréhension permet de suivre le cours de statistique inférentielle. Il est à préciser que ce cours est suivi d'un cours de simulation qui présente des applications informatiques des notions de probabilité. La manipulation informatique de ces notions suppose une certaine maitrise de la théorie de ces concepts.

Le cours est constitué de cinq chapitres :
Le premier chapitre traite la notion de probabilité.
Le deuxième chapitre étudie le concept mathématique de la variable aléatoire et des lois de probabilités.
Le troisième chapitre présente les caractéristiques numériques des variables aléatoires.
Le quatrième chapitre étudie des opérateurs mathématiques qui présentent des propriétés très intéressantes.
Le cinquième chapitre aborde le comportement asymptotique des variables aléatoires et présente des théorèmes importants de la statistique.

1. Probabilité

La probabilité peut être définie comme étant la mesure de la chance de réalisation d'un certain événement et cette mesure est comprise entre 0 et 1.

Expérience stochastique, Evénement

Une expérience est dite stochastique ou aléatoire, s'il est impossible de prévoir son résultat. On admet qu'elle peut être répétée indéfiniment dans des conditions identiques, son résultat peut donc varier d'une réalisation à l'autre.

Exemples :

1. On jette un dé et observe le résultat obtenu : $\{1, 2, \ldots, 6\}$
2. On lance trois fois de suite une pièce de monnaie : 8 résultats possibles : PPP, PPF, …, FFF.
3. On prélève un lot de n pièces et on s'intéresse au nombre de pièces défectueuses.
4. On jette un dé numéroté et on s'intéresse au nombre de jets nécessaires jusqu'à ce que le n°6 apparaît.
5. Dans une station de bus, on mesure la durée du temps qui sépare le départ d'un bus et l'arrivée du prochain.

Ensemble fondamental, événement

L'ensemble, noté Ω, de tous les résultats possibles d'une expérience aléatoire est dit ensemble fondamental. Selon la nature de cette expérience Ω peut être :

- fini (exemples 1, 2, 3)
- infini dénombrable (exemple 4)
- infini non dénombrable (exemple 5)

On appelle événement tout sous ensemble de Ω.

Un événement A est réalisé lors d'une expérience aléatoire, si le résultat de cette dernière appartient à A.

Si $A = \{w\}$, on parle d'événement élémentaire.
Ø l'ensemble vide est l'événement impossible qui n'est jamais réalisé.
Ω est l'événement certain qui est toujours réalisé.

Opérations sur les événements

Les événements associés à une expérience aléatoire, sont des sous-ensembles de l'ensemble Ω, il est naturel de définir des opérations sur les événements à l'image des opérations bien connues en théorie des ensembles :

Intersection : $AB = A \cap B \Leftrightarrow \forall w \in \Omega,\ w \in A$ et $w \in B$
Réunion : $A + B = A \cup B \Leftrightarrow \forall w \in \Omega,\ w \in A$ ou $w \in B$
Inclusion : $A \subset B \Leftrightarrow \forall w \in \Omega,\ w \in A$ alors $w \in B$
$\overline{A}$ est l'événement complémentaire de A dans $\Omega \Leftrightarrow \forall w \in \Omega,\ w \in A$ alors $w \notin \overline{A}$
Commutativité : $A + B = B + A$
Associativité : $A + (B+C) = (A+B) + C$
Distributivité : $A(B+C) = AB + AC$
Complémentarité : $\overline{\overline{A}} = A$
Loi de Morgan : $\overline{AB} = \overline{A} + \overline{B}$
$A\Omega = A$
$A\overline{A} = \emptyset$

Evénements mutuellement exclusifs

Une famille d'événements (A_i), $i = 1,\ldots,n$, est composé d'événements mutuellement exclusifs si et seulement si :

$$A_iA_j = \begin{cases} A_i \text{ si } i = j \\ \emptyset \text{ si } i \neq j \end{cases}$$

Un élément de Ω n'est pas inclus dans plus d'un événement de la liste

Evénements collectivement exhaustifs

Une famille d'événements (A_i), $i = 1,\ldots, n$, est composée d'événements collectivement exhaustifs si et seulement si :

$$A_1 + A_2 + \ldots + A_n = \Omega$$

Un élément de Ω est inclus dans au moins un événement de la liste.

Probabilité, Espace probabilisé

Tribu (ou σ-algèbre) : Définition

Soit Ω un ensemble. On appelle **tribu** (ou **σ-algèbre**) des parties de Ω, tout sous ensemble $\mathcal{A}$ de $\mathcal{P}(\Omega)$ tel que :
1) $\Omega \in \mathcal{A}$
2) Pour tout $A \in \mathcal{A}$, $\overline{A} \in \mathcal{A}$
3) Pour toute famille finie ou infinie dénombrable (A_i) $i \in I$ $(I \subset N)$ d'éléments de $\mathcal{A}$,
$\bigcup_{i \in I} A_i \in \mathcal{A}$
$(\Omega, \mathcal{A})$ est alors appelé espace probabilisable.

Exemple
Si $A \subset \Omega$, $A \neq \emptyset$ et $A \neq \Omega$ alors $\{\emptyset, A, \overline{A}, \Omega\}$ est une tribu des parties de Ω

En pratique, lorsque Ω est fini ou dénombrable, alors $\mathcal{A} = P(\Omega)$.

Proposition

Soit $\mathcal{A}$ une tribu des parties de Ω
(1) $\emptyset \in \mathcal{A}$
(2) Pour tout $(A, B) \in \mathcal{A}^2$, $A \cup B$, $A \cap B$, $A \backslash B$ sont des éléments de $\mathcal{A}$
(3) Pour toute famille finie ou dénombrable (A_i) $i \in I$ ($I \subset N$) d'éléments de $\mathcal{A}$, $\bigcap_{i \in I} A_i \in \mathcal{A}$

Démonstration

(1) $\emptyset = \overline{\Omega}$
(2) Pour tout $(A, B) \in \mathcal{A}^2$, $A \cup B \in \mathcal{A}$ par définition
$A \cap B = \overline{\overline{A} \cup \overline{B}}$ or $(\overline{A}, \overline{B}) \in \mathcal{A}^2$, donc $\overline{A} \cup \overline{B} \in \mathcal{A}$ et $\overline{\overline{A} \cup \overline{B}} \in \mathcal{A}$
$A \backslash B = A \cap \overline{B} \in \mathcal{A}$
(3) $\bigcap_{i \in I} A_i = \overline{\bigcup_{i \in I} \overline{A_i}}$
Or pour tout $i \in I$, $\overline{A_i} \in \mathcal{A}$, donc $\bigcup_{i \in I} \overline{A_i} \in \mathcal{A}$ et $\overline{\bigcup_{i \in I} \overline{A_i}} \in \mathcal{A}$

Définition :
Soit $(\Omega, \mathcal{A})$ un espace probabilisable, on appelle **probabilité** sur $(\Omega, \mathcal{A})$ tout application P de $\mathcal{A}$ vers $[0, 1]$, vérifiant les axiomes suivants :
(1) $P(\Omega) = 1$
(2) Pour toute famille finie d'événements mutuellement exclusifs, $A_1, A_2, \dots, A_n$, on a
$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$

Ces deux axiomes sont suffisants pour introduire une probabilité lorsque l'ensemble fondamental Ω est fini.
Si Ω est infini :
(2') Pour toute famille infinie dénombrable d'événements mutuellement exclusifs, A_i :

$$P\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} P(A_i)$$

$(\Omega, \mathcal{A}, P)$ est appelé espace probabilisé et pour tout $A \in \mathcal{A}$, $P(A)$ est la **probabilité** de l'événement A.

La série de terme général $P(A_i)$ est convergente car quel que soit le rang r :

$$\sum_{i=r}^{+\infty} P(A_i) \leq 1$$

Propriétés élémentaires d'une loi de probabilité

(1) $P(A) = 1 - P(\overline{A})$ donc $P(\emptyset) = 0$.

(2) $P(A) \leq P(B)$ si $A \subset B$: $B = A + A\overline{B}$

(3) $P(A+B) = P(A) + P(B) - P(AB)$
$B \setminus A$: $w \in B$ et $w \in A$
$B = AB + B\setminus A \rightarrow P(B) = P(AB) + P(B\setminus A)$
$A + B = A + B\setminus A \rightarrow P(A+B) = P(A) + P(B\setminus A)$

(4) $P(A + B + C) = P(A) + P(B) + P(C) + P(ABC) - (P(AB) + P(BC) + P(CA))$
(4) peut être déduite de (3).

(5) Soit A_i $i = 1,\ldots,n$ une famille de n événements mutuellement exclusifs, et collectivement exhaustifs on a :

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) = 1$$

Si les événements A_i sont équiprobables alors $P(A_i) = 1/n$, $i = 1,\ldots,n$
Un événement E constitué par r événements A_i, vérifie alors :

$$P(E) = P(\bigcup_{i=1}^{r} A_i) = \sum_{i=1}^{r} P(A_i) = \frac{r}{n} = \frac{\text{nbre de cas favorables}}{\text{nbre de cas possibles}}$$

Probabilité Conditionnelle

Théorème et définition

Soient A et B deux événements de l'espace Ω avec $P(B) \neq 0$.

$$P_B(A) = \frac{P(AB)}{P(B)}$$

P_B est une probabilité définie sur les sous ensembles de Ω, appelée probabilité conditionnelle sachant B.
$P_B(A)$ est aussi noté $P(A/B)$.

P_B est une probabilité car elle vérifie les 3 axiomes :

(1) $0 \leq P_B(A) \leq 1$ car $P(AB) \leq P(B)$

(2) $P_B(\Omega) = \dfrac{P(B\Omega)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$

(3) Soit A_i, $i \in N$, une famille d'événements mutuellement exclusifs

$$P_B(\bigcup_{i=1}^{\infty} A_i) = \frac{P(B\bigcup_{i=1}^{n} A_i)}{P(B)} = \frac{P(\bigcup_{i=1}^{n} BA_i)}{P(B)} = \frac{\sum_{i=1}^{n} P(BA_i)}{P(B)} = \sum_{i=1}^{n} P_B(A_i)$$

, les événements BA_i, $i \in N$ sont mutuellement exclusifs.

Théorème des probabilités totales

Soit A_i $i = 1, \ldots, n$, une famille d'événements mutuellement exclusifs et collectivement exhaustifs. Pour tout événement A on a :

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

En effet :

$$A = A\Omega = A\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} AA_i$$

Les événements AA_i sont mutuellement exclusifs :

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

Théorème de multiplication

Soit A_i $i = 1,\ldots,n$ une famille d'événements telle que $P(A_1\ A_2 \ldots A_n) \neq 0$

$P(A_1\ A_2 \ldots A_n) = P(A_1)P(A_2/A_1)P(A_3/A_1A_2)\ldots.P(A_n/A_1 \ldots A_{n-1})$

Exemple

On dispose d'un lot comprenant 10 pièces dont la moitié est défectueuse. On prélève sans remise un échantillon de taille 3. Quelle est la probabilité pour que l'échantillon ne comprenne aucune pièce défectueuse ?

Solution

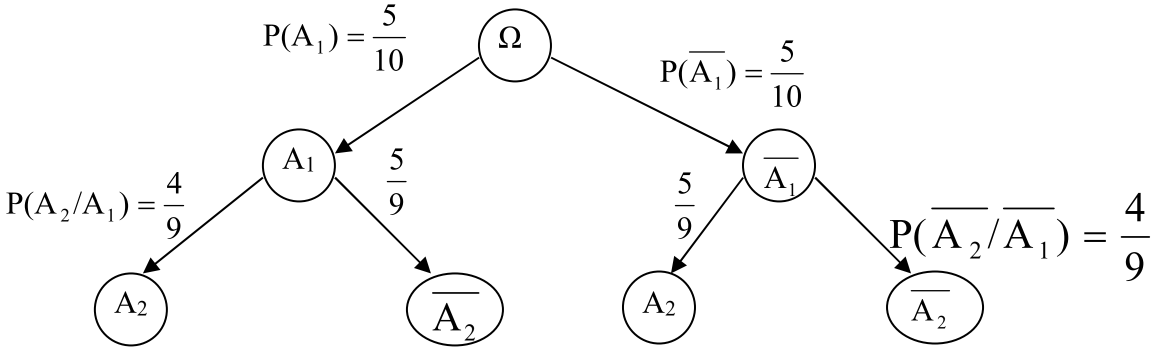
Soit A_i l'événement tel que la ième pièce prélevée soit bonne $i = 1, 2, 3$

$P(A) = P(A_1\ A_2\ A_3) = P(A_1)P(A_2/A_1)P(A_3/A_1A_2)$

$$P(A) = \frac{5}{10}\frac{4}{9}\frac{3}{8} = \frac{1}{12}$$

Arbre des probabilités

L'arbre des probabilités est un moyen qui permet de décrire et d'étudier des expériences stochastiques qui se déroulent en plusieurs étapes.
Reprenons l'exemple précédent, dans lequel il s'agit de prélever sans remise, trois pièces et intéressons nous à la probabilité que la deuxième soit bonne :



$$P(A_2) = \frac{5}{10}\frac{4}{9} + \frac{5}{10}\frac{5}{9} = \frac{1}{2}$$

Théorème de Bayes

Soit A_i i = 1, …, n, une famille d'événements mutuellement exclusifs et collectivement exhaustifs et B un événement tel que P(B) ≠ 0 :

$$P(A_i / B) = \frac{P(A_i B)}{P(B)} = \frac{P(B / A_i) P(A_i)}{\sum_{i=1}^{n} P(B / A_i) P(A_i)}$$

Exemple : illustration du théorème de Bayes

Trois machines A, B et C produisent respectivement 50%, 30% et 20% du nombre total des pièces fabriquées dans une usine. Les pourcentages des pièces défectueuses de ces machines sont 3%, 4% et 5%.

(1) Calculer la probabilité pour qu'une pièce prise au hasard soit défectueuse ?

(2) On choisi une pièce défectueuse, quelle est la probabilité pour qu'elle soit produite par la machine A ?

Solution

(1) Soit X l'événement correspondant à une pièce défectueuse,
P(X) = P(XA + XB + XC) = P(XA) + P(XB) + P(XC)
= P(X/A)P(A) + P(X/B)P(B) + P(X/C)P(C)
P(X) = 0,03*0,5 + 0,04*0,3 + 0,05*0,2 = 0,037

$$\text{(2) } P(A/X) = \frac{P(X/A)P(A)}{P(X)} = \frac{0{,}03 * 0{,}5}{0{,}037} = \frac{15}{37}$$

Indépendance des événements

Définition : Intuitivement, deux événements A et B sont dits indépendants, si la probabilité de l'un n'est pas influencée par la réalisation ou la non réalisation de l'autre.

En terme de probabilité conditionnelle :
P(A/B) = P(A) ou P(B/A) = P(B) à condition que P(A) ≠ 0 et P(B) ≠ 0

Soit que : P(AB) = P(A) P(B) cette relation ne nécessite pas que P(A) ≠ 0 et P(B) ≠

Exemples

(1) On jette deux fois de suite une pièce de monnaie pour obtenir pile ou face. Les deux jets se font dans les mêmes conditions et le résultat du premier jet est indépendant du second.

(2) On tire sans remise deux fois de suite une carte d'un jeu de 52 cartes
A : Obtenir un as au 1er tirage
B : Obtenir un valet au second tirage

Les conditions de réalisations de A et de B ne sont pas les mêmes :
A : choisir une carte parmi 52
B: choisir une carte parmi 51 et P(B) dépend de P(A)

Remarques

Si A et B sont indépendants alors il en est de même pour A et $\overline{B}$, B et $\overline{A}$, $\overline{B}$ et $\overline{A}$.
Pour démontrer ces propriétés, il suffit d'écrire la relation : $P(A) = P(AB) + P(A\overline{B})$

Il faut distinguer entre deux événements indépendants et deux événements exclusifs :
Si P(AB) = 0 alors nécessairement A et B sont dépendants si P(A) et P(B) sont différents de 0.

Si $B \subset A$ et $P(A) < 1$, alors A et B sont dépendants.

Indépendance de n événements :

Soit $A_1,\dots,A_n$, n événements, on dit que $A_1,\dots,A_n$ sont deux à deux indépendants si pour tout i et j de $1,\dots,n$ tels que $i \neq j$, $P(A_iA_j) = P(A_i)\,P(A_j)$.

On dit que $A_1,\dots,A_n$ sont mutuellement indépendants si pour tout ensemble I choisi parmi $1,\dots,n$: $P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$

Si une famille d'événements sont mutuellement indépendants alors nécessairement, ils sont deux à deux indépendants. La réciproque n'est pas toujours vérifiée, comme l'indique le contre exemple suivant :

$\Omega = \{w1, w2, w3, w4\}$ avec P(w1) = P(w2) = P(w3) = P(w4) = 1/4
A = {w1, w2}, B = {w1, w3} et C = {w1, w4}
P(AB) = P(w1) = 1/4 = P(A) P(B)
P(AC) = P(w1) = 1/4 = P(A) P(C)
P(BC) = P(w1) = 1/4 = P(B) P(C)
Car P(A) = P(B) = P(C) = 1/2
$P(ABC) = P(w1) = 1/4 \neq P(A)P(B)P(C)$
Les événements A, B et C sont deux à deux indépendants mais ne sont pas mutuellement indépendants.

Indépendance conditionnelle :

La notion d'indépendance de deux événements est liée à une probabilité et cette probabilité peut être conditionnelle

Soient A, B, C trois événements d'un ensemble fondamental Ω sur lequel on définit une probabilité conditionnelle sachant C ($P(C) \neq 0$).

A et B sont conditionnellement indépendants sachant C si et seulement si P(AB/C) = P(A/C)P(B/C).

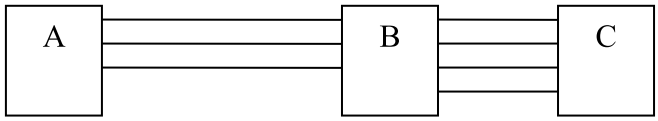
L'indépendance conditionnelle n'implique pas et n'est pas impliqué par une indépendance liée à une autre probabilité (voir exercice 2 du chapitre 3).

Analyse combinatoire

L'analyse combinatoire est une branche des mathématiques qui étudie les techniques de dénombrements particulièrement utiles en théorie des probabilités.

Principe de multiplication

Pour aller d'une ville A à une ville B, il existe trois chemins possibles, et pour aller de la ville B à la ville C, il en existe quatre.
Quel est le nombre total de chemins possibles pour aller de A vers C ?



Un tel calcul utilise le principe de multiplication :
En effet, pour chaque chemin possible entre A et B, il existe quatre possibilités d'aller de B vers C. Le nombre total de chemins possibles entre les villes A et C est douze.

Théorème :

Considérons n opérations successives, si chaque opération k peut se dérouler de m_k manières différentes (k = 1,…,n), alors les n opérations peuvent être effectuées dans l'ordre indiqué de $m_1 m_2 \ldots m_n$ manières différentes.

Dénombrement sans répétition

Permutations

Etant donné n objets discernables (a,b, … ,l), on appelle permutation d'ordre n toute suite formée de ces n objets, dans un ordre choisi arbitrairement, deux permutations correspondant à des ordres différents sont considérées comme différentes.

Exemple : le nombre de permutations qu'il est possible de faire avec trois éléments a, b et c est 6 : (a,b,c), (a,c,b), (b,a,c), (b,c,a), (c,a,b), (c,b,a)
$P_3 = 3*2*1 = 6 = 3!$

Le nombre de permutations différentes, d'ordre n est égale à $P_n = n*(n-1) *\ldots* 2*1 = n!$

Arrangements

Etant donné n objets discernables (a,b, … ,l) et un entier $p \leq n$, on appelle arrangement d'ordre p, toute suite formée par p objets choisis parmi (a,b, … ,l), un objet choisi ne peut être utilisé

une deuxième fois, deux suites ne différant que par l'ordre des objets sont considérées comme différentes.
Par exemple, (a,c,d), ((a,d,c), (b,c,e) sont des arrangements d'ordre 3, distincts de (a,b,c,d,e).

Le nombre d'arrangements d'ordre p, de n objets, est égale à :

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

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

Exemple : le nombre d'arrangements d'ordre 2 des trois éléments a,b et c sont : (a,b), (b,a), (b,c), (c,b), (a,c), (c,a).

$$A_3^2 = 3*2 = \frac{3!}{(3-2)!} = 6$$

Remarque : $A_n^n = P_n = n!$, ceci conduit à admettre que 0! = 1.

Combinaisons

Etant donné n objets discernables (a,b, … ,l) et un entier p ≤ n, on appelle combinaison d'ordre p, de ces n objets, un ensemble de p objets choisis parmi (a, b, … ,l), deux ensembles ne différant que par l'ordre des objets choisis sont considérés comme identiques.

Par exemple {a,e,c}, {a,c,e}, {c,a,e} représentent une même combinaison d'ordre 3.

Le nombre de combinaisons d'ordre p, de n objets, est égale à : $C_n^n = \binom{n}{p} = \frac{A_n^p}{P_n} = \frac{n!}{(n-p)!\,p!}$

Remarque : $C_n^n = 1 = \frac{n!}{0!\,n!}$, on admet que 0! = 1.

Dénombrement avec répétition

Permutations

Etant donné α objets a, indiscernables
β objets b, indiscernables
…
λ objets l, indiscernables
avec α + β + … + λ = n.

Une permutation de ces n objets, est une suite formée avec ces n objets, dans un ordre choisi arbitrairement, deux permutations diffèrent par l'ordre de succession des objets discernables, par contre l'échange de deux a, ou de deux b, … ne conduit pas à une nouvelle permutation.

Exemple : considérons l'ensemble a, a, b, b, b, c
Les permutations comportant c à la dernière place sont :

a a … c	. a a .. c	.. a a . c	… a a c
a . a .. c	. a . a . c	.. a . a c	
a .. a . c	. a .. a c		
a … a c			

Les b sont remplacés par des '.' pour une meilleure lisibilité.
Il était possible de mettre c dans chacune des cinq autres places. Le nombre total des permutations est 10*6 = 60.
D'une manière générale, le nombre des permutations des n objets est égale à :

$$P_{\alpha,\beta,\ldots,\lambda} = \frac{n!}{\alpha!\beta!\ldots\lambda!} = \frac{(\alpha+\beta+\ldots+\lambda)!}{\alpha!\beta!\ldots\lambda!}$$

Dans l'exemple : $P_{2,3,1} = \frac{6!}{2!3!1!} = \frac{4*5*6}{2!} = 60$

Arrangements

Etant donné n objets discernables (a_1, a_2, … ,a_n) et un entier p, on appelle arrangement avec répétition, une suite formée par p objets choisis parmi (a_1, a_2, … ,a_n) et chacun des objets pouvant être choisi jusqu'à p fois. Deux arrangements qui différent que par l'ordre des objets discernables sont considérés comme différents.

Exemple : étant donné dix chiffres 0, 1, 2, … , 9, les arrangements avec répétition d'ordre 3 sont les nombres entiers compris entre 000 et 999.

Le nombre d'arrangements avec répétition est $\mathrm{A}_n^p = n^p$

Preuve : démonstration par récurrence sur p.

Si p = 1 $\mathrm{A}_n^1 = n = n^1$

Supposons que $\mathrm{A}_n^p = n^p$, il y a n^p arrangements avec répétition d'ordre p des n objets.
Pour chaque arrangement d'ordre p, il y a n possibilités d'ajouter 1 objet pour obtenir un arrangement d'ordre p+1.
Donc $\mathrm{A}_n^{p+1} = n\mathrm{A}_n^p = n^{p+1}$

Combinaisons

Etant donné n objets discernables (x_1, x_2, … ,x_n) et un entier p, on appelle combinaison avec répétition CAR d'ordre p de ces n objets, un ensemble de p objets choisis parmi (x_1, x_2, … ,x_n), le même objet pouvant être choisi jusqu'à p fois. Deux combinaisons ne différant que par l'ordre des objets sont considérées comme identiques.
Exemple :
Les CAR d'ordre 2 de 3 objets (a,b,c) sont : aa, bb, cc, ab, ac, bc.

Le nombre de CAR d'ordre p de n objets est noté : K_n^p

$K_n^1 = n$, le nombre de CAR d'ordre 1 de n objets.

$K_n^2 = C_n^2 + n$, c'est le nombre de combinaison d'ordre 2, plus le nombre de couples constitués du même élément.

$K_n^3 = C_n^3 + 2C_n^2 + n$, c'est le nombre de combinaison d'ordre 3, plus le nombre de triplets ou chaque élément est répété deux fois, plus le nombre de triplets constitués du même élément.

Le nombre total d'éléments dans tous les CAR d'ordre p, choisis parmi n éléments est : pK_n^p

Chaque élément figure $\frac{p}{n}K_n^p$ dans les CAR d'ordre p choisies parmi n éléments.

Soit x un objet donné parmi les n objets et considérons le nombre de fois ou x figure dans les combinaisons d'ordre p et les combinaisons d'ordre p-1, on a la relation suivante :

$$\frac{p}{n}K_n^p = K_n^{p-1} + \frac{(p-1)}{n}K_n^p$$

K_n^{p-1} : Le nombre de fois où l'objet x est ajouté aux CAR d'ordre p-1.

$\frac{(p-1)}{n}K_n^p$: Le nombre de fois où l'objet x figure dans les CAR d'ordre p-1.

$$\frac{p}{n}K_n^p = K_n^{p-1} + \frac{(p-1)}{n}K_n^{p-1}$$

$$K_n^p = \frac{(n+p-1)}{p}K_n^{p-1}$$

$$K_n^p = \frac{(n+p-1)}{p}\frac{(n+p-2)}{p-1}\cdots\frac{n+1}{2}K_n^1 = \frac{(n+p-1)}{p}\frac{(n+p-2)}{p-1}\cdots\frac{n+1}{2}n$$

$$K_n^p = \frac{(n+p-1)!}{p!(n-1)!} = C_{n+p-1}^p$$

Exercices

Exercice 1
Utiliser les axiomes de l'algèbre des événements pour démontrer les relations suivantes :
1) $\overline{E} = \varnothing$
2) $A + B = A + \overline{A}B + ABC$
3) $A + E = E$
4) $P(A) = \sum_{i=1}^{n} P(AE_i)$, pour tout ensemble E_i, $i = 1,\ldots,n$, de parties mutuellement exclusives et collectivement exhaustives.
5) $P(A_1A_2\ldots A_n) = P(A_1)P(A_2/A_1)P(A_3/A_1A_2)\ldots P(A_n/ A_1A_2\ldots A_{n-1})$

Exercice 2
Expliquer en détail vos réponses aux questions suivantes :
1) Si les événements A et B sont mutuellement exclusifs et collectivement exhaustifs, est-ce que $\overline{A}$ et $\overline{B}$ les ont aussi ?

2) Si les événements A et B sont mutuellement exclusifs mais pas collectivement exhaustifs, est-ce que $\overline{A}$ et $\overline{B}$ sont collectivement exhaustifs ?

3) Si les événements A et B sont collectivement exhaustifs mais pas mutuellement exclusifs, est-ce que $\overline{A}$ et $\overline{B}$ sont collectivement exhaustifs ?

Exercice 3
Pour trois jets d'une pièce de monnaie, supposée parfaite, déterminer les probabilités :
1) d'obtenir PPP (P = ''pile'', F = ''face''),
2) d'obtenir PFP,
3) d'avoir deux piles et une face,
4) d'avoir plus de piles que de faces.
Déterminer la probabilité conditionnelle de :
5) ''plus de piles que de faces'' sachant que ''au moins une face'',
6) ''plus de piles que de faces'' sachant que ''moins de deux faces''.

Exercice 4
Casimir est un fou avec une probabilité de 0.6, un voleur avec une probabilité de 0.7 et ni l'un ni l'autre avec une probabilité de 0.25.
1) Quelle est la probabilité qu'il soit fou ou voleur mais pas les deux ?
2) Quelle est la probabilité conditionnelle qu'il soit un voleur sachant qu'il n'est pas fou ?

Exercice 5
Un jeu commence par le choix d'un des deux dés A et B tel que A est choisi avec une probabilité p. Le dé choisi est ensuite lancé jusqu'à ce qu'une face blanche apparaisse, auquel cas le jeu est terminé.
Le dé A a 4 faces rouges et 2 faces blanches, le dé B a 2 faces rouges et 4 faces blanches.
Après que le jeu ai été joué longtemps, on observe que la probabilité que le jeu s'achève en exactement trois lancers est 7/81. Déterminer p.

Exercice 6

La maison de la famille Durand abrite le père, la mère, 4 enfants, 2 chats et 3 chiens. Toutes les six heures, la maisonnée forme un groupe de 5 (personnes et animaux) avec les règles suivantes :

- Il y a au moins un parent et un animal dans chaque groupe.
- Il n'y a jamais un chien et un chat dans le même groupe à moins que les deux parents n'y soient.
- Tous les groupes sont équiprobables.

Sachant qu'exactement un des parents se trouve dans le groupe de 18 heures, quelle est la probabilité que Rover, le chien le plus âgé, y soit ?

2. Les variables aléatoires et lois de probabilités

Définition

Une variable aléatoire (VA) est définie par une application de Ω (ensemble fondamental) vers R :

$$X : \Omega \to R$$
$$w \mapsto X(w)$$

$X(\Omega) = \{y \in R \,/\, \exists\ w \in \Omega, X(w) = y\}$

Si $X(\Omega)$ est un sous ensemble finie de R alors X est une variable aléatoire réelle discrète finie.
Si $X(\Omega)$ est un sous ensemble dénombrable de R, alors X est une Variable aléatoire discrète infinie
Si $X(\Omega)$ est un intervalle de R alors X est une Variable aléatoire continue.
$X(\Omega)$ est appelé support de la variable aléatoire X.

Notation :
X désigne une VA (application de Ω vers R)
x désigne un réel qui appartient à $X(\Omega)$

Fonction de répartition

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

$$F_X : R \to R$$
$$x \mapsto P(X \le x)$$

Propriétés

(1) $\forall\ x \in R,\ 0 \le F_X(x) \le 1$
(2) F_X est croissante sur R
(3) $\lim\limits_{x\to-\infty} F_X(x) = 0$ et $\lim\limits_{x\to+\infty} F_X(x) = 1$ et F_X est continue à droite de tout point de R
(4) Si $a \le b$, $P(a < x \le b) = F_X(b) - F_X(a)$

En effet :

(1) $F_X(x) = P(X \le x)$ donc $0 \le F_X(x) \le 1$

(2) Si $a \le b$ alors $\{X \le a\} \subset \{X \le b\}$
$\{X \le b\} = \{X \le a\} \cup \{a < X \le b\}$: deux événements exclusifs ou disjoints
donc $F_X(a) \le F_X(b)$

(4) $F_X(b) = F_X(a) + P(a < x \le b)$

(3) F_X est une fonction croissante sur R et bornée sur [0, 1], donc elle admet une limite au voisinage de $-\infty$ et au voisinage de $+\infty$
$A_n = (X \le n),\ n \in N$

$$\bigcup_{n\in N} A_n = (X \in R) = \Omega \Rightarrow \lim_{n\to+\infty} \mathrm{P}(\mathrm{A_n}) = 1 \Rightarrow \lim_{n\to+\infty} \mathrm{F(n)} = 1$$

Vu l'existence de la limite en + ∞, alors $\lim_{x\to+\infty} \mathrm{F_X(x)} = 1$

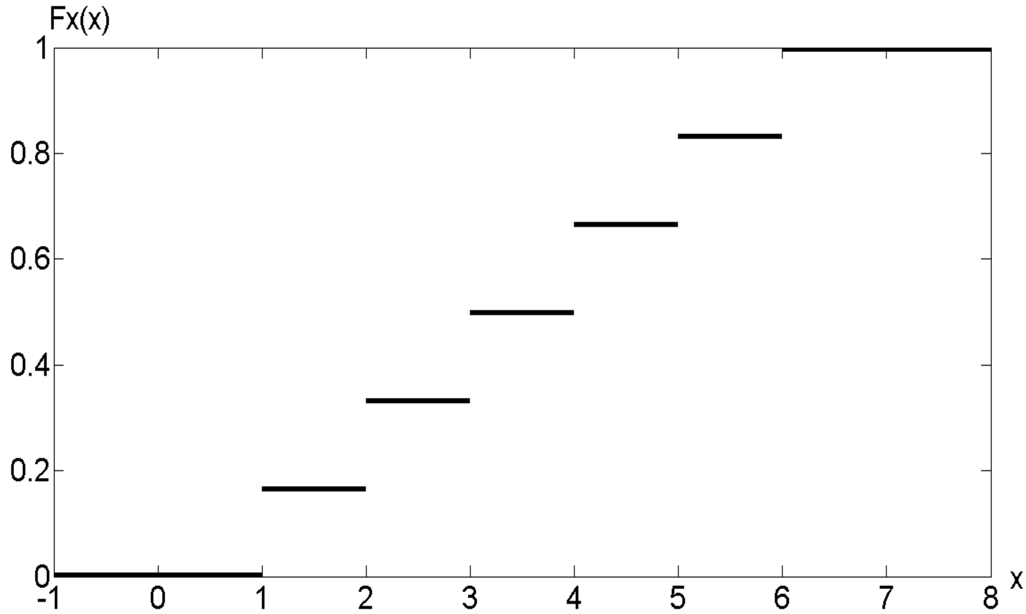
Un raisonnement similaire peut être fait au voisinage de -∞, en considérant la suite décroissante $B_n = (X \le -n)$, $n \in N$, pour montrer que $\lim_{x\to-\infty} \mathrm{F_X(x)} = 0$

Soit $C_n = (X \le x_0 + 1/n)$, $(C_n)_{n\in N}$ est une suite décroissante d'événements et $\bigcap_{n\in N} C_n = (X \le x_0)$

Donc $\lim_{n\to+\infty} P(C_n) = P(X \le x_0)$, c'est à dire $\lim_{n\to+\infty} F(x_0 + 1/n) = F(x_0)$, la continuité de F_X à droite de tout x_0 de R est ainsi déduite.

Exemple

Considérons le cas d'une variable aléatoire X associée à l'expérience aléatoire qui consiste à jeter un dé parfait numéroté de 1 à 6 et relever le numéro obtenu.
On à $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$ et $P(X = x_i) = 1/6$.



Lois discrètes

Soit X une VA discrète sur Ω. L'ensemble des valeurs prises par X est $X(\Omega) = \{x_1, x_2, \ldots, x_n, \ldots\}$. A chaque x_n, on peut associer $P(X = x_n)$ et on a : $\sum_n P(X = x_n) = P(\Omega) = 1$

La loi discrète est caractérisée par l'ensemble X(Ω) et par la donnée de la probabilité P(X= xn) associée à chaque valeur X.

Si X(Ω) est fini, il est possible de représenter la loi discrète X sous forme d'un tableau :

X	x_1	x_2	…	x_j	…	x_n
P(X)	p_1	p_2	…	p_j	…	p_n

La première ligne du tableau indique les différentes valeurs possibles de la variable aléatoire X et la deuxième ligne présente les probabilités : $P(X = x_j)$, j = 1, …, n.

Exemples de lois discrètes finies

Loi de Bernoulli

C'est une expérience aléatoire qui admet deux résultats possibles :
$P(A) = p$ et $P(\overline{A}) = 1 - p$, La variable aléatoire associée admet deux valeurs possibles : {0, 1}

X	0	1
P(X)	1 - p	p

Exemples :

(1) Jet d'une pièce de monnaie équilibrée, a chaque jet il y a deux résultats possibles :
P(pile) = 1/2 et P(face) = 1/2

(2) Une urne contient b boules blanches et n boules noires, on choisit au hasard une boule qui ne peut être que blanche ou noire :
P(boule blanche) = b/(n+b), P(boule noire) = n/(n+b)

(3) Une pièce fabriquée par une machine, deux résultats possibles : pièce acceptable ou pièce défectueuse.

La loi de Bernoulli est notée par B(p).

Loi binomiale

Soit E une expérience aléatoire liée à une loi de Bernoulli. E admet deux résultats possibles R_1 et R_2 : $P(R_1) = p$ et $P(R_2) = 1-p$

On répète E **n fois** dans des conditions identiques (p est constant) et indépendantes (faire E la ième fois est indépendant avec faire E la jème fois).

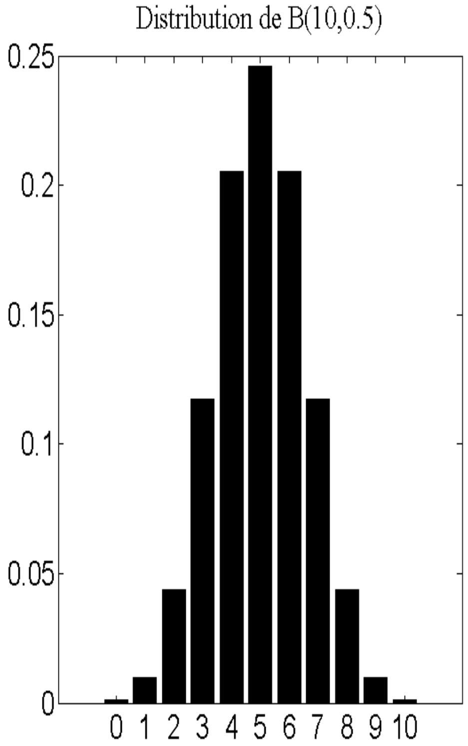
Soit X la VA qui désigne le nombre de fois où le résultat R_1 est obtenu, $X(\Omega) = \{0, 1, …., n\}$
$X = k \equiv k$ fois R_1 et (n-k) fois R_2

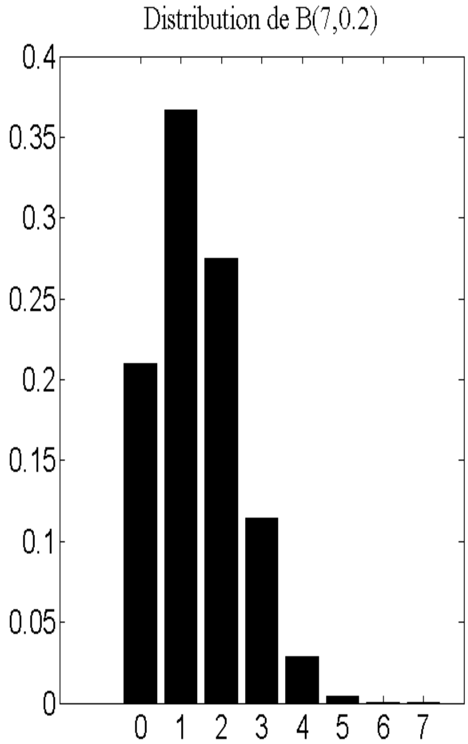
Le nombre de fois où cette configuration est obtenue est égale au nombre de permutations avec répétition d'ordre n où R_1 figure k fois et R_2 figure (n-k) fois :

$$P_{k,n-k} = \frac{n!}{k!(n-k)!} = C_n^k$$

$$P(X = k) = P_{k,n-k}\, p^k (1-p)^{n-k} = C_n^k p^k (1-p)^{n-k}$$

On dit que X suit une loi binomiale de paramètre n et p et on note B(n,p).





On vérifie facilement que la somme des probabilités est égale à 1, en utilisant la formule du binôme :

$$\sum_{k=0}^{n} P(X=k) = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = (p+(1-p))^n = 1$$

Loi uniforme discrète

La variable aléatoire associée à une loi uniforme discrète admet n valeurs 1, …., n et pour lesquelles on $P(X = i) = 1/n$, $i = 1,\ldots,n$

Exemple : jet d'un dé équilibré à 6 faces.

X	1	2	3	4	5	6
P(X)	1/6	1/6	1/6	1/6	1/6	1/6

Loi hypergéométrique

Soit une urne contenant N boules, dont n_1 sont blanches et $N - n_1 = n_2$ sont noires. On tire successivement et sans remise n boules, soit X la variable aléatoire qui désigne le nombre de boules blanches obtenues.

Parmi les n boules prélevées, il y a k boules blanches et (n-k) boules noires

Conditions sur k : $0 \le k \le n_1$ et $0 \le n-k \le n_2$

$\sup(0, n-n_2) \le k \le \inf(n, n_1)$

Calculons la probabilité de la configuration B,…,B,N,…,N : k boules blanches et n-k boules noires.
Les tirages sont dépendants, on utilise le Théorème de multiplication :
P(B,…,B,N,…,N) = P(B)P(B/B)…P(N/B,…,B)….P(N/B,…,B,N,…,N)

$$= \frac{n_1}{N}\frac{n_1-1}{N-1}\cdots\frac{n_1-(k-1)}{N-(k-1)}\frac{n_2}{N-k}\frac{n_2-1}{N-(k+1)}\cdots\frac{n_2-(n-k-1)}{N-(n-1)}{}_2$$

$$= \frac{n_1!}{(n_1-k)!}\frac{n_2!}{(n_2-(n-k))!}\frac{(N-n)!}{N!}$$

Or le nombre de configurations B,…,B,N,…,N est le nombre de permutations avec répétition d'ordre n où la boule blanche figure k fois et la boule noire figure (n-k) fois : $\frac{n!}{k!(n-k)!}$

$$P(X=k) = \frac{n_1!}{(n_1-k)!}\frac{n_2!}{(n_2-(n-k))!}\frac{(N-n)!}{N!}\frac{n!}{k!(n-k)!}$$

$$= \frac{n_1!}{k!(n_1-k)!}\frac{n_2!}{(n-k)!(n_2-(n-k))!}\frac{n!(N-n)!}{N!}$$

$$= \frac{C_{n_1}^{k}C_{n_2}^{n-k}}{C_N^n}$$

Il est possible de vérifier que $\sum_{k\in\Theta} P(X=k) = \sum_{k\in\Theta}\frac{C_{n_1}^{k}C_{n_2}^{n-k}}{C_N^n} = 1,\ \Theta = \{\sup(0, n - n_2),...,\inf(n, n_1)\}$

Pour cela, on utilise : $(1+x)^N = (1+x)^{n1}(1+x)^{n2}$

$$\sum_{n=0}^{N} C_N^n x^n = \sum_{j=0}^{n_1} C_{n_1}^{j} x^j \sum_{k=0}^{n_2} C_{n_2}^{k} x^k$$

Ce qui entraine que : $C_N^n x^n = \sum_{j+k=n} C_{n_1}^{j} C_{n_2}^{k} x^n$

D'où l'égalité : $C_N^n = \sum_{j+k=n} C_{n_1}^{j} C_{n_2}^{k}$, avec $\left|\begin{array}{l} 0 \le j \le n_1 \\ 0 \le k \le n_2 \end{array}\right.$

$$C_N^n = \sum_{k\in\Theta} C_{n_1}^{k} C_{n_2}^{n-k}$$

La loi hypergéométrique est caractérisée par trois paramètres : N, p et n, avec $p = n_1/N$.
Elle est notée par $\mathcal{H}$(N,n,p).

Si q = 1 – p alors $P(X=k) = \frac{C_{Np}^{k}C_{Nq}^{n-k}}{C_N^n}$

Lorsque N est assez grand et n est petit devant N, la loi hypergéométrique de paramètres N, p et n converge vers la loi binomiale B(n,p), en effet :

$$P(X=k) = \frac{Np!}{(Np-k)!k!}\frac{Nq!}{(Nq-n+k)!(n-k)!}\frac{(N-n)!n!}{N!}$$

$$\Rightarrow P(X=k) = C_n^k \frac{Np!}{(Np-k)!}\frac{Nq!}{(Nq-n+k)!}\frac{(N-n)!}{N!}$$

$$\frac{Np!}{(Np-k)!} = Np(Np-1)...(Np-k+1) \xrightarrow[N\text{ grand}]{} (Np)^k$$

$$\frac{Nq!}{(Nq-n+k)!} = Nq(Nq-1)...(Nq-n+k+1) \xrightarrow[N\text{ grand}]{} (Nq)^{n-k}$$

$$\frac{(N-n)!}{N!} = \frac{1}{N(N-1)...(N-n+1)} \xrightarrow[N\text{ grand}]{} \frac{1}{\mathrm{N^n}}$$

Donc $P(X=k) \xrightarrow[N\text{ grand}]{} \mathrm{C_n^k}\dfrac{(Np)^k(Nq)^{n-k}}{N^n} = \mathrm{C_n^k}p^k q^{n-k}$, c'est l'expression de la loi B(n,p).

En pratique, la loi $\mathcal{H}$(N,n,p) peut être approchée par la loi B(n,p) lorsque N $\geq$ 10n.

Lois discrètes infinies

Loi géométrique de paramètre p : elle décrit la probabilité d'obtenir le premier succès lorsqu'on exécute des épreuves indépendantes, chacune d'entre elle ayant la probabilité de succès p :

$\mathrm{p(n) = (1\text{-}p)^{n\text{-}1}p}$ pour $\mathrm{n} \in \{1, 2, \ldots\}$

$\sum\limits_{n=1}^{+\infty} p(n) = \sum\limits_{n=1}^{+\infty}(1-p)^{n-1}p = \dfrac{1}{1-(1-p)} = 1$, c'est la somme d'une suite géométrique.

Loi binomiale négative de paramètres p et r

Une expérience aléatoire dont la probabilité de succès est p. Cette expérience est répétée dans des conditions identiques et indépendantes jusqu'à ce qu'on obtienne r succès.
Soit n le nombre d'expériences réalisées :
$\mathrm{p(n) = C_{n-1}^{r-1}p^r(1-p)^{n-r}}$, $\mathrm{n} \in \{\mathrm{r, r+1}, \ldots\}$.

La justification de cette expression vient du fait qu' au cours de n réalisations on a r succès et le dernier résultat est un succès, cela suppose qu'au cours des n-1 réalisations, on r-1 succès.

Il est possible de montrer que : $\sum\limits_{\mathrm{n=r}}^{+\infty}\mathrm{p(n)} = \sum\limits_{\mathrm{n=r}}^{+\infty}\mathrm{C_{n-1}^{r-1}p^r(1-p)^{n-r}} = 1$

La démonstration peut se faire par récurrence :

Si r = 1, on a bien $\sum\limits_{\mathrm{n=1}}^{+\infty}\mathrm{C_{n-1}^{0}p(1-p)^{n-1}} = \dfrac{\mathrm{p}}{1-\mathrm{p}}\sum\limits_{\mathrm{n=1}}^{+\infty}\mathrm{(1-p)^n} = 1$

Supposons que : $\sum\limits_{\mathrm{n=r}}^{+\infty}\mathrm{C_{n-1}^{r-1}(1-p)^{n-r}} = \dfrac{1}{\mathrm{p^r}}$

Dérivons les deux égalités ci-dessus par rapport à p : $\sum\limits_{\mathrm{n=r}}^{+\infty}\mathrm{-(n-r)C_{n-1}^{r-1}(1-p)^{n-r-1}} = \dfrac{\mathrm{-r}}{\mathrm{p^{r+1}}}$

$$\Rightarrow \sum_{\mathrm{n=r}}^{+\infty}\frac{\mathrm{n-r}}{\mathrm{r}}\mathrm{C_{n-1}^{r-1}(1-p)^{n-r-1}} = \sum_{\mathrm{n=r+1}}^{+\infty}\frac{\mathrm{n-r}}{\mathrm{r}}\mathrm{C_{n-1}^{r-1}(1-p)^{n-r-1}} = \frac{1}{\mathrm{p^{r+1}}}$$

Or $\Rightarrow \dfrac{\mathrm{n-r}}{\mathrm{r}}\mathrm{C_{n-1}^{r-1}} = \dfrac{\mathrm{n-r}}{\mathrm{r}}\dfrac{\mathrm{(n-1)!}}{\mathrm{(r-1)!(n-r)!}} = \dfrac{\mathrm{(n-1)!}}{\mathrm{r!(n-r-1)!}} = \mathrm{C_{n-1}^{r}}$

Donc $\sum_{n=r+1}^{+\infty} C_{n-1}^{r} p^{r+1}(1-p)^{n-r-1} = 1$

Loi de Poisson

C'est une loi discrète infinie qui dépend d'un paramètre λ strictement positif et est notée par P(λ), cette loi a pour expression : $P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$

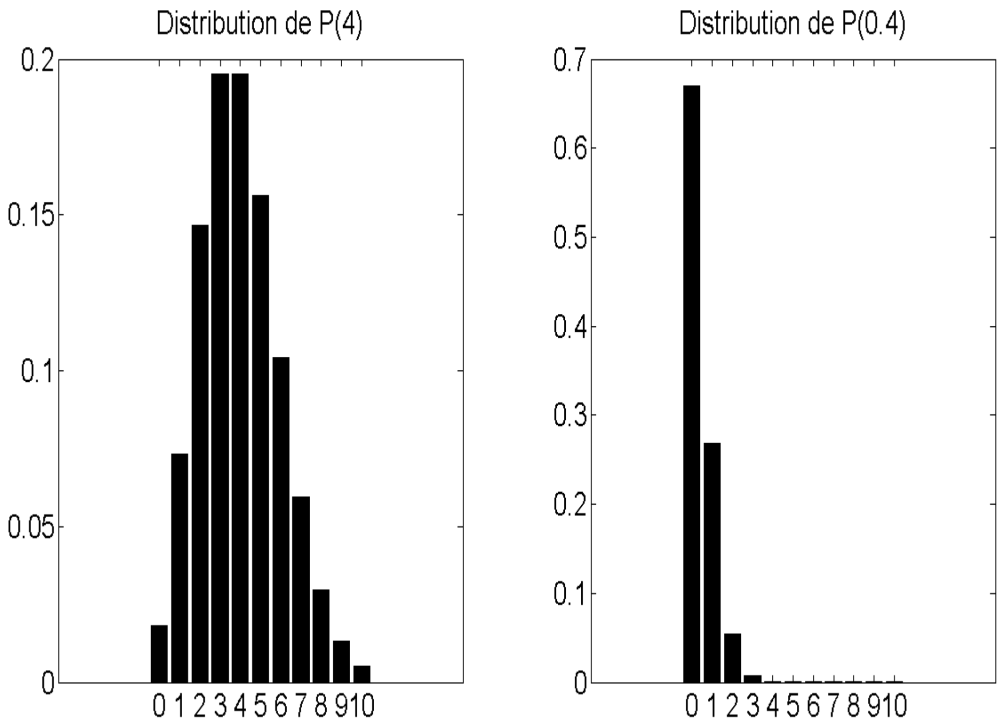
Le développement en série entière de la fonction exponentielle : $e^{\lambda} = \sum_{k=0}^{+\infty}\dfrac{\lambda^k}{k!}$ permet de démontrer facilement que $\sum_{k=0}^{+\infty} P(X=k) = 1$

On peut déterminer quel est l'entier le plus probable en formant le rapport :

$$\frac{P(X=k)}{P(X=k-1)} = \frac{\lambda^k}{k!}\frac{(k-1)!}{\lambda^{k-1}} = \frac{\lambda}{k}, k \geq 1$$

Ce qui montre que pour tous les entiers inférieurs à λ, on a P(X = k) > P(X = k-1), P(X = k) est croissante, puis décroissante pour les entiers k > λ, le maximum est atteint pour l'entier k = [λ]. Pour les cas où [λ] = 0 (λ < 1), les valeurs de P(X = k) sont décroissantes et la valeur maximale est P(X = 0). Si λ est un entier, il y a deux probabilités maximales qui sont égales : P(X = λ) et P(X = λ-1).

Les figures ci-dessous présentent les distributions de P(4) et P(0.4) sur l'ensemble {0, …10}.



La loi de Poisson est utilisée pour décrire le nombre de réalisations d'un événement dont la probabilité est faible :

- Le nombre d'accidents se produisant dans une usine au cours d'une période de temps donnée.
- Le nombre d'articles défectueux compris dans un envoi de bonne qualité.
- Le nombre d'appels téléphoniques pendant un intervalle de temps donné.

Théorème : Approximation d'une loi binomiale par une loi de Poisson
Soit X une variable aléatoire obéissant à une loi binomiale B(n,p) et supposons que p varie avec n : p = λ/n où λ est une constante positive, alors :

$$P(X=k) \xrightarrow[\mathrm{n}\to+\infty]{} \frac{\lambda^k e^{-\lambda}}{k!} \quad (\mathrm{k}=0,1,2,...)$$

Pour démontrer ce résultat on a :

$$P(X=k) = \mathrm{C}_\mathrm{n}^\mathrm{k} (\frac{\lambda}{n})^k (1-\frac{\lambda}{n})^{n-k}$$

Si n → +∞, on constate que : $\mathrm{C}_\mathrm{n}^\mathrm{k}\left(\frac{\lambda}{n}\right)^k = \frac{n(n-1)...(n-k+1)}{n^k}\frac{\lambda^k}{k!} \xrightarrow[n\to+\infty]{} \frac{\lambda^k}{k!}$

$$(1-\frac{\lambda}{n})^{n-k} = (1-\frac{\lambda}{n})^n (1-\frac{\lambda}{n})^{-k}$$

$$(1-\frac{\lambda}{n})^{-k} \xrightarrow[n\to+\infty]{} 1 \text{ et } \log((1-\frac{\lambda}{n})^n) = n\log(1-\frac{\lambda}{n}) \xrightarrow[n\to+\infty]{} -n\frac{\lambda}{n} = -\lambda$$

Donc $(1-\frac{\lambda}{n})^n \xrightarrow[n\to+\infty]{} e^{-\lambda}$

Ce qui prouve le théorème.

En pratique cette approximation est considérée comme acceptable si $\mathrm{n} \geq 30$, $\mathrm{p} \leq 0.1$ et $\mathrm{np} \leq 15$.

Loi de probabilité d'une fonction de variable aléatoire

Soit X une variable aléatoire discrète, g une fonction définie sur X(Ω), Calculons la loi de probabilité de Y = g(X) :

$$\mathrm{g} : \mathrm{X}(\Omega) \to \mathrm{R}$$
$$\mathrm{x_i} \mapsto \mathrm{g}(\mathrm{x_i})$$

Le caractère aléatoire de X est transmis à Y à travers la loi de X
Le support de Y ou l'ensemble des valeurs prises par la variable Y est :
$Y(\Omega) = \{y_j \ / \ \exists \ x_i \in X(\Omega), g(x_i) = y_j \}$
Soit $Ij = \{x_i \in X(\Omega) \ / \ g(x_i) = y_j\}$
La loi de Y :
$P(Y = y_j) = P(g(X) = y_j) = \sum_{x_i \in Ij} P(X = x_i)$

Exemple :

Loi de X :

x_i	-1	1	2
$P(X=x_i)$	1/4	1/2	1/4

Loi de $Y = e^X$

y_j	e^{-1}	e	e^2
$P(Y= y_j)$	1/4	1/2	1/4

Loi de $Z = X^2$
$Z = 1 \equiv X = 1$ ou $X = -1$, donc $P(Z=1) = P(X=-1) + P(X=1)$

z_k	1	4
$P(Z= z_k)$	3/4	1/4

Variables aléatoires continues à densité

Soit X une variable aléatoire, dans le cas où X(Ω) est un intervalle de R alors X est dite variable aléatoire continue.

On suppose qu'il existe une fonction f tel que $\forall a,b \in R, P(X \in [a,b]) = \int_a^b f(x)dx$

La fonction f est appelée alors densité de probabilité ou densité de la variable aléatoire continue.

Pour qu'une fonction f soit une densité de probabilité, il faut et il suffit donc qu'elle satisfasse aux deux conditions suivantes :

$$f(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Remarques

i) La probabilité d'un événement de type P(X = a) = 0 :

$$P(a < X < b) = \int_a^b f(x)dx \Rightarrow P(X = a) = \int_a^a f(x)dx = 0$$

$$P(X \leq a) = P(X < a) = \int_{-\infty}^a f(x)dx$$

ii) Pour chaque x de X(Ω) ou la fonction densité f_X est continue, alors la fonction de répartition F_X est dérivable et on a $F_X'(x) = f_X(x)$

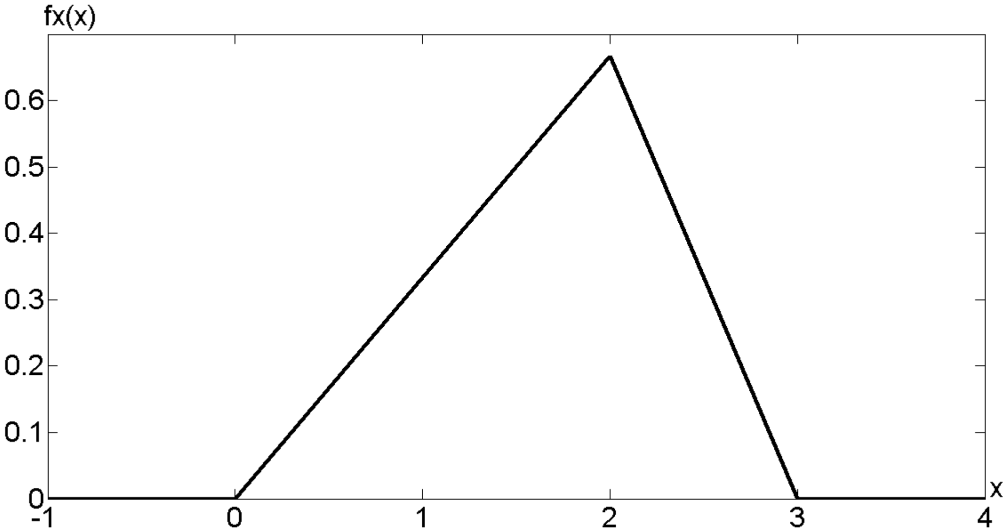
Exemple : on suppose que la longueur X d'une pièce fabriquée en série est distribuée selon une densité de probabilité donnée par :

$$f_X(x) = \begin{cases} x/3 & \text{si } 0 \leq x \leq 2 \\ 2 - 2x/3 & \text{si } 2 \leq x \leq 3 \\ 0 \text{ ailleurs} & \end{cases}$$

Vérifier que f_X est une densité de probabilité et représenter son graphe.

Calculer $P(1 \le X \le 2{,}5)$ ainsi que $P(1 \le X \le 2{,}5 / X \le 2)$

f_X est une fonction positive et on a $\int_0^3 f_X(x)dx = 1$



$$P(1 \le X \le 2{,}5) = \int_1^2 (x/3)dx + \int_2^{2,5} (2 - 2x/3)dx = 3/4$$

$$P(1 \le X \le 2{,}5/X \le 2) = \frac{P(1 \le X \le 2{,}5 \text{ et } X \le 2)}{P(X \le 2)} = \frac{\int_1^2 (x/3)dx}{\int_0^2 (x/3)dx} = 3/4$$

Densité d'une fonction de variable aléatoire

Soit X une variable aléatoire continue de densité f_X et soit g une fonction strictement monotone (décroissante ou croissante) et dérivable. Calculons la densité de la variable aléatoire $Y = g(X)$.

F_X et F_Y sont respectivement les fonctions de répartition de X et Y.

$F_Y(y) = P(Y<y) = P(g(X) < y)$ avec $y \in Y(\Omega) = \{ y \in R \,/\, \exists \; x \in X(\Omega), g(x) = y\}$

Si g est croissante alors $g(X) < y \equiv X < g^{-1}(y)$

donc $F_Y(y) = F_X(g^{-1}(y))$ et pour trouver la relation entre les fonctions densités, il suffit de dériver les fonctions de répartition : $f_Y(y) = (g^{-1})'(y)\, f_X(g^{-1}(y))$

Si g est décroissante alors $g(X) < y \equiv X > g^{-1}(y)$

donc $F_Y(y) = 1 - F_X(g^{-1}(y))$

$f_Y(y) = -(g^{-1})'(y)\, f_X(g^{-1}(y))$

D'où $f_Y(y) = \left| (g^{-1})'(y) \right| f_X(g^{-1}(y))$

Cas d'une fonction non bijective
Soit X une variable aléatoire continue de densité f_X, calculons la densité de la variable aléatoire $Y = X^2$.

$F_Y(y) = P(Y<y) = P(X^2<y)$
Si $y<0$ alors $P(X^2<y) = 0 \rightarrow F_Y(y) = 0$
Si $y>0$, $F_Y(y) = P(X^2<y)$, $(X^2 < y) \Leftrightarrow (-\sqrt{y} < X < \sqrt{y})$
$F_Y(y) = P(-\sqrt{y} < X < \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$

$$F_Y(y) = \begin{cases} 0 \text{ si } y < 0 \\ F_X(\sqrt{y}) - F_X(-\sqrt{y}) \text{ si y} \geq 0 \end{cases}$$

$$f_Y(y) = F_Y'(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}}$$

Variable aléatoires continues dont la densité est symétrique
Soit X une variable aléatoire continue de densité f_X symétrique et de fonction de répartition F_X, alors la loi de $-X$ est identique à celle de X et $\forall$ x $\in$ X(Ω), $F_X(-x) = 1 - F_X(x)$.

En effet : soit x $\in$ X(Ω), $P(-X<x) = P(X>-x)$ donc $F_{-X}(x) = 1 - F_X(-x)$
La dérivée des fonction de répartition implique : $f_{-X}(x) = f_X(-x) = f_X(x)$ car la fonction f_X est symétrique.

Soit x $\in$ X(Ω), $F_X(-x) = \int_{-\infty}^{-x} f_X(t)dt$, posons u = -t

$$F_X(-x) = -\int_{+\infty}^{x} f_X(-u)du = \int_{x}^{+\infty} f_X(u)du = 1 - F_X(x)$$

Exemples de variables aléatoires à densités

<u>Loi uniforme sur [a, b], U[a,b]</u>

La loi uniforme est caractérisée par une fonction densité constante sur l'intervalle de définition : [a, b], b > a

Cette constante est calculée de telle manière que : $\int_a^b f_U(u)du = cte\int_a^b du = cte(b-a) = 1$

$$f_U(u) = \begin{cases} \dfrac{1}{b-a} \text{ si } u \in [a,b] \\ 0 \text{ ailleurs} \end{cases}$$

$$F_U(u) = \begin{cases} 0 \text{ si u} < \text{a} \\ \dfrac{\text{u - a}}{\text{b - a}} \text{ si a} \leq \text{u} \leq \text{b} \\ 1 \text{ si u} > 1 \end{cases}$$

Loi normale N(μ,σ)

Cette loi constitue un modèle fréquemment utilisé dans divers domaines : variation du diamètre d'une pièce dans une fabrication industrielle, répartition des erreurs de mesure autour de la vraie valeur. La densité de cette loi est :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ pour } x \in R$$

La fonction de répartition est l'intégral de cette densité.
La loi normale centrée réduite N(0,1) est caractérisée par une fonction de répartition égale :

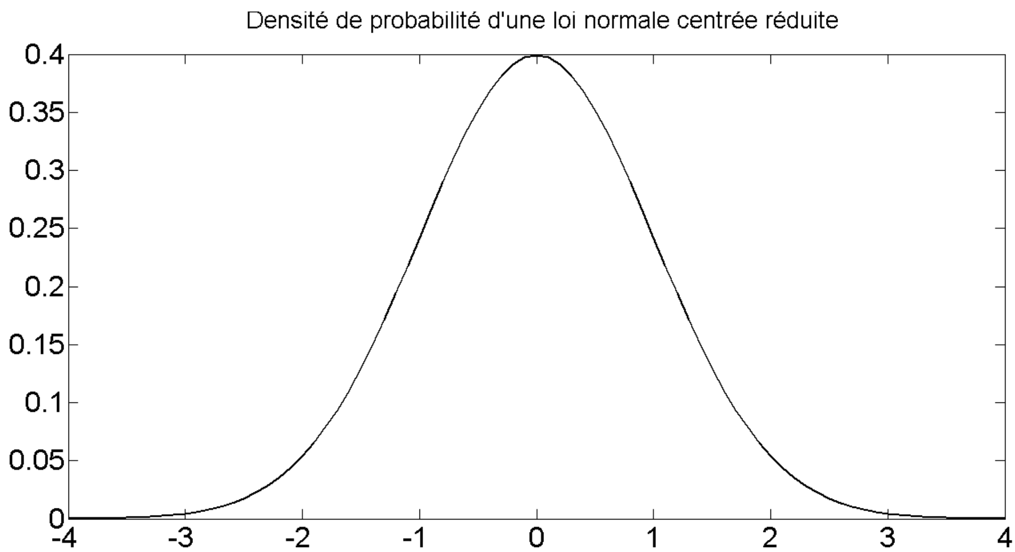
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} dx \text{ pour } x \in R$$

Théorème :
Si X est une variable aléatoire normale N(μ,σ), alors la variable aléatoire continue Y = αX + β est normale, de loi N(αμ+β, ασ) avec α ≠ 0.

En effet : La fonction g(X) = αX + β est strictement monotone, et $g^{-1}(y) = \frac{y-\beta}{\alpha}$
On utilise la relation : $f_Y(y) = \left| (g^{-1})'(y) \right| f_X(g^{-1}(y))$

$$f_Y(y) = \frac{1}{|\alpha|\sqrt{2\pi}\sigma} e^{-\frac{(y-\beta-\alpha\mu)^2}{2\sigma^2\alpha^2}} \Rightarrow \text{Y :N}(\alpha\mu + \beta, \alpha\sigma)$$

Application : $\alpha = \frac{1}{\sigma}$ et $\beta = -\frac{\mu}{\sigma}$
Dans ce cas αμ + β = 0 et ασ = 1
Toute loi normale N(μ, σ) peut être transformée en une loi normale centrée réduite.



La fonction densité de probabilité d'une loi normale centrée réduite est symétrique, par suite $\forall$ x $\in$ R alors : $\Phi(-x) = 1 - \Phi(x)$

Il existe des tables pour calculer les valeurs de la fonction Φ, fonction de répartition d'une loi normale centrée réduite.

Parmi les propriétés de la loi normale centrée réduite N(0, 1) :

P(X $\in$ [-1, 1])	0,68
P(X $\in$ [-2, 2])	0,95
P(X $\in$ [-3, 3])	0,99

Loi exponentielle

Définition :

Soit λ un réel strictement positif.
Une variable aléatoire X suit la loi exponentielle de paramètre λ (notée $\mathcal{E}(\lambda)$) lorsque X admet pour densité de probabilité la fonction f définie par $f(x) = \begin{cases} \lambda e^{-\lambda x} \text{ si } x \geq 0 \\ 0 \text{ si } x < 0 \end{cases}$

En effet f est une fonction positive et $\int_{-\infty}^{+\infty} f(x)dx = 1$

La fonction de répartition de la loi exponentielle est donnée par :

$$F(x) = \begin{cases} 1 - e^{-\lambda x} \text{ si x} \geq 0 \\ 0 \text{ si x} < 0 \end{cases}$$

La loi exponentielle à la propriété d'être sans mémoire :

P(X>x+u/X>u) = P(X>x)

$$P(X > x + u / X > u) = \frac{P(X > x + u \text{ et } X > u)}{P(X > u)} = \frac{P(X > x + u)}{P(X > u)} = \frac{1 - F(x + u)}{1 - F(u)}$$

Donc $P(X > x + u / X > u) = \dfrac{e^{-\lambda(x+u)}}{e^{-\lambda u}} = e^{-\lambda x} = P(X > x)$

Interprétation : Si X modélise la durée de vie d'un individu A, la propriété que X est sans mémoire exprime que A ne vieillit pas : si A a vécu u années, la probabilité pour qu'il vive encore x années est la même que la probabilité pour un individu similaire à A, qui vient de naitre, vive aussi x années

La loi Gamma

Définition :
Soit b et t deux nombres réels strictement positifs.
Une variable aléatoire X suit la loi gamma de paramètres b et t, lorsqu'elle admet pour densité la fonction f définie par $\begin{cases} f(x) = \dfrac{e^{-\frac{x}{b}} x^{t-1}}{\Gamma(t) b^t} \text{ si } x > 0 \\ f(x) = 0 \text{ si } x \leq 0 \end{cases}$, avec $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$

f est une fonction densité de probabilité car elle est positive et $\int_0^{+\infty} f(x)dx = 1$, en effet :

On pose x = bu et $\int_0^{+\infty} f(x)dx = \dfrac{1}{\Gamma(t)b^t} \int_0^{+\infty} e^{-u} b^{t-1} u^{t-1} b du = \dfrac{\Gamma(t)b^t}{\Gamma(t)b^t} = 1$
La fonction Γ vérifie la propriété : Γ(x+1) = xΓ(x) et si n est un entier alors Γ(n+1) = n!, avec Γ(1) = 1.
La loi gamma de paramètres b et 1, est une loi exponentielle de paramètre 1/b.

Distribution simultanée ou loi du couple

Couples aléatoires discrets

Soient X et Y deux VA discrètes définies sur l'ensemble fondamental Ω
$X(\Omega) = \{x_1, x_2, \ldots, x_i, \ldots\}$, $i \in I_X$
$Y(\Omega) = \{y_1, y_2, \ldots, y_j, \ldots\}$, $j \in I_Y$
La loi de probabilité jointe des deux variables aléatoires ou la distribution bidimensionnelle du vecteur aléatoire (X, Y) est alors définie par des probabilités pij affectées aux couples (x_i, y_j) :
$p_{ij} = P((X = x_i) \text{ et } (Y = y_j))$ c'est la loi du couple (X,Y)

L'ensemble des valeurs prises par (X, Y) sont les couples de X(Ω) × Y(Ω), c'est l'ensemble de tous les couples dont le premier élément appartient à X(Ω) et le deuxième appartient à Y(Ω).
$$\sum_{i \in I_X} \sum_{j \in I_Y} p_{ij} = \sum_{i \in I_X} \sum_{j \in I_Y} P(X = x_i \text{ et } Y = y_j) = \sum_{i \in I_X} P(X = x_i) = 1$$

La famille ($Y=y_j$), $j \in I_Y$, est collectivement exhaustive sur Y(Ω)
Il en est de même pour ($X=x_i$), $i \in I_X$

Les distributions marginales de X et de Y sont définies par :
$p_{i.} = P(X = x_i)$ et $p_{.j} = P(Y = y_j)$

Lorsque les variables aléatoires X et Y sont finies, le nombre de couples possibles est card(X(Ω))*card(Y(Ω)), il est alors possible de représenter le couple aléatoire (X,Y) par un tableau :

Y / X	y1	…	yj	…	ym	Probabilités marginales
x1						p1.
⋮						⋮
xi			pij			pi.
⋮						⋮
xn						pn.
Probabilités marginales	p.1	…	p.j	…	p.m	1

La fonction de répartition conjointe des deux V A X et Y, $F_{X,Y}(x,y)$ est définie par :

$$F_{X,Y}(x,y) = P(X \le x, Y \le y) = \sum_{x_i \le x} \sum_{y_j \le y} P(X = x_i, Y = y_j) = \sum_{x_i \le x} \sum_{y_j \le y} pij$$

Exemple

Considérons l'expérience stochastique qui consiste à jeter trois fois de suite, une pièce de monnaie équilibrée :

- X = nombre de ‘‘face’’; X(Ω) = {0, 1, 2, 3}
- Y = numéro du jet dans lequel apparaît ‘‘face’’ pour la première fois (on posera par exemple Y = 4 s'il y a trois ‘‘pile’’.

Y / X	1	2	3	4	pi.	
0	0	0	0	1/8	1/8	PPP
1	1/8	1/8	1/8	0	3/8	FPP, PFP, PPF
2	2/8	1/8	0	0	3/8	FFP, FPF, PFF
3	1/8	0	0	0	1/8	FFF
p.j	4/8	2/8	1/8	1/8	1	
	FPP FFP FFF	PFP PFF	PPF	PPP		

Couple aléatoire continue

X et Y sont deux variables aléatoires à densité sur respectivement X(Ω) et Y(Ω) deux sous ensembles de R.
Les valeurs prises par le couple (X, Y) décrivent une partie du plan de dimension 2
Les variables aléatoires X et Y sont conjointement continues s'il existe une fonction $f_{X,Y}$ appelée densité conjointe de (X, Y) telle que :

$$\forall C \in R^2, P((X,Y) \in C) = \iint_C f_{X,Y}(x,y)dxdy$$

$$\forall A, B \subset R, P((X,Y) \in A \times B) = \iint_{A\times B} f_{X,Y}(x,y)dxdy$$

La fonction densité conjointe vérifie les deux propriétés suivantes :

- $f_{X,Y}(x,y) \geq 0$ pour tout $(x, y) \in R^2$
- $\iint_{R^2} f_{X,Y}(x,y)dxdy = 1$

La fonction de répartition de (X,Y) est définie par :

$$F_{X,Y}(a,b) = P(X \leq a, Y \leq b) = \int_{-\infty}^{a}\int_{-\infty}^{b} f_{X,Y}(x,y)dydx$$

Ce qui permet d'exprimer la densité conjointe de (X, Y) par : $f_{X,Y}(x,y) = \dfrac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$

Les densités marginales de X et Y sont données par :

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy \text{ et } f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx$$

Exemple

Soit la densité conjointe du couple (X, Y) donnée par :

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y} \text{ si } x, y \in R^+ \\ 0 \text{ sinon} \end{cases}$$

Calculer : P(X>1,Y<1), P(X<a), $F_{X,Y}(x,y)$

La fonction f est positive et on vérifie que $\int_0^{+\infty}\int_0^{+\infty} f_{X,Y}(x,y)dxdy = 1$

En effet : $\int_0^{+\infty} 2e^{-x}\int_0^{+\infty} e^{-2y}dxdy = \int_0^{+\infty} 2e^{-x}[\frac{-e^{-2y}}{2}]_0^{+\infty} dx = \int_0^{+\infty} e^{-x}dx = 1$

P(X>1,Y<1) = $\int_1^{+\infty}\int_{-\infty}^{1} f_{X,Y}(x,y)dxdy = \int_1^{+\infty} e^{-x}\int_0^1 2e^{-2y}dydx = [-e^{-x}]_1^{+\infty}[-e^{-2y}]_0^1 = e^{-1}(1-e^{-2})$

Pour calculer P(X<a), il faut déterminer la densité marginale f_X :

$$f_X(x) = \begin{cases} \int_0^{+\infty} 2e^{-x}e^{-2y}dy = e^{-x}[-e^{-2y}]_0^{+\infty} = e^{-x} \text{ si } x \geq 0 \\ 0 \text{ sinon} \end{cases}$$

$$P(X < a) = \int_0^a f_X(x)dx = [-e^{-x}]_0^a = 1 - e^{-a}$$

$$F_{X,Y}(x,y) = \int_0^a \int_0^b f_{X,Y}(x,y)dydx = \int_0^a \int_0^b 2e^{-x}e^{-2y}dydx$$

$$= [-e^{-x}]_0^a[-e^{-2y}]_0^b = (1-e^{-a})(1-e^{-2b})$$

On peut vérifier que : $f_{X,Y}(a,b) = \dfrac{\partial^2 F_{X,Y}(a,b)}{\partial x \partial y}$

Indépendance de deux variables aléatoires

Soient X et Y deux variables aléatoires, la notion d'indépendance entre X et Y peut être exprimée au niveau des probabilités :
$\forall$ A, B $\subset$ R, P(X $\in$ A, Y $\in$B) = P(X $\in$A)P(Y $\in$B)
Cette notion d'indépendance peut être également exprimée au niveau des fonctions de répartition : $\forall$ (x,y) $\in$ R^2 $F_{X,Y}(x,y) = F_X(x)F_Y(y)$

Dans le cas discret, une condition nécessaire et suffisante pour l'indépendance de deux variables aléatoires X et Y est :
$\forall$ $x_i \in X(\Omega)$ et $\forall$ $y_j \in Y(\Omega)$, $P(X=x_i \text{ et } Y=y_j) = P(X=x_i)P(Y=y_j)$

Cette égalité des probabilités entraine celle des fonctions de répartition :

$$F_{X,Y}(x,y) = \sum_{X \leq x}\sum_{Y \leq y} P(X = x, Y = y) = \sum_{X \leq x}\sum_{Y \leq y} P(X = x)P(Y = y)$$

$$= \sum_{X \leq x} P(X = x) \sum_{Y \leq y} P(Y = y) = F_X(x)F_Y(y)$$

Dans l'exemple du couple discret (X,Y) précité, on a :
P(X = 0, Y = 1) = 0 $\neq$ P(X = 0)P(Y = 1) = (1/8)(4/8) = 1/16, les variables aléatoires X et Y sont dépendantes.

Dans le cas continu, une condition nécessaire et suffisante pour l'indépendance de deux variables aléatoires X et Y est : $f_{X,Y}(x,y) = f_X(x)f_Y(y)$.

Nous avons l'équivalence suivante : $F_{X,Y}(x,y) = F_X(x)F_Y(y) \Leftrightarrow f_{X,Y}(x,y) = f_X(x)f_Y(y)$.

La première implication peut être démontrée en dérivant deux fois les fonctions de répartition par rapport à x et par rapport à y.
L'implication inverse peut être démontrée en intégrant doublement les fonctions densité.

Dans l'exemple du couple continu (X,Y) précité, on a :

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y} \text{ si } x, y \in R^+ \\ 0 \text{ sinon} \end{cases}$$

$$f_X(x) = \begin{cases} \int_0^{+\infty} 2e^{-x}e^{-2y}dy = e^{-x}[-e^{-2y}]_0^{+\infty} = e^{-x} \text{ si } x \geq 0 \\ 0 \text{ sinon} \end{cases}$$

$$f_Y(y) = \begin{cases} \int_0^{+\infty} 2e^{-x}e^{-2y}dx = 2e^{-y}[-e^{-x}]_0^{+\infty} = 2e^{-2y} \text{ si } x \geq 0 \\ 0 \text{ sinon} \end{cases}$$

Les variables aléatoires X et Y sont indépendantes.

Lois conditionnelles

Cas discret

Soit X une V A discrète et A un événement de probabilité non nulle. La loi de X sachant A ou conditionnée par A est l'ensemble des couples ($x \in X(\Omega)$, $P(X = x/A)$).

X/A	x_1	x_2	…	x_j	…	x_n
P(X/A)	$P(X=x_1/A)$	$P(X=x_2/A)$	…	$P(X=x_j/A)$	…	$P(X=x_n/A)$

Cas particulier : Si Y est une V A discrète, la loi de X sachant Y = y est l'ensemble des couples ($x \in X(\Omega)$, $P(X = x/Y = y)$).

$P(X = x/Y = y) = \dfrac{P(X = x, Y = y)}{P(Y = y)}$, c'est le rapport de la probabilité conjointe et de la probabilité marginale.

Reprenons l'exemple du couple discret et calculons la loi conditionnelle X/ Y = 2 :

X/Y = 2	0	1	2	3
$P(X=x_i/Y=2)$	0	1/2	1/2	0

En effet : P(X = 0, Y = 2) = P(X=3, Y = 2) = 0 et P(Y = 2) = 2/8
P(X = 1, Y = 2) = P(X=2, Y = 2) = 1/8.

Il est également possible de définir une loi conjointe conditionnelle, si (X, Y) est un couple discret et A un événement de probabilité non nulle, on a :

$$p_{X,Y/A}(x, y) = \begin{cases} \dfrac{P(X = x, Y = y)}{P(A)} \text{ si } (x, y) \in A \\ 0 \text{ sinon} \end{cases}$$

Cas continu

Comme pour le cas discret, il est possible de définir une fonction densité conjointe conditionnelle, si X et Y sont deux V A à densité et A un événement de probabilité non nulle :

$$f_{X,Y/A}(x,y)=\begin{cases}\dfrac{f_{X,Y}(x,y)}{P(A)} \text{ si } (x,y)\in A\\ 0 \text{ sinon}\end{cases}$$

Il est possible également de déterminer la densité conditionnelle de X sachant Y = y à condition que $f_Y(y) \neq 0$: $f_{X/Y}(x/y)=\dfrac{f_{X,Y}(x,y)}{f_Y(y)}$

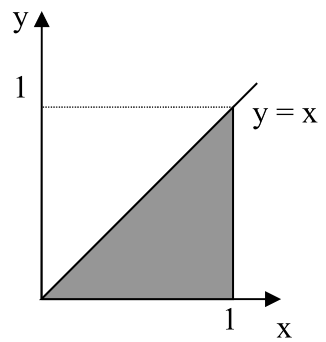
Exemple

Soit la densité conjointe de (X,Y) donnée par : $f_{X,Y}(x,y)=\begin{cases}3x \text{ si } 0\le y\le x\le 1\\ 0 \text{ sinon}\end{cases}$

Calculer la densité conditionnelle de Y sachant X = x et la densité marginale de Y.

Pour cela, il faut calculer d'abord la densité marginale de X :

Le domaine de définition D de la fonction $f_{X,Y}$ est



$$f_X(x)=\begin{cases}\int_0^x 3x\,dy = 3x^2 \text{ si } 0\le x\le 1\\ 0 \text{ sinon}\end{cases}$$

La densité conditionnelle de Y sachant X = x est :

$$f_{Y/X}(y/x)=\begin{cases}\dfrac{f_{X,Y}(x,y)}{f_X(x)}=\dfrac{1}{x} \text{ si } 0\le y\le x\le 1\\ 0\end{cases}$$

La densité marginale de Y est : $f_Y(y)=\begin{cases}\int_y^1 3x\,dx=[3\frac{x^2}{2}]_y^1 = \frac{3}{2}(1-y^2) \text{ si } 0\le y\le 1\\ 0 \text{ sinon}\end{cases}$

Les V A X et Y sont dépendantes car $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$.

Loi de la somme de deux variables aléatoires

<u>Cas discret</u>

X et Y sont deux V A discrètes et soit Z = X + Y, calculons la densité de Z.

P(Z=z) = $\sum_{x\in X(\Omega)} P(X=x, Y=z-x) = \sum_{y\in Y(\Omega)} P(X=z-y, Y=y)$

Si les V A X et Y sont indépendantes alors :
P(Z=z) = $\sum_{x \in X(\Omega)} P(X = x)P(Y = z - x) = \sum_{y \in Y(\Omega)} P(X = z - y)P(Y = y)$

Application : Calculer la loi de la somme de deux lois de Poisson indépendantes de paramètres λ et μ.

P(Z=z) = $\sum_{x=0}^{+\infty} \lambda^x \frac{e^{-\lambda}}{x!} \mu^{z-x} \frac{e^{-\mu}}{(z-x)!} = e^{-\lambda-\mu} \sum_{x=0}^{z} \frac{\lambda^x}{x!} \frac{\mu^{z-x}}{(z-x)!}$, car z-x = y ≥ 0
P(Z=z) = $\frac{e^{-\lambda-\mu}}{z!} \sum_{x=0}^{z} z! \frac{\lambda^x}{x!} \frac{\mu^{z-x}}{(z-x)!} = \frac{e^{-\lambda-\mu}}{z!} (\lambda + \mu)^z$

La somme de deux lois de Poisson indépendantes de paramètres λ et μ, est une loi de Poisson de paramètre λ + μ.
Ce résultat peut être généralisé à la somme de n lois de Poisson indépendantes, de paramètres λ_i, est une loi de Poisson de paramètre $\lambda_1 + \lambda_2 + \ldots + \lambda_n$.

Cas continu
X et Y sont deux V A à densité et soit Z = X + Y, calculons la densité de Z.
$F_Z(z) = P(Z<z) = P(X+Y < z) = \iint_{D_{x,y}} f_{X,Y}(x, y)dxdy$, avec $D_{x,y} = \{(x,y) \in \mathbb{R}^2 / x + y < z\}$

$F_Z(z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z-x} f_{X,Y}(x, y)dydx$

Faisons le changement de variables suivant : $\left| \begin{array}{l} u = x \\ v = x + y \end{array} \right. \Leftrightarrow \left| \begin{array}{l} x = u \\ y = v - u \end{array} \right.$

$(x,y) \in D_{x,y} \Leftrightarrow (u,v) \in D_{u,v}$ avec $D_{u,v} = \{(u,v) \in \mathbb{R}^2 / v < z\}$
$F_Z(z) = \iint_{D_{x,y}} f_{X,Y}(x, y)dxdy = \iint_{D_{u,v}} f_{X,Y}(x(u,v), y(u,v)) \left| \frac{D(x, y)}{D(u,v)} \right| dudv$

$$\frac{D(x, y)}{D(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

$F_Z(z) = \int_{-\infty}^{z} \int_{-\infty}^{+\infty} f_{X,Y}(u, v - u)dudv = \int_{-\infty}^{z} f_Z(v)dv$, ce qui permet d'obtenir l'expression de la densité de Z : $f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(u, z - u)du = \int_{-\infty}^{+\infty} f_{X,Y}(x, z - x)dx$, d'une manière analogue, il est possible de montrer que $f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(z - y, y)dy$
Si en plus les V A X et Y sont indépendantes alors la densité de Z = X +Y, a pour expression :
$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z - x)dx = \int_{-\infty}^{+\infty} f_X(z - y) f_Y(y)dy$.

Application : Calculer la loi de la somme de deux V A indépendantes, normales centrées réduites :

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \text{ et } f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

$$f_Z(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} e^{-\frac{(z-x)^2}{2}} dx = \frac{1}{\sqrt{2}\sqrt{2\pi}} e^{-\frac{z^2}{2*2}} \left[\frac{1}{\frac{1}{\sqrt{2}}\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-z/2)^2}{2*1/2}} d(x) \right]$$

Cette égalité est vérifiée car $-\frac{x^2}{2} - \frac{(z-x)^2}{2} = -\frac{z^2}{4} - (x - z/2)^2$

Le terme entre crochets vaut 1 car c'est l'intégrale sur R de la densité d'une loi N(z/2, $\frac{1}{\sqrt{2}}$) et la loi Z est une loi normale N(0, $\sqrt{2}$)

Loi du produit de deux variables aléatoires à densité

X et Y sont deux V A à densité et soit Z = XY, calculons la densité de Z.
$F_Z(z) = P(Z<z) = P(XY < z) = \iint_{D_{x,y}} f_{X,Y}(x,y)dxdy$, avec $D_{x,y} = \{(x,y) \in R^2 / xy < z\}$

Faisons le changement de variables suivant : $\left| \begin{array}{l} u = x \\ v = xy \end{array} \right. \Leftrightarrow \left| \begin{array}{l} x = u \\ y = \frac{v}{u} \end{array} \right.$

$(x,y) \in D_{x,y} \Leftrightarrow (u,v) \in D_{u,v}$ avec $D_{u,v} = \{(u,v) \in R^2 / v < z\}$

$F_Z(z) = \iint_{D_{x,y}} f_{X,Y}(x,y)dxdy = \iint_{D_{u,v}} f_{X,Y}(x(u,v), y(u,v)) \left| \frac{D(x,y)}{D(u,v)} \right| dudv$

$$\frac{D(x,y)}{D(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -\frac{v}{u^2} & \frac{1}{u} \end{vmatrix} = \frac{1}{u}$$

$F_Z(z) = \int_{-\infty}^{z} \int_{-\infty}^{+\infty} f_{X,Y}(u, \frac{v}{u}) \frac{dudv}{|u|} = \int_{-\infty}^{z} f_Z(v)dv$, ce qui permet d'obtenir l'expression de la densité de Z : $f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(u, \frac{z}{u}) \frac{du}{|u|} = \int_{-\infty}^{+\infty} f_{X,Y}(x, \frac{z}{x}) \frac{dx}{|x|}$, d'une manière analogue, il est possible de montrer que $f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(\frac{z}{y}, y) \frac{dy}{|y|}$

Si en plus les V A X et Y sont indépendantes alors la densité de Z = XY, a pour expression :
$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(\frac{z}{x}) \frac{dx}{|x|} = \int_{-\infty}^{+\infty} f_X(\frac{z}{y}) f_Y(y) \frac{dy}{|y|}$.

Application : Calculer la loi du produit de deux V A indépendantes, uniformes sur [-1 , 1] :

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{si } -1 \le x \le 1 \\ 0 & \text{sinon} \end{cases} \text{ et } f_Y(y) = \begin{cases} \frac{1}{2} & \text{si } -1 \le y \le 1 \\ 0 & \text{sinon} \end{cases}$$

$x \in [-1 , 1]$ et $y \in [-1 , 1]$ alors $z = xy \in [-1 , 1]$

$$f_Z(z)=\int_{-1}^{+1} f_X(x) f_Y(\frac{z}{x})\frac{dx}{|x|} \text{ et } f_Y(\frac{z}{x})\neq 0 \Leftrightarrow -1\leq \frac{z}{x}\leq 1 \Leftrightarrow |z|\leq |x|$$

Si $z \geq 0$ alors $f_Z(z)=\int_{-1}^{-z} -\frac{dx}{4x}+\int_{z}^{1}\frac{dx}{4x}=-\frac{1}{2}\log(z)$

Si $z \leq 0$ alors $f_Z(z)=\int_{-1}^{z} -\frac{dx}{4x}+\int_{-z}^{1}\frac{dx}{4x}=-\frac{1}{2}\log(-z)$

Finalement $f_Z(z)=\begin{cases} -\frac{1}{2}\log(|z|) \text{ si } -1\leq z\leq 1 \\ 0 \text{ sinon} \end{cases}$

Il est possible de vérifier que $\int_{-1}^{1} f_Z(z)dz=1$, en effet :

$$-\frac{1}{2}\int_{-1}^{1}\log(|z|dz=-\int_{0}^{1}\log(z)dz=-[z\log(z)]_0^1+\int_0^1 dz=1$$

Loi du rapport de deux variables aléatoires à densité

X et Y sont deux V A à densité et soit $Z=\frac{X}{Y}$, calculons la densité de Z.

$F_Z(z) = P(Z<z) = P(\frac{X}{Y}<z) = \iint_{D_{x,y}} f_{X,Y}(x,y)dxdy$, avec $D_{x,y} = \{(x,y) \in \mathbb{R}^2 / \frac{x}{y} < z\}$

Faisons le changement de variables suivant : $\begin{vmatrix} u=x \\ v=\frac{x}{y}\end{vmatrix} \Leftrightarrow \begin{vmatrix} x=u \\ y=\frac{u}{v}\end{vmatrix}$

$(x,y) \in D_{x,y} \Leftrightarrow (u,v) \in D_{u,v}$ avec $D_{u,v} = \{(u,v) \in \mathbb{R}^2 / v < z\}$

$F_Z(z) = \iint_{D_{x,y}} f_{X,Y}(x,y)dxdy=\iint_{D_{u,v}} f_{X,Y}(x(u,v),y(u,v))\left|\frac{D(x,y)}{D(u,v)}\right|dudv$

$$\frac{D(x,y)}{D(u,v)}=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}=\begin{vmatrix} 1 & 0 \\ \frac{1}{v} & \frac{-u}{v^2}\end{vmatrix}=\frac{-u}{v^2}$$

$F_Z(z) = \int_{-\infty}^{z}\int_{-\infty}^{+\infty} f_{X,Y}(u,\frac{u}{v})\left|\frac{-u}{v^2}\right|dudv=\int_{-\infty}^{z} f_Z(v)dv$, ce qui permet d'obtenir l'expression de la densité de Z : $f_Z(z)=\int_{-\infty}^{+\infty} f_{X,Y}(u,\frac{u}{z})\left|\frac{-u}{z^2}\right|du=\int_{-\infty}^{+\infty} f_{X,Y}(x,\frac{x}{z})\left|\frac{-x}{z^2}\right|dx$.

Il est possible de montrer que $f_Z(z)=\int_{-\infty}^{+\infty} f_{X,Y}(zy,y)|y|dy$, lorsque le changement de variables suivant est utilisé : $\begin{vmatrix} u=y \\ v=\frac{x}{y}\end{vmatrix} \Leftrightarrow \begin{vmatrix} y=u \\ x=uv\end{vmatrix}$ et $\frac{D(x,y)}{D(u,v)}=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}=\begin{vmatrix} v & u \\ 1 & 0\end{vmatrix}=-u$

Si en plus les V A X et Y sont indépendantes alors la densité de Z = $\frac{X}{Y}$, a pour expression :

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(\frac{x}{z}) \left|\frac{-x}{z^2}\right| dx = \int_{-\infty}^{+\infty} f_X(zy) f_Y(y) |y| dy .$$

Application : Calculer la loi du rapport de deux V A indépendantes, normales centrées réduites :

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \text{ et } f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

x $\in$ R et y $\in$ R alors z = $\frac{x}{y}$ $\in$ R

$$f_Z(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} e^{-\frac{x^2}{2z^2}} \left|\frac{-x}{z^2}\right| dx = \frac{1}{\pi} \int_{0}^{+\infty} e^{-\frac{x^2}{2}} e^{-\frac{x^2}{2z^2}} \frac{x}{z^2} dx$$, posons u = x^2

$$f_Z(z) = \frac{1}{2\pi} \int_{0}^{+\infty} e^{-(1+\frac{1}{z^2})\frac{u}{2}} \frac{du}{z^2} = -\left[\frac{1}{\pi z^2 (1+\frac{1}{z^2})} e^{-(1+\frac{1}{z^2})\frac{u}{2}} \right]_0^{+\infty} = \frac{1}{\pi(1+z^2)}$$, c'est la loi de Cauchy.

Lois de probabilité déduites de la normale centrée réduite

La loi de khi deux χ_n^2

La loi de khi deux χ_n^2 est une VA positive qui dépend de l'entier n

$\chi_n^2 = \sum_{i=1}^{n} X_i^2$, avec $X_i \to N(0, 1)$

Il est possible de démontrer par récurrence que la densité de la loi de χ_n^2 est :

$$f(x) = \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma(n/2)} \text{ avec } \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$$

On utilisera pour cela le fait que $\int_0^{+\infty} t^{r-1} (1-t)^{s-1} dt = \frac{\Gamma(r)\Gamma(s)}{\Gamma(s+r)}$

Loi de Fisher Snedecor F(m,n)

C'est une V A positive qui dépend de deux entiers m et n :

$F(m,n) = \frac{\frac{X}{m}}{\frac{Y}{n}}$, avec X $\to$ χ_m^2 et Y $\to$ χ_n^2

La densité d'une loi Fisher Snedecor F(m,n) : $f(z) = \dfrac{z^{\frac{m}{2}-1}\left(\frac{m}{n}\right)^{\frac{m}{2}}\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)\left(\frac{mz}{n}+1\right)^{\frac{m+n}{2}}}$

La loi de Student

C'est une V A qui dépend de l'entier n,

$T_n = \dfrac{Z}{\sqrt{\frac{X}{n}}}$, Z → N(0, 1) et X → χ_n^2

Le domaine de variation de la loi de Student est l'ensemble des réels R, comme pour la loi normale.

La densité d'une loi de Student T_n : $f(t) = \dfrac{1}{\sqrt{n\pi}} \dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\left(\frac{t^2}{n}+1\right)^{\frac{n+1}{2}}}$

Il est possible de procéder par deux manières pour établir ce résultat :
(1) une loi de Student est le rapport de deux variables aléatoires.
(2) considérer le fait que $T_n^2 = F(1,n)$ (le carré d'une loi de Student T_n est une loi de Fisher Snedecor F(1,n)).

Exercices

Exercice 1

Soit K une V A qui suit une loi géométrique :

$$P(K=k)=\begin{cases} C(1-p)^{k-1}, \mathrm{k}=1,2,... \\ 0 \,\text{sinon} \end{cases}$$

1) Déterminer C
2) Soit N $\in \mathcal{N}$, calculer la probabilité pour que K soit supérieure strictement à N.
3) Sachant que la V A, K est plus grande que N, quelle est la probabilité (conditionnelle) qu'elle soit plus grande que 2N ?
4) Quelle est la probabilité pour que K soit un multiple 3 ?

Exercice 2

La probabilité qu'une ampoule particulière claque durant le k-ième mois d'utilisation est donnée par la loi : P(K= k) = $\left(\frac{1}{5}\right)\left(\frac{4}{5}\right)^{k-1}$, k =1,2,...

Quatre ampoules sont testées simultanément. Déterminer la probabilité que :
1) Aucune des quatre ampoules ne claque durant le premier mois d'utilisation.
2) Exactement 2 ampoules ont claqué à la fin du troisième mois.
3) Exactement une ampoule a claqué durant chacun des trois premiers mois.
4) Exactement une ampoule a claqué à la fin du second mois et exactement deux ampoules fonctionnent encore au début du cinquième mois.

Exercice 3

On jette deux dés. Quelle est la probabilité pour que la somme des points donnés par la valeur des dés soit i = 2, 3, …, 11, 12.

Exercice 4

Une urne contient trois boules rouges et sept boules noires. Les joueurs A et B tirent une boule à tour de rôle jusqu'à ce qu'une rouge sorte et A commence le premier. Trouver la probabilité pour que le joueur A tire la première boule rouge, les boules tirées ne sont pas remises.

Exercice 5

On admet que 5% des hommes et 0,25% des femmes sont daltoniens. On sélectionne une personne daltonienne au hasard. Quelle est la probabilité qu'il s'agisse d'un homme ? On admettra que les hommes sont aussi nombreux que les femmes. Si au contraire il y en avait deux fois plus que de femmes, que deviendrait le résultat ?

Exercice 6

Le nombre de clients d'un bureau de poste en l'espace d'une journée est une variable aléatoire poissonienne de paramètre λ. La probabilité qu'un homme rentre dans le bureau est p. Montrer que le nombre des hommes et celui des femmes parmi les clients quotidiens sont des variables aléatoires poissoniènnes de paramètres respectifs λp et λ(1-p) et qu'elles sont indépendantes.

Exercice 7

Soient X et Y deux variables aléatoires indépendantes de fonctions de répartition F_X et F_Y, calculer la fonction de répartition des variables aléatoires Min(X,Y) et Max(X,Y).

Application à X uniforme sur $\left[-\frac{1}{2},\frac{1}{2}\right]$ et Y uniforme sur [-1,1]

Exercice 8

Soit la fonction densité conjointe de X et Y donnée par :

$f_{X,Y}(x,y) = c(y^2-x^2)e^{-y}$ pour $-y \leq x \leq y$ et $0 \leq y \leq +\infty$

Calculer c et les densités marginales de X et de Y

Exercice 9

Les moteurs d'un avion ont une probabilité de 1-p de défaillance en cours du vol, et ce indépendamment les uns des autres. Un avion a besoin d'une majorité de ses moteurs (supérieur ou égale à la moitié) pour pouvoir terminer son vol. Pour quelles valeurs de p un avion à quatre (respectivement cinq) moteurs est-il préférable à un bimoteur (respectivement trimoteur) ?

3. Les Moments d'une loi de probabilité

Les moments sont des caractéristiques numériques des variables aléatoires.

Espérance mathématique

L'espérance mathématique d' V A est aussi désignée par le moment d'ordre 1 de la variable. Pour une variable discrète X, elle est égale à : $E(x) = \sum_{x_i \in X(\Omega)} x_i P(X = x_i)$, cela suppose évidemment l'existence de cette somme surtout pour une variable discrète infinie.

Pour une variable à densité X, elle est égale à : $E(X) = \int_R x f_X(x) dx$, cela suppose évidemment l'existence de l'intégrale.

Proposition 1

Soit Y une variable aléatoire à densité, pour laquelle l'espérance existe alors :

$$E(Y) = \int_0^{+\infty} P(Y > y) dy - \int_0^{+\infty} P(Y < -y) dy$$

Démonstration :

$$\int_0^{+\infty} P(Y > y) dy - \int_0^{+\infty} P(Y < -y) dy = \int_0^{+\infty} (1 - F_Y(y)) dy - \int_0^{+\infty} F_Y(-y) dy$$

$$\int_0^{+\infty} (1 - F_Y(y)) dy = [y(1 - F_Y(y)]_0^{+\infty} + \int_0^{+\infty} y f_Y(y) dy$$

$$y(1 - F_Y(y)) = y \int_y^{+\infty} f(t) dt < \int_y^{+\infty} t f(t) dt \xrightarrow[y \to +\infty]{} 0 \text{ car E(Y) existe}$$

$$\int_0^{+\infty} F_Y(-y) dy = [y F_Y(-y)]_0^{+\infty} + \int_0^{+\infty} y f_Y(-y) dy \text{, posons y = -s :}$$

$$\int_0^{+\infty} y f_Y(-y) dy = \int_0^{-\infty} s f_Y(s) ds$$

$$y F_y(-y) = -s F_Y(s) = |s| F_Y(s) = |s| \int_{-\infty}^{s} f_Y(t) dt < -\int_{-\infty}^{s} t f_Y(t) dt \xrightarrow[s \to -\infty]{} 0 \text{ car E(Y) existe}$$

Donc $\int_0^{+\infty} P(Y > y) dy - \int_0^{+\infty} P(Y < -y) dy = \int_0^{+\infty} y f_Y(y) dy - \int_0^{-\infty} s f_Y(s) ds = E(Y)$

La proposition 1 peut s'écrire également dans le cas discret : $E(Y) = \sum_{i=0}^{+\infty} P(Y > i)$, en effet :

P(Y > 0) = P(Y = 1) + P(Y = 2) + … + P(Y = n) + …
P(Y > 1) = + P(Y = 2) + … + P(Y = n) + …
…
P(Y > n-1) = P(Y = n) + …
…
En sommant ces lignes, on trouve :

P(Y>0) + P(Y>1) + ... + P(Y>n-1) + .. = P(Y=1) + 2P(Y=2) + ... + nP(Y=n) + ... = E(Y)

Proposition 2

Soit g une fonction réelle et X une V A discrète, alors : $E(g(X)) = \sum_i g(x_i)P(X = x_i)$

Si X est une V A continue de densité f_X alors : $E(g(X)) = \int_R g(x)f_X(x)dx$

Démonstration :

Commençons par le cas discret et considérons deux V A discrètes X et Y avec Y = g(X) :

$P(Y = y_j) = \sum_{i \in I_j} P(X = x_i)$, $I_j = \{i \in N / g(x_i) = y_j\}$ et $I = \{i \in N / x_i \in X(\Omega)\}$

$$E(Y) = \sum_j y_j P(Y = y_j) = \sum_j y_j \sum_{i \in I_j} P(X = x_i) = \sum_j \sum_{i \in I_j} y_j P(X = x_i)$$

$E(Y) = \sum_i g(x_i)P(X = x_i)$, les ensembles I_j constituent une partition de l'ensemble I.

Si X et Y sont deux V A à densité avec g(X) = Y :

$\int_0^{+\infty} P(Y > y)dy = \int_0^{+\infty} P(g(X) > y)dy = \int_0^{+\infty} \int_{x:g(x)>y} f_X(x)dxdy$, par application du théorème de Fubini

on a : $\int_0^{+\infty} P(Y > y)dy = \int_{x:g(x)>0} \int_0^{g(x)} f_X(x)dydx = \int_{x:g(x)>0} g(x)f_X(x)dx$

$\int_0^{+\infty} P(Y < -y)dy = \int_0^{+\infty} P(g(X) < -y)dy = \int_0^{+\infty} \int_{x:g(x)<-y} f_X(x)dxdy$, posons s = -y

$$-\int_0^{-\infty} \int_{x:g(x)<s} f_X(x)dxds = \int_{-\infty}^{0} \int_{x:g(x)<s} f_X(x)dxds = \int_{x:g(x)<0} \int_{g(x)}^{0} f_X(x)dsdx = -\int_{x:g(x)<0} g(x)f_X(x)dx$$

Finalement $\int_0^{+\infty} P(Y > y)dy - \int_0^{+\infty} P(Y < -y)dy = E(Y) = E(g(X))$

<u>Cas particulier :</u> g(X) = aX + b, a, b deux réels

Dans le cas discret : $E(aX + b) = \sum_i (ax_i + b)P(X = x_i) = aE(X) + b$

Dans le cas continu : $E(aX + b) = \int_R (ax + b)f_X(x)dx = aE(X) + b$

Remarques :

(1) L'intérêt de la proposition 2 réside dans le fait que l'espérance de g(X) peut être calculée à partir des caractéristiques de X (on peut se passer du calcul de la densité de g(X) ou de la loi de g(X)).

(2) E(g(X)) est en général différent de g(E(X)), en effet :
Exemple : g(X) = 1/X

X	1	2
P(X)	1/2	1/2

1/X	1	1/2
P(g(X))	1/2	1/2

E(X) = 1*1/2 + 2*1/2 = 3/2 et E(g(X)) = 1*1/2 + 1/2*1/2 = 3/4
g(E(x)) = 2/3 ≠ 3/4

Espérance d'une combinaison linéaire de variables aléatoires

X et Y sont deux V A, et α, β sont deux réels, on se propose de calculer E(αX + βY).

Si X et Y sont des variables discrètes alors :

$$\begin{aligned}E(\alpha X+\beta Y) &= \sum_i\sum_j(\alpha x_i+\beta y_j)P(X=x_i,Y=y_j)\\ &= \sum_i\sum_j \alpha x_i P(X=x_i,Y=y_j)+\sum_j\sum_i \beta y_j P(X=x_i,Y=y_j)\\ &= \sum_{\mathrm{i}} \alpha x_i p_{i.}+\sum_j \beta y_j p_{.j}\\ &= \alpha \mathrm{E(X)}+\beta \mathrm{E(Y)}\end{aligned}$$

Si X et Y sont des variables à densité alors :

$$\begin{aligned}E(\alpha X+\beta Y) &= \iint_{R\times R}(\alpha x+\beta y)f_{X,Y}(x,y)dxdy\\ &= \iint_{R\times R}\alpha x f_{X,Y}(x,y)dydx+\iint_{R\times R}\alpha y f_{X,Y}(x,y)dxdy\\ &= \int_{\mathrm{R}}\alpha x \mathrm{f_X}(x)dx+\int_{\mathrm{R}}\beta y \mathrm{f_Y}(y)dy\\ &= \alpha \mathrm{E(X)}+\beta \mathrm{E(Y)}\end{aligned}$$

Une caractéristique des variables aléatoires indépendantes

X et Y deux V A indépendantes et f et g sont deux fonctions réelles, on se propose de calculer E(f(X)g(Y)) :
Si X et Y sont des variables à densité alors :
$E(f(X)g(Y))=\iint\limits_{R\times R}f(x)g(y)f_{X,Y}(x,y)dxdy$, les variables X et Y sont indépendantes, ce qui implique que la densité conjointe est égale au produit des densités marginales :

$$E(f(X)g(Y))=\iint\limits_{R\times R}f(x)g(y)f_X(x)f_Y(y)dxdy=\int\limits_R f(x)f_X(x)dx\int\limits_R g(y)f_Y(y)dy=E(f(X))E(g(Y))$$

Si X et Y sont des variables discrètes alors :
$E(f(X)g(Y))=\sum_i\sum_j f(x_i)g(y_j)P(X=x_i,Y=y_j)$, l'hypothèse d'indépendance des deux variables aléatoires entraine que :

$$\begin{aligned}E(f(X)g(Y)) &= \sum_i\sum_j f(x_i)g(y_j)P(X=x_i)P(Y=y_j)=\sum_i f(x_i)P(X=x_i)\sum_j g(y_j)P(Y=y_j)\\ &= \mathrm{E(f(X))E(g(Y))}\end{aligned}$$

Dans le cas particulier où les fonctions f et g sont la fonction identité, on a :

E(XY) = E(X)E(Y)

Moments d'ordre n

Définition
Le moment d'ordre n de la V A discrète X, de loi de probabilité P(X = x) est défini par :

$E(X^n) = \sum_i x_i^n P(X = x_i)$, sous réserve de l'existence de cette quantité.

Le moment d'ordre n de la V A continue X, de densité f_X, est défini par :

$E(X^n) = \int_R x^n f_X(x)dx$, sous réserve de l'existence de cette quantité.

Définition

Le moment centré d'ordre n d'une V A, est défini par : $E[(X - E(X))^n]$

Variance : C'est le moment centré d'ordre 2

$Var(X) = E[(X - E(X))^2] = E[X^2 - 2XE[X] + E(X)^2] = E[X^2] - E(X)^2$

En effet : si a, b $\in$ R et X, Y deux V A, alors :

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

Var(X) est une caractéristique numérique de la V A, qui est **positive** par construction.

On désigne par $\sigma(X)$ l'écart-type de X tel que : $\sigma(X) = \sqrt{\text{var}(X)}$.

Propriétés

(1) L'espérance mathématique E(X) minimise la fonction $\Phi : x \mapsto E[(X-x)^2]$

$\Phi(x) = x^2 - 2xE(X) + E(X^2)$

$\Phi'(E(X)) = 0$ et $\Phi''(x) = 2 > 0$

Le minimum est atteint à x = E(X) et sa valeur = Var(X)

(2) La variance est un opérateur non linéaire : $var(aX + b) = a^2 var(X)$

$E(aX + b) = aE(X) + b$

$Var(aX + b) = E[(a(X+b) - aE(X) - b)^2 = a^2E[(X - E(X))^2]$

(3) La variance est un indicateur de dispersion, plus l'écart entre les valeurs est grand plus la variance est grande, en effet, considérons les trois variables aléatoires Y, Z et W :

Y	-1	1
P(Y)	1/2	1/2

Z	-100	100
P(Z)	1/2	1/2

W	0
P(W)	1

$E(Y) = E(Z) = E(W) = 0$

$E(Y^2) = 1/2 + 1/2 = 1$, ce qui entraine que var(Y) = 1

$E(Z^2) = 10000*1/2 + 10000*1/2 = 10000$, ce qui entraine que var(Y) = 10000

$E(W^2) = 0$, ce qui entraine que var(W) = 0

On remarque que var(W) < var(Y) < var(Z) , la variance croit avec l'écart entre les valeurs de la variable

Var(Z) peut être calculée à partir de la relation Z = 100Y, ce qui entraine que :

var(Z) = 10000var(Y)

Moments des lois usuelles

Loi normale $N(\mu, \sigma)$

$$E(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Posons $u = \dfrac{x-\mu}{\sigma} \Rightarrow \sigma du = dx$

$$E(X) = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u e^{-\frac{u^2}{2}} dx + \mu = \frac{\sigma}{\sqrt{2\pi}} \left[-e^{-\frac{u^2}{2}} \right]_{-\infty}^{+\infty} + \mu = \mu$$

$Var(X) = E\left[(X - E(X))^2\right] = \dfrac{1}{\sqrt{2\pi}\sigma} \displaystyle\int_{-\infty}^{+\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$, posons $u = \dfrac{x-\mu}{\sigma} \Rightarrow \sigma du = dx$

$$Var(X) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u^2 e^{-\frac{u^2}{2}} du = \frac{\sigma^2}{\sqrt{2\pi}} \left[-u e^{-\frac{u^2}{2}} \right]_{-\infty}^{+\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du = \sigma^2$$

Loi uniforme U(a, b)

$$E(X) = \int_a^b \frac{x}{b-a} dx = \left[\frac{x^2}{2(b-a)} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

$$E(X^2) = \int_a^b \frac{x^2}{b-a} dx = \left[\frac{x^3}{3(b-a)} \right]_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ab + a^2}{3}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{b^2 + ab + a^2}{3} - \frac{(b+a)^2}{4} = \frac{(b-a)^2}{12}$$

Loi exponentielle $\mathcal{E}(\lambda)$

$$E(X) = \int_0^{+\infty} \lambda x e^{-\lambda x} dx = \left[-x e^{-\lambda x} \right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx = \left[-\frac{e^{-\lambda x}}{\lambda} \right]_0^{+\infty} = \frac{1}{\lambda}$$

$$E(X^2) = \int_0^{+\infty} \lambda x^2 e^{-\lambda x} dx = \left[-x^2 e^{-\lambda x} \right]_0^{+\infty} + \int_0^{+\infty} 2x e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$Var(X) = E(X^2) - E(X)^2 = \dfrac{1}{\lambda^2}$

Loi de Cauchy

La loi de Cauchy ayant pour densité la fonction $f(x) = \dfrac{1}{\pi(1+x^2)}$ n'admet aucun moment car l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \frac{x}{\pi(1+x^2)} dx$ diverge.

Loi de Bernoulli B(p)

X	0	1
P(X)	1-p	p

$E(X) = 0*(1-p) + p*1 = p$, $E(X^2) = 0*(1-p) + p*1 = p$
$Var(X) = E(X^2) – E(X)^2 = p – p^2 = p(1-p)$.

Loi Binomiale B(n, p)

Si Y est une loi Binomiale B(n, p) alors elle peut être considérée comme la somme de n variables indépendantes de Bernoulli B(p), X_i

$$Y = \sum_{i=1}^{n} X_i \Rightarrow E(Y) = E(\sum_{i=1}^{n} X_i) = np$$

$$Var(Y) = \mathrm{var}(\sum_{i=1}^{n} X_i) = np(1-p)$$

Loi de Poisson P(λ)

$$E(X) = \sum_{k=0}^{+\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda} = \lambda\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!}e^{-\lambda} = \lambda$$

$$E(X^2) = \sum_{k=0}^{+\infty} k^2\frac{\lambda^K}{k!}e^{-\lambda} = \lambda\sum_{k=1}^{+\infty} k\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda} = \lambda\sum_{i=0}^{+\infty}(i+1)\frac{\lambda^i}{i!}e^{-\lambda} = \lambda e^{-\lambda}\left(\sum_{i=0}^{+\infty} i\frac{\lambda^i}{i!} + \sum_{i=0}^{+\infty}\frac{\lambda^i}{i!}\right)$$

$E(X^2) = \lambda^2 + \lambda$ et $var(X) = E(X^2) – E(X)^2 = \lambda$

Covariance

Définition

Soient X et Y deux variables aléatoires, la covariance de X et Y notée Cov(X, Y) est définie par : cov(X, Y) = E[(X- E(X))(Y- E(Y))] = E(XY) – E(X)E(Y)

Remarques

(1) La covariance est symétrique : Cov(X, Y) = Cov(Y, X)
(2) Cov(X, X) = Var(X)
(3) Cov(aX + b, Y) = aCov(X, Y)
(4) La covariance de deux variables indépendantes est nulle, car dans ce cas E(XY) = E(X)E(Y), mais la réciproque est fausse, en effet, considérons le couple (X,Y) discret dont la loi est la suivante :

Y \ X	0	1	2	$p_{i.}$
0	0	1/2	0	1/2
1	1/4	0	1/4	1/2
$p_{.j}$	1/4	1/2	1/4	

Calculons la loi de Z = XY :
$Z(\Omega) = \{0, 1, 2\}$
Z=0 ≡ X=0 ou Y=0, donc P(Z=0) = P(X=0) + P(Y=0) – P(X=0, Y=0) = 1/4 + 1/2 = 3/4
P(Z=1) = P(X=1,Y=1) = 0
P(Z=2) = P(X=2,Y=1) = 1/4

La loi de Z est donc la suivante :

Z	0	1	2
P(Z)	3/4	0	1/4

Les variables X et Y sont dépendantes car $P(X=0,Y=0) = 0 \neq P(X=0)P(Y=0) = 1/8$
$E(XY) = E(Z) = 0*3/4 + 1*0 + 2*1/4 = 1/2$
$E(X) = 0*1/4 + 1*1/2 + 2*1/4 = 1$
$E(Y) = 0*1/2 + 1*1/2 = 1/2$
$Cov(X,Y) = E(XY) – E(X)E(Y) = 1/2 – 1/2 = 0$.

Variance d'une somme de variables aléatoires

Soient X et Y deux variables aléatoires : $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$
En effet :
$Var(X + Y) = E[(X + Y)^2] – [E(X + Y)]^2$
$E[(X + Y)^2] = E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2)$ (1)
$[E(X + Y)]^2 = [E(X) + E(Y)]^2 = E(X)^2 + 2E(X)E(Y) + E(Y)^2$ (2)
$(1) – (2) = Var(X) + Var(Y) + 2Cov(X, Y)$

La variance de la somme de X_i i = 1, … ,n V A :

$$\operatorname{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{var}(X_i) + 2\sum_{1\le i<j\le n} \operatorname{cov}(X_i, X_j)$$

La corrélation linéaire

Le coefficient de corrélation linéaire de deux V A X et Y, est définie par :

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\operatorname{var}(Y)}} \text{ si } \operatorname{var}(X)\operatorname{var}(Y) \neq 0$$

Lorsque $\rho(X, Y) = 0$, les V A X et Y sont dits non corrélées.
Remarque : Deux V A indépendantes sont nécessairement non corrélées.

Proposition

X et Y étant deux V A de variances non nulles, leur coefficient de corrélation est toujours de module inférieur à 1 :

(1) $-1 \le \rho(X,Y) \le 1$

(2) $\rho(X, Y) = \pm 1 \Leftrightarrow Y = aX + b$ avec $a, b \in R$

Démonstration

(1) $\forall \lambda \in R$, $Var(X + \lambda Y) \geq 0$
Or $Var(X + \lambda Y) = Var(X) + 2\lambda Cov(X, Y) + \lambda^2 Var(Y) \geq 0$
Le discriminent de cette équation de second degré est négatif ou nul :

$$\Delta' = Cov(X, Y)^2 – Var(X)Var(Y)$$

$$\Delta' = \operatorname{var}(X)\operatorname{var}(Y)\left(\frac{\operatorname{cov}(X,Y)^2}{\operatorname{var}(X)\operatorname{var}(Y)} - 1\right) = \operatorname{var}(X)\operatorname{var}(Y)(\rho^2(X,Y) - 1)$$

$$\Delta' \le 0 \Leftrightarrow \rho(X, Y)^2 – 1 \le 0 \Leftrightarrow -1 \le \rho(X, Y) \le 1$$

(2) supposons que $Y = aX + b$, a et b deux réelles,
$Cov(X, Y) = Cov(X, aX + b) = aVar(X)$
$Var(Y) = Var(aX + b) = a^2Var(X)$

$$\rho(X,Y)^2 = \frac{\operatorname{cov}(X,Y)^2}{\operatorname{var}(X)\operatorname{var}(Y)} = \frac{a^2 \operatorname{var}(X)^2}{a^2 \operatorname{var}(X)^2} = 1$$

Soit que $\rho(X, Y) = \pm 1$

Supposons que $\rho(X, Y) = \pm 1$ alors $\Delta' = 0$

$\Delta' = 0 \Rightarrow \exists\, \lambda_0 \in R$, tel que $Var(X + \lambda_0 Y) = 0$

$\Rightarrow X + \lambda_0 Y = cst$

$\Rightarrow Y = aX + b$

Espérance conditionnelle

X et Y sont deux V A, considérons X/Y, la distribution de X sachant Y=y.

Si X et Y sont deux V A discrètes alors la loi conditionnelle est définie par :

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

On définit l'espérance conditionnelle par : E(X/Y=y) $= \sum_{x \in X(\Omega)} xP(X = x / Y = y)$, sous réserve de l'existence de cette somme.

Si X et Y sont deux V A à densité alors la loi conditionnelle est définie par :

$$f_{X/Y}(x/y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

On définit l'espérance conditionnelle par : E(X/Y=y) $= \int_R x f_{X/Y}(x/y)dx$, sous réserve de l'existence de cette intégrale.

Remarque : L'espérance conditionnelle de la distribution de X/Y est une fonction de y

Théorème : Calcul de l'espérance par conditionnement

X et Y étant deux V A alors : $E(X) = E(E(X/Y=y)) = E_Y(E_X(X/Y=y))$

Démonstration

X et Y deux V A à densité :

$$E_X(X) = \int_R x f_X(x)dx = \int_R\int_R x f_{X,Y}(x,y)dydx = \int_R\int_R x f_{X/Y}(x/y) f_Y(y)dxdy$$

$$= \int_R E_X(X/Y = y) f_Y(y)dy = E_Y(E_X(X/Y = y))$$

X et Y sont deux V A discrètes :

$$E_X(X) = \sum_x xP(X = x) = \sum_x\sum_y xP(X = x, Y = y) = \sum_x\sum_y xP(X = x / Y = y)P(Y = y)$$

$$= \sum_y E_X(X/Y = y)P(Y = y) = E_Y(E_X(X/Y = y))$$

Variance conditionnelle

Définition : $Var(X/Y=Y) = E((X - E(X/Y))^2/Y = y) = E(X^2/Y=y) - E(X/Y=y)^2$

La variance conditionnelle se définit comme le moment d'ordre 2 de la distribution conditionnelle X/Y=y

Proposition
X, Y étant deux V A et g une fonction réelle, la variance conditionnelle de Y sachant X minimise : $E[(Y - g(X))^2/X=x]$.

Démonstration
$E[(Y - g(X))^2/X=x] = E[(Y - E(Y/X=x) + E(Y/X=x) - g(X))^2/X=x]$
$= E[(Y - E(Y/X))^2/X=x] + 2E[(Y - E(Y/X))(E(Y/X) - g(X))/X=x] + E[(E(Y/X) - g(X))^2/X]$
$E(Y/X) - g(X)$ ne dépend que x, c'est donc une constante par rapport à la variable Y, il s'en suit alors que : $E[(Y - E(Y/X))(E(Y/X) - g(X))/ X] = (E(Y/X) - g(X))E[(Y - E(Y/X))/X] = 0$
Finalement, nous obtenons l'égalité suivante :
$E[(Y - g(X))^2/X] = E[(Y - E(Y/X))^2/X] + E[(E(Y/X) - g(X))^2/X]$
Le minimum de $E[(Y - g(X))^2/X]$ est égale à $E[(Y - E(Y/X))^2/X]$, il est atteint si $g(X) = E(Y/X)$

Théorème
Soit X et Y deux variables aléatoires, alors : $Var_X(X) = E_Y[Var_X(X/Y)] + Var_Y(E_X[X/Y])$.

Démonstration

$$\begin{aligned} E_Y[Var_X(X/Y)] &= E_Y[E_X[X^2/Y] - E_X[X/Y]^2] \\ &= E_Y[E_X[X^2/Y]] - E_Y[E_X[X/Y]^2] \\ &= E_X[X^2] - E_Y[E_X[X/Y]^2] \end{aligned}$$

De même : $Var_Y(E_X[X/Y]) = E_Y[E_X[X/Y]^2] - E_Y[E_X[X/Y]]^2$
$= E_Y[E_X[X/Y]^2] - E_X[X]^2$
L'addition des deux égalités permet de trouver le résultat.

L'intérêt de ce théorème est la possibilité de calculer Var(X) à partir de quantités conditionnelles qui peuvent être connues.

Somme d'un nombre aléatoire de variables indépendantes et identiquement distribuées (iid)

On considère la V A S, qui est la somme d'un nombre aléatoire de V A X_i, i = 1, …, n.
$S = X_1 + …. + X_n$, les variables X_i sont indépendantes et suivent la même loi, elles sont donc identiquement distribuées. On se propose de calculer l'espérance et la variance de S.
Cette situation est fréquemment rencontrée en assurance si les variables X_i désignent les montants des dommages causés par les sinistres survenus dans une période de temps donnée. S est alors le montant total de tous les sinistres. L'hypothèse supposée est que tous les sinistres sont indépendants et suivent la même loi de probabilité. Le nombre des sinistres survenus durant la période de temps donnée est une variable aléatoire.

$$E(S/N=n) = E\left(\sum_{i=1}^{n} X_i\right) = nE(X)$$
$$E(S) = E(E(S/N)) = E(NE(S)) = E(N)E(S)$$

$Var(S/N=n) = nvar(X)$ et $E(var(S/N)) = E(Nvar(X)) = E(N)var(X)$

$E(S/N=n) = nE(X)$ et $var(E(S/N)) = var(NE(X)) = E(X)^2 var(N)$

Donc $var(S) = E(N)var(X) + E(X)^2 var(N)$

Usage de la table de la fonction de répartition de la loi normale centrée réduite

Connaissant l'abscisse u, la table permet de donner F(u) = P(U<u) si u est positif, par exemple si u=1,53 alors F(u) se lit sur l'intersection de la ligne 1,5 et la colonne 0,03 et qui correspond à la valeur de 0,937

Connaissant une probabilité donnée comprise entre 0,5 et 1, il est possible de déterminer l'abscisse u de la fonction de répartition.

Calculons par exemple u tel que F(u) = 0,3, dans ce cas u est négatif et P(U<-u) = 0,7, la valeur de -u peut être lue directement sur la table : $0{,}52 \leq -u \leq 0{,}53$
Pour estimer une valeur de –u, on peut faire une interpolation linéaire qui consiste à considérer qu'il existe une relation linéaire entre p et –u, cela revient à résoudre le système suivant : $\begin{cases} 0{,}52 = a0{,}6985 + b \\ 0{,}53 = a0{,}7019 + b \end{cases}$

$$a = \frac{\begin{vmatrix} 0{,}52 & 1 \\ 0{,}53 & 1 \end{vmatrix}}{\begin{vmatrix} 0{,}6985 & 1 \\ 0{,}7019 & 1 \end{vmatrix}} = \frac{0{,}52 - 0{,}53}{0{,}6985 - 0{,}7019} = \frac{-0{,}01}{-0{,}0034} = 2{,}9411$$

$$b = \frac{\begin{vmatrix} 0{,}6985 & 0{,}52 \\ 0{,}7019 & 0{,}53 \end{vmatrix}}{\begin{vmatrix} 0{,}6985 & 1 \\ 0{,}7019 & 1 \end{vmatrix}} = \frac{0{,}53 * 0{,}6985 - 0{,}52 * 0{,}7019}{0{,}6985 - 0{,}7019} = \frac{0{,}0052}{-0{,}0034} = -1{,}5344$$

Après avoir déterminé a et b, on a : -u = 0,7a+b → -u = 0,5243, soit que u = -0,5243

Soit X →N(m,σ), on cherche x tel que P(X<x) = 0,83

$X < x \Leftrightarrow \frac{X-m}{\sigma} < \frac{x-m}{\sigma}$ avec $\frac{X-m}{\sigma} \rightarrow N(0,1)$

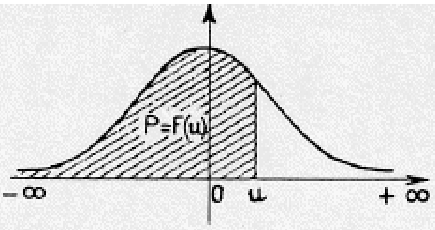
La table nous indique que : $0{,}95 < \frac{x-m}{\sigma} < 0{,}96$, on peut faire une interpolation linéaire pour estimer une valeur de $\frac{x-m}{\sigma}$, pour cela, le système suivant doit être résolu :

$$\begin{cases} 0{,}95 = a0{,}8289 + b \\ 0{,}96 = a0{,}8315 + b \end{cases}$$

Après avoir déterminé a et b, on a $\frac{x-m}{\sigma} = a0{,}83 + b$ = 0,9542 → x = 0,9542σ + m

Le tableur Excel peut être utilisé pour obtenir la probabilité p d'une loi normale centrée réduite, connaissant l'abscisse u : =LOI.NORMALE.STANDARD(u)
A l'inverse, connaissant p, il est possible de déterminer l'abscisse u :
=LOI.NORMALE.STANDARD.INVERSE(p)

Table de la fonction de répartition de la loi normale centrée réduite



u	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

Table pour les grandes valeurs de u

u	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,8	4,0	4,5
F(u)	0,99865	0,99904	0,99931	0,99952	0,99966	0,99976	0,999841	0,999928	0,999968	0,999997

Exercices

Exercice 1
Une paire de dé à quatre faces (1,2,3,4) est lancée une seule fois. X est la variable aléatoire discrète représentant le produit des faces inférieures des dés. Déterminer la variance conditionnelle de X^2 sachant que la somme des faces est plus grande que leur produit.

Exercice 2
Les variables aléatoires X et Y sont décrites par leur densité conjointe :

$$f_{X,Y}(x,y) = \begin{cases} K \text{ si } x + y \leq 1, x > 0, y > 0 \\ 0 \text{ sinon} \end{cases}$$

1) X et Y sont-elles indépendantes ?
2) X et Y sont-elles conditionnellement indépendantes sachant que max(X,Y) $\leq$ 0.5 ?
3) Soit R = XY, calculer E(R).
4) Si A et B sont les événements ‘‘2(Y-X) $\geq$ Y + X’’ et ‘‘Y $\geq$ 3/4’’, calculer P(A), P(B), P($\overline{A}$B), P($\overline{\overline{A}\,\overline{B}}$). Déterminer et faire le digramme de la densité conditionnelle $f_{X/\overline{A}\overline{B}}(x/\overline{A}\overline{B})$.

Exercice 3
Les variables aléatoires X et Y ont la densité conjointe :

$$f_{X,Y}(x,y) = \begin{cases} Ax \text{ si } 1 \leq x \leq y \leq 2 \\ 0 \text{ sinon} \end{cases}$$

1) Calculer A
2) Calculer la fonction densité de probabilité de Y : $f_Y(y)$
3) Quelle est $E\left[\frac{1}{X} / Y = \frac{3}{2}\right]$?
4) Quelle est $f_Z(z)$, la densité de probabilité de la variable aléatoire Z = X – Y ?

Exercice 4
Un ensemble a un cardinal de n éléments. On choisit de façon aléatoire et équiprobable un sous-ensemble non vide. Soit X la variable aléatoire discrète, cardinal du sous-ensemble choisi. Montrer que :

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

$$Var(X) = \frac{n2^{2n-2} - n(n+1)2^{n-2}}{\left(2^n - 1\right)^2}$$

Montrer que pour n→+∞, var(X) ≈ $\frac{n}{4}$, comparer ce résultat avec la forme limite que prend var(Y) pour P(Y = i) = $\frac{1}{n}$

On démontrera ou on utilisera les égalités suivantes :

$\sum_{k=1}^{n} kC_n^k = n2^{n-1}$ et $\sum_{k=1}^{n} k^2 C_n^k = n(n+1)2^{n-2}$

Exercice 5

Les variables aléatoires X et Y sont décrites par la densité conjointe :

$$f_{X,Y}(x,y) = \begin{cases} 1 \text{ si } 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 \text{ si non} \end{cases}$$ et la variable Z est définie par Z = XY.

Déterminer le second moment conditionnel de Z sachant que l'équation $r^2 + xr + y = 0$ a des racines réelles en r.

Exercice 6

Les variables aléatoires indépendantes X et Y sont décrites par leur densité marginale :

$$f_X(z) = f_Y(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \text{ avec } z \in R$$

Déterminer la loi de $Q = \frac{Y}{X}$, calculer E(Q) et Var(Q)

Exercice 7

Une pièce de monnaie a p pour probabilité de tomber sur face. On la lance indéfiniment. Calculer l'espérance du nombre de jets qu'il faudra jusqu'à ce qu'une chaine de r résultats consécutifs de type face apparaisse.

4. Les Transformées

Les transformées sont des opérateurs calculés au niveau des variables aléatoires. Elles permettent d'établir des propriétés très intéressantes.

Pour une variable discrète X, on défini la transformée en z :

$p_X^T(z) = E(z^X) = \sum_{n=0}^{+\infty} z^n P(X = n)$ Cette définition a un sens au mois lorsque $|z| \leq 1$.

Pour une variable à densité X, on défini la transformée en s :

$$f_X^T(s) = E(e^{-sX}) = \int_R e^{-sx} f_X(x)dx$$

On ne s'intéressera qu'aux cas où cette intégrale existe.

Propriétés relatives aux probabilités et aux moments

<u>Cas de la transformée en z</u>

$p_X^T(1) = 1$

$P(X = n) = \frac{1}{n!}\frac{d^n p_X^T(0)}{dz^n}$ pour $n \in N$

$E(X) = \frac{dp_X^T(1)}{dz}$

$Var(X) = \left[\frac{d^2 p_X^T(1)}{dz^2} + \frac{dp_X^T(1)}{dz}\right] - \left[\frac{dp_X^T(1)}{dz}\right]^2$

$\frac{d^n p_X^T(1)}{dz^n} = E(X(X-1)...(X-n+1))$

<u>Cas de la transformée en s</u>

$f_X^T(0) = 1$

$E(X) = -\frac{df_X^T(0)}{ds}$

$Var(X) = \frac{d^2 f_X^T(0)}{ds^2} - \left[\frac{df_X^T(0)}{ds}\right]^2$

$\frac{d^n f_X^T(0)}{ds^n} = (-1)^n E(X^n)$

Les démonstrations de toutes ces relations sont simples et abordables.
Ces propriétés nous montre que la connaissance de la transformée en z ou de la transformée en s d'une variable aléatoire, permet de déterminer la loi de cette variable ainsi que ses moments.

Transformées de lois discrètes :

Loi de Bernoulli B(p)

X	0	1
P(X)	1-p	p

$p_X^T(z) = z^0P(X=0) + zP(X=1) = 1\text{-}p + pz$

Loi de Poisson P(λ)

$$p_X^T(z) = \sum_{k=0}^{+\infty} z^k P(X=k) = \sum_{k=0}^{+\infty} z^k e^{-\lambda}\frac{\lambda^k}{k!} = e^{\lambda(z-1)}$$

Transformée en s de la loi normale N(μ, σ)

$$f_X^T(s) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{-sx} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$\frac{-2\sigma^2 sx}{2\sigma^2} - \frac{x^2-2x\mu+\mu^2}{2\sigma^2} = -\frac{x^2-2x(\mu-s\sigma^2)+(\mu-s\sigma^2)^2}{2\sigma^2} + \frac{(\mu-s\sigma^2)^2-\mu^2}{2\sigma^2}$$

$$= -\frac{(x-(\mu-s\sigma))^2}{2\sigma^2} + \frac{(\mu-s\sigma^2)^2-\mu^2}{2\sigma^2}$$

$$f_X^T(s) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{-\frac{(x-(\mu-s\sigma))^2}{2\sigma^2}} e^{\frac{(\mu-s\sigma^2)^2-\mu^2}{2\sigma^2}}dx = \frac{e^{\frac{(\mu-s\sigma^2)^2-\mu^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{-\frac{(x-(\mu-s\sigma))^2}{2\sigma^2}}dx$$

$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{-\frac{(x-(\mu-s\sigma))^2}{2\sigma^2}}dx = 1$, c'est l'intégrale de la densité de la loi N((μ-sσ),σ)

$f_X^T(s) = e^{\frac{s^2\sigma^2-2\mu s}{2}}$, ayant calculé la transformée en s, il est possible de déduire facilement l'espérance et la variance de la loi normale, par les formules :

$$E(X) = -\frac{df_X^T(0)}{ds} \text{ et } Var(X) = \frac{d^2 f_X^T(0)}{ds^2} - \left[\frac{df_X^T(0)}{ds}\right]^2$$

Propriétés de la transformée en s

Soit f la fonction définie par : $\begin{array}{l} f : R^+ \to R \\ t \mapsto t \end{array}$

Calculons la transformée en s de la fonction f :

$$f^T(s) = \int_0^{+\infty} te^{-st}dt = \left[-\frac{te^{-st}}{s}\right]_0^{+\infty} + \int_0^{+\infty}\frac{e^{-st}}{s}dt = \frac{1}{s^2}$$

On considère la fonction : $\begin{array}{l} f : R^+ \to R \\ t \mapsto \dfrac{t^n}{n!},\ n \in N^* \end{array}$

Il est possible de démontrer par récurrence que la transformée en s de la fonction f est :

$$f^T(s) = \frac{1}{s^{n+1}}$$

Soit f une fonction définie sur R^+, dont la transformée en s est Φ(f) :

$$\Phi(f)(s) = \int_0^{+\infty} e^{-st} f(t)dt$$

On définit la fonction : $f \circ \tau_a : [-a,+\infty[\to R$

$$y \mapsto f \circ \tau_a(y) = f(y+a)$$

La transformée en s de la fonction $f_o\tau_a$ est :

$$\Phi(f \circ \tau_a)(s) = \int_{-a}^{+\infty} e^{-sy} f \circ \tau_a(y)dy = \int_{-a}^{+\infty} e^{-sy} f(y+a)dy$$

Faisons le changement de variable : t = y+a

$$\Phi(f \circ \tau_a)(s) = \int_{0}^{+\infty} e^{-st+sa} f(y)dy = e^{sa}\Phi(f)(s)$$

Transformée en s de la somme de deux V A indépendantes

Proposition :
X et Y sont deux V A à densité et indépendantes, si Z est la V A telle que Z = X + Y, alors la transformée en s de Z est égale au produit des transformées en s des V A, X et Y :

$$f_Z^T(s) = f_X^T(s) f_Y^T(s)$$

Démonstration :

$$f_Z^T(s) = E(e^{-sZ}) = E(e^{-s(X+Y)}) = \int_R\int_R e^{-s(X+Y)} f_{X,Y}(x,y)dxdy = \int_R\int_R e^{-sx} f_X(x) e^{-sy} fy(y)dxdy = f_X^T(s) f_Y^T(s)$$

Application :

Supposons que les V A, X et Y suivent une loi uniforme U[0,1] :

$$f_X^T(s) = f_Y^T(s) = \int_0^1 e^{-st}dt = \left[-\frac{e^{-st}}{s}\right]_0^1 = \frac{1-e^{-s}}{s}$$

Dans ce cas $f_Z^T(s) = \left(\frac{1-e^{-s}}{s}\right)^2 = \frac{1-2e^{-s}+e^{-2s}}{s^2}$

Connaissant la transformée en s de la variable Z, essayons de calculer la fonction densité de probabilité f_Z. Pour cela décomposons $f_Z^T(s)$ en éléments dont on connait la réciproque de la transformée en s.

Si f_1 est la fonction définie par : $f_1 : R^+ \to R$

$$t \mapsto t$$

On a $\Phi(f_1)(s) = \frac{1}{s^2}$

Si f_2 est la fonction définie par : $f_2 : [1,+\infty[\to R$

$$t \mapsto -2(t-1)$$

On a $\Phi(f_2)(s) = \frac{-2e^{-s}}{s^2}$

Si f_2 est la fonction définie par : $f_3 : [2,+\infty[\to R$

$$t \mapsto t-2$$

On a $\Phi(f_3)(s) = \frac{e^{-2s}}{s^2}$

La fonction densité de probabilité de Z a pour expression :
Si $0 \le t < 1$, alors $f_Z(t) = f_1(t) = t$
Si $1 \le t < 2$, alors $f_Z(t) = f_1(t) + f_2(t) = t - 2t + 2 = -t + 2$
Si $2 \le t$, alors $f_Z(t) = f_1(t) + f_2(t) + f_3(t) = t - 2t + 2 + t - 2 = 0$

Ce résultat peut être retrouvé en utilisant le produit de convolution des V A, X et Y, en effet on a : $f_Z(z) = \int_0^1 f_X(x) f_Y(z-x)dx$, remarquons que z $\in [0,2]$ car (x,y) $\in [0,1]^2$

$0 \le z-x \le 1 \Leftrightarrow -z \le -x \le 1-z \Leftrightarrow z-1 \le x \le z$

Si $0 \le z \le 1$ alors $f_Z(z) = \int_0^z dx = z$

Si $1 \le z \le 2$ alors $f_Z(z) = \int_{z-1}^1 dx = 1-(z-1) = 2-z$

Somme d'un nombre aléatoire de V A indépendantes et identiquement distribuées (iid)

On considère la suite de V A indépendantes X_i, à densité et de même loi de probabilité. On pose $R = \sum_{i=1}^{N} X_i$ et N est une V A discrètes, R est la somme aléatoire de V A à densité.

Calcul de la densité de R

$$P(R < r) = \sum_{n \in N^*} P(R < r, N = n) = \sum_{n \in N^*} P(R < r / N = n) P(N = n)$$

$$F_R(r) = \sum_{n \in N^*} F_{R/N}(r/n) P(N = n)$$

$$f_R(r) = \sum_{n \in N^*} f_{R/N}(r/n) P(N = n)$$

Calcul de la transformée en s de R

$$f_R^T(s) = \int_{-\infty}^{+\infty} e^{-rs} f_R(r)dr = \int_{-\infty}^{+\infty} e^{-rs} \sum_{n \in N^*} f_{R/N}(r/n) P(N = n)dr = \sum_{n \in N^*} P(N = n) \int_{-\infty}^{+\infty} e^{-rs} f_{R/N}(r/n)dr$$

$$f_R^T(s) = \sum_{n \in N^*} P(N = n) \left(f_X^T(s)\right)^n$$

Exercices

Exercice 1

Soit U_i, i = 1,2,3 trois variables aléatoires indépendantes et uniformes sur [0, 1], calculer de deux manières différentes, la densité de la variable aléatoire $S = U_1 + U_2 + U_3$

Exercice 2

Toutes les parties de cet exercice requièrent des réponses numériques.

1) Si $f_Y^T(s) = K/(2+s)$, calculer K et $E(Y^3)$

2) Si $p_X^T(z) = (1+z^2)/2$, calculer $P(X = E(X))$ et σ_X

3) Si $f_Y^T(s) = 2(2 - e^{-s/2} - e^{-s})/3s$, calculer $E(e^{2X})$

4) Si $p_X^T(z) = A(1+3z)^3$, calculer $P(X = 2)$ et $E(X^3)$

Exercice 3

Déterminer si les expressions suivantes sont ou non des transformées en z valides pour une variable aléatoire discrète qui ne peut prendre que des valeurs entières positives.

1) $z^2 + 2z - 2$
2) $2 - z$
3) $(2 - z)^{-1}$

Exercice 4

Montrer qu'aucune des expressions suivantes ne constitue une transformée en s :

1) $(1 - e^{-5s})/s$
2) $7(4 + 3s)^{-1}$

Exercice 5

Soit L une variable aléatoire discrète dont les valeurs possibles sont positives. On donne :

$$p_L^T(z) = K\left[\frac{14 + 5z - 3z^2}{8(2-z)}\right]$$

Déterminer les valeurs de E(L), P(L = 1) et l'espérance conditionnelle de L sachant $L \neq 0$.

Exercice 6

Des boites sont rangées dans des cartons, eux-mêmes rangés dans des caisses. Le poids en grammes d'une boite est une variable aléatoire continue, de distribution :

$$f_X(x) = \lambda e^{-\lambda x}, \text{ avec } x \geq 0$$

Le nombre de boites dans chaque carton, K, est une variable aléatoire avec la distribution :

$$P(K = k) = \frac{\mu^k e^{-\mu}}{k!}, \text{ pour } k = 0, 1, 2, \dots$$

Le nombre de cartons dans chaque caisse, N, est une variable aléatoire avec la distribution :

$$P(N = n) = p^{n-1}(1 - p), \text{ pour } n = 1, 2, 3, \dots$$

Les variables X, K et N sont mutuellement indépendantes. Déterminer :

1) La probabilité qu'une caisse sélectionnée au hasard contienne exactement un carton, une boite,

2) la transformée en s de la distribution pour le poids total des boites dans une caisse,

3) la probabilité qu'une caisse sélectionnée au hasard contienne un nombre impair de boites.

5. Le comportement asymptotique des variables aléatoires

Nous commençons ce chapitre par traiter des inégalités qui ont vont plutôt servir à démontrer des propriétés qui concernent le comportement asymptotique ou la convergence des variables aléatoires.

Les Inégalités élémentaires

Inégalité de Markov

Soit X une V A positive, $\forall$ a > 0, P(X $\geq$a) $\leq \frac{E(X)}{a}$

Démonstration :

Si X est une variable à densité, alors :

$$E(X) = \int_0^{+\infty} x f_X(x)dx \geq \int_a^{+\infty} x f_X(x)dx \geq a\int_a^{+\infty} f_X(x)dx = aP(X > a)$$

Si X est une variable discrète, alors :

$$E(X) = \sum_{x>0} xP(X = x) < \sum_{x>a} xP(X = x) < a\sum_{x>a} P(X = x) = aP(X > a)$$

Il est possible que a soit assez faible de telle sorte que $\frac{E(X)}{a} > 1$, dans ce cas, le résultat apporté par l'inégalité de Markov n'a aucun intérêt.

D'un autre côté, l'inégalité de Markov présente une précision très large, surtout si la loi de probabilité de X est connue, en effet si on considère la loi $\mathcal{E}(1)$ et a = 2, on a :

E(X) = 1 et la borne de l'inégalité de Markov est 1/2, qui est environ 3.7 fois supérieure à P(X>2) = e^{-2} = 0.135

Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire de moyenne μ et de variance σ^2, alors :

$$\forall \text{ k} > 0, \text{P}(|\text{X} - \mu| \geq k) \leq \left(\frac{\sigma}{k}\right)^2$$

Démonstration :

On utilise l'inégalité de Markov appliquée à la variable aléatoire $(X - \mu)^2$

$\forall$ k > 0, $\text{P}((\text{X} - \mu)^2 > k^2) \leq \frac{E(X-\mu)^2}{k^2}$

Or l'événement $(X-\mu)^2 > k^2 \equiv |X-\mu| > k$

Donc $\forall$ k > 0, $\text{P}(|X-\mu| > k) \leq \frac{E(X-\mu)^2}{k^2} = \left(\frac{\sigma}{k}\right)^2$

Inégalité de Tchebychev unilatérale

Soit X une variable aléatoire d'espérance nulle et de variance σ^2. On a alors pour tout réel a>0 : $P(X \geq a) \leq \frac{\sigma^2}{a^2+\sigma^2}$

Démonstration :

Soit b > 0, on a $X \geq a \equiv X + b \geq a + b$ et $P(X + b \geq a + b) \leq P(|X + b| \geq a + b)$
Or $P(|X + b| \geq a + b) = P((X + b)^2 \geq (a + b)^2)$
Appliquons l'inégalité de Markov à la V A positive $(X + b)^2$:

$$P\left((X+b)^2 \geq (a+b)^2\right) \leq \frac{E(X+b)^2}{(a+b)^2} = \frac{\sigma^2 + b^2}{(a+b)^2}$$

On peut facilement vérifier que la grandeur $\frac{\sigma^2 + b^2}{(a+b)^2}$ est minimale si $b = \sigma^2/a$ et ce minimum est égale à $\frac{\sigma^2}{a^2 + \sigma^2}$

Exemple d'utilisation des inégalités de Markov et Tchebychev :
On suppose que le nombre de pièces sortant d'une usine en l'espace d'une semaine, est une variable aléatoire d'espérance 50.
a) Que peut-on dire de la probabilité que la production hebdomadaire dépasse 75 pièces ?
b) On sait de plus que la variance de la production hebdomadaire est égale à 25, quelle est la probabilité que la production hebdomadaire soit comprise entre 40 et 60 pièces ?

Solution : Désignons par X le nombre de pièces produites en une semaine
a) L'inégalité de Markov donne : $P(X > 75) \leq \frac{50}{75} = 0.66$
b) Si de plus la variance est connue, on peut appliquer l'inégalité de Tchebychev :
$40 \leq X \leq 60 \equiv -10 \leq X - 50 \leq 10 \equiv |X - 50| \leq 10$
$P(|X - 50| \geq 10) \leq \frac{25}{100} = 0.25$
Donc $P(|X - 50| \leq 10) \geq 1 - 0.24 = 0.75$.

En reprenant la première question, la connaissance de la variance permet d'améliorer la borne 0.66 trouvée, l'inégalité de Tchebychev symétrique donne :
$X > 75 \equiv X - 50 > 25$
$P(X - 50 > 25) \leq P(|X - 50| > 25) \leq \frac{25}{25^2} = \frac{1}{25}$
L'inégalité de Tchebychev unilatérale donne :
$P(X - 50 \geq 25) \leq \frac{25}{25 + 25^2} = \frac{1}{26}$
On remarque que cette dernière borne égale à $\frac{1}{26}$ est légèrement meilleure que la première borne égale à $\frac{1}{25}$

Inégalité de Jensen
Soit X une variable aléatoire et Φ une fonction définie et convexe sur $X(\Omega)$, alors si $E(X)$ et $E(\Phi(X))$ existent alors on a : $E(\Phi(X)) \geq \Phi(E(X))$

Démonstration :
Considérons d'abord le cas où X est une V A de densité f_X.
Soit $(x, y) \in X(\Omega)^2$ tel que $E(X) \in]x, y[$, on a :

$$\frac{\Phi(x)-\Phi(E(X))}{x-E(X)} \leq \Phi_g^{'}(E(X)) \leq \frac{\Phi(y)-\Phi(E(X))}{y-E(X)}$$

La première inégalité implique que : $\Phi(x)-\Phi(E(X)) \geq (x-E(X))\Phi_g^{'}(E(X))$

La deuxième inégalité implique que : $\Phi(y)-\Phi(E(X)) \geq (y-E(X))\Phi_g^{'}(E(X))$

On déduit alors que $\forall x \in X(\Omega),\ \ \mathrm{h(x)} = \Phi(x)-\Phi(E(X)) \geq (x-E(X))\Phi_g^{'}(E(X)) = g(x)$,

Autrement : $\forall x \in X(\Omega),\ \ \mathrm{h(x)} \geq g(x) \Rightarrow \mathrm{E}(\mathrm{h}(\mathrm{X})) \geq E(g(X))$

$$\Rightarrow \mathrm{E}(\Phi(X))-\Phi(E(X))) \geq E((X-E(X))\Phi_g^{'}(E(X)))$$

On obtient alors le résultat recherché.

L'inégalité de Jensen peut être démontrée d'une autre manière, mais en supposant que la fonction Φ est deux fois dérivable pour appliquer la formule de Taylor-Lagrange d'ordre 2 au voisinage de E(X) :

$$E(\Phi(X)) = \int_R \Phi(x) f_X(x)dx$$

$$= \int_{\mathrm{R}} \left(\Phi(E(X)) + (x-E(X))\Phi'(E(X)) + (x-E(X))^2\Phi''(\theta_x)\right) f_X(x)dx,$$

θ_x étant compris entre x et E(X)

$$E(\Phi(X)) = \int_{\mathrm{R}} \Phi(E(X)) f_X(x)dx + \underbrace{\int_R (x-E(X))\Phi'(E(X)) f_X(x)dx}_{0} + \underbrace{\int_R \frac{(x-E(X))^2}{2!}\Phi''(\theta_x) f_X(x)dx}_{\geq 0}$$

$$\geq \Phi(\mathrm{E(X)})$$

Φ'' étant positive car la fonction Φ est convexe.

Dans le cas où X est une V A discrète, il s'agit de démontrer que :

$$\sum_i \Phi(x_i)P(X=x_i) \geq \Phi\left(\sum_i x_i P(X=x_i)\right)$$

Nous allons faire tout simplement une démonstration par récurrence sur i,

Si i = 2, alors $\Phi(x_1P(X=x_1) + x_2P(X=x_2)) \leq \Phi(x_1)P(X=x_1) + \Phi(x_2)P(X=x_2)$

Supposons que la propriété est vérifiée pour i = n,

$\Phi(x_1P(X=x_1) + x_2P(X=x_2) + \ldots + x_nP(X=x_n) + x_{n+1}P(X=x_{n+1}))$

$= \Phi(x_1p_1 + x_2p_2 + \ldots + x_np_n + x_{n+1}p_{n+1})$

$$= \Phi((p_1+p_2)(\frac{p_1}{p_1+p_2}x_1 + \frac{p_2}{p_1+p_2}x_2) + \ldots + x_np_n + x_{n+1}p_{n+1})$$

$$= \Phi((p_1+p_2)y + \ldots + x_np_n + x_{n+1}p_{n+1}) \leq (p_1+p_2)\Phi(y) + \ldots + p_n\Phi(x_n) + p_{n+1}\Phi(x_{n+1})$$

avec $y = \dfrac{p_1}{p_1+p_2}x_1 + \dfrac{p_2}{p_1+p_2}x_2$

Finalement :

$\Phi(x_1p_1 + x_2p_2 + \ldots + x_np_n + x_{n+1}p_{n+1}) \leq p_1\Phi(x_1) + p_2\Phi(x_2) + \ldots + p_n\Phi(x_n) + p_{n+1}\Phi(x_{n+1})$

Exemple d'application de l'inégalité de Jensen

Considérons la fonction $\Phi(x) = x^2$, qui est une fonction convexe sur R et X une variable aléatoire. L'application de l'inégalité de Jensen permet de justifier un résultat bien connu : $E(X)^2 \leq E(X^2)$ qui traduit que la variance $Var(X) = E(X^2) - E(X)^2$ est toujours positive.

Le comportement asymptotique des suites de variables aléatoires

Il s'agit d'étudier le comportement des suites de variables aléatoires au voisinage de l'infini. Ce comportement peut-être vu de plusieurs manières, si on considère une suite de variables aléatoires X_n, $n \geq 1$, il est possible de s'intéresser à la probabilité de l'écart entre X_n et sa limite, à la limite d'un moment centré de X_n ou la loi de la variable limite.

Convergence en probabilité d'une variable aléatoire

Définition : soit X_n, $n \geq 1$, une suite de variables aléatoires, on dit que X_n converge en probabilité vers X si et seulement si $\forall \varepsilon > 0,\ \lim\limits_{n\to+\infty} P(|X_n - X| > \varepsilon) = 0$

On note $X_n \underset{n\to+\infty}{\overset{P}{\to}} X$

Cette notion de convergence peut s'interpréter de la manière suivante : Pour tout écart ε fixé, lorsque n devient très grand, il est de plus en plus peu probable d'observer un écart, supérieur à l'écart donné, entre X et X_n.

Proposition :

Soit X_n, $n \geq 1$, une suite de variables aléatoires.

Si $\lim\limits_{n\to+\infty} E(X_n) = a$ et $\lim\limits_{n\to+\infty} Var(X_n) = 0$ alors $X_n \underset{n\to+\infty}{\overset{P}{\to}} a$

Démonstration :

On a $\forall n \geq 1, |X_n - a| \leq |X_n - E(X_n)| + |E(X_n) - a|$

Soit ε > 0, ∃ N ∈ N, ∀ n > N, $|E(X_n) - a| < \varepsilon/2$

Donc ∀ n > N, on a $|X_n - a| > \varepsilon \Rightarrow |X_n - E(X_n)| > \varepsilon/2$

Il s'en suit que ∀ n > N, on a $P(|X_n - a| > \varepsilon) \leq P(|X_n - E(X_n)| > \varepsilon/2)$

Par application de l'inégalité de Bienaymé-Tchebychev, on a :

∀ n > N, on a $P(|X_n - E(X)| > \varepsilon/2) \leq 4\dfrac{Var(X_n)}{\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$

D'où ∀ ε > 0, ∃ N ∈ N, ∀ n > N, $P(|X_n - a| > \varepsilon) \xrightarrow[n\to+\infty]{} 0$

Comme conséquence de cette proposition, pour que $X_n \underset{n\to+\infty}{\overset{P}{\to}} X$, il suffit que :

$\lim\limits_{n\to+\infty} E(X_n - X) = 0$ et $\lim\limits_{n\to+\infty} Var(X_n - X) = 0$

Pour cela, on considère $Y_n = X_n - X \underset{n\to+\infty}{\overset{P}{\longrightarrow}} 0$

Convergence en moyenne quadratique

Définition : Une suite de V A, X_n, $n \geq 1$, converge en moyenne quadratique vers la V A, X si :

$$\lim_{n\to+\infty} E(X_n - X)^2 = 0$$

On note $X_n \underset{n\to+\infty}{\overset{m.q}{\to}} X$

Propriétés :

1) La convergence en moyenne quadratique entraine la convergence en probabilité.

2) Soit X_n, $n \geq 1$, une suite de variables aléatoires, si $\lim\limits_{n\to+\infty} E(X_n) = a$ et $\lim\limits_{n\to+\infty} Var(X_n) = 0$ alors

$$X_n \underset{n\to+\infty}{\overset{m.q}{\to}} a$$

Démonstration :

1) Soit $\varepsilon > 0$ $\forall$ $n \geq 1$, l'événement $|X_n - X| > \varepsilon \equiv |X_n - X|^2 > \varepsilon^2$

Donc $\forall$ $n \geq 1$, $P(|X_n - X| > \varepsilon) = P(|X_n - X|^2 > \varepsilon^2)$

$(X_n - X)^2$ est une V A positive, par application de l'inégalité de Markov on a :

$$\forall\, n \geq 1,\ P(|X_n - X|^2 > \varepsilon^2) \leq \frac{E(X_n - X)^2}{\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$$

2) $\lim\limits_{n\to+\infty} E(X_n - a)^2 = \lim\limits_{n\to+\infty} E(X_n^2) - 2a \lim\limits_{n\to+\infty} E(X_n) + a^2 = \lim\limits_{n\to+\infty} E(X_n^2) - \lim\limits_{n\to+\infty} E(X_n)^2$

$= \lim\limits_{n\to+\infty} Var(X_n) = 0$

Convergence en loi

Définition : Soient X_n et X des variables aléatoires de fonctions de répartition respectives F_n et F, on dit que la suite X_n converge en loi vers la variable aléatoire X et on note $X_n \underset{n\to=\infty}{\overset{L}{\longrightarrow}} X$, si en tout point x ou F est continue, la suite $F_n(x)$ converge vers F(x).

Propriétés :

1) Si X_n et X sont des V A discrètes, alors X_n converge en loi vers X si et seulement si :

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

2) La convergence en probabilité entraine la convergence en loi, $X_n \underset{n\to+\infty}{\overset{P}{\to}} X \Rightarrow X_n \underset{n\to=\infty}{\overset{L}{\longrightarrow}} X$

Preuve :

1) Supposons que pour tout $k \in \mathbb{N}$, $\lim\limits_{n\to+\infty} P(X_n = k) = P(X = k)$

Pour tout x de $\mathbb{R}^+$ où F est continue, on a un nombre fini d'entiers naturels k inférieurs à x, donc $F_n(x) = P(X_n \leq x) = \sum\limits_{k\leq x} P(X_n = k)$

Or $\lim\limits_{n\to+\infty} \sum\limits_{k\leq x} P(X_n = k) = \sum\limits_{k\leq x} P(X = k) = P(X \leq x) = F(x)$

De plus si $x<0$, $F_n(x) = 0$ donc $\lim\limits_{n\to+\infty} F_n(x) = F(x)$

Donc $X_n \underset{n\to=\infty}{\overset{L}{\longrightarrow}} X$

Supposons que X_n converge en loi vers X

Pour tout $k \in \mathbb{N}$, $k + 1/2$ et $k - 1/2$ sont des points de continuité de la fonction F et :

$P(X_n = k) = P(k - 1/2 < X_n < k + 1/2) = F_n(k + 1/2) - F_n(k - 1/2)$

Donc $\lim\limits_{n\to+\infty} P(X_n = k) = F(k + 1/2) - F(k - 1/2) = P(X = k)$

Nous avons déjà établi des exemples de convergence en loi :

La convergence de la loi Binomiale B(n, p) vers la loi de Poisson P(np) et la convergence de la loi Hypergéométrique $\mathcal{H}$(N,n,p) vers la loi Binomiale B(n,p).

2) Commençons tout d'abord par établir l'inégalité suivante :
A et B sont deux V A, c est un réel et $\varepsilon > 0$

$P(A \leq c) = P(A \leq c, B \leq c + \varepsilon) + P(A \leq c, B > c + \varepsilon)$
$\quad = P(A \leq c / B \leq c + \varepsilon)P(B \leq c + \varepsilon) + P(A \leq c, B - \varepsilon > c)$
$\quad \leq P(B \leq c + \varepsilon) + P(A \leq c, B - \varepsilon > c)$

On a $\begin{cases} A \leq c \\ B - \varepsilon > c \end{cases} \Leftrightarrow \begin{cases} A \leq c \\ \varepsilon - B < -c \end{cases} \Rightarrow A - B < -\varepsilon$

Donc $P(A \leq c, B - \varepsilon > c) \leq P(A - B < -\varepsilon) \leq P(A - B < -\varepsilon) + P(A - B > \varepsilon) = P(|A - B| > \varepsilon)$

Finalement on a : $P(A \leq c) \leq P(B \leq c + \varepsilon) + P(|A - B| > \varepsilon)$

Si on pose A = X_n, B = X et c = x, on a : $P(X_n \leq x) \leq P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon)$ (a)

Si on pose A = X, B = X_n et c = x – ε, on a : $P(X \leq x - \varepsilon) \leq P(X_n \leq x) + P(|X_n - X| > \varepsilon)$

Ce qui implique que : $P(X_n \leq x) \geq P(X \leq x - \varepsilon) - P(|X_n - X| > \varepsilon)$ (b)

Soit x un réel ou F est continu, soit $\eta > 0$, $\exists\ \varepsilon > 0$ tel que :

$$|F(x+\varepsilon) - F(x)| < \eta / 2 \text{ et } |F(x-\varepsilon) - F(x)| < \eta / 2$$

Puisque que F est croissante alors : $F(x+\varepsilon) - F(x) < \eta / 2$ et $F(x-\varepsilon) - F(x) > -\eta / 2$

La convergence en probabilité de X_n vers X implique que :

$$\exists\, n_0,\ \forall\, n > n_0\ P(|X_n - X| > \varepsilon) < \eta / 2$$

L'inégalité (a) implique que : $\forall\, n > n_0$

$$F_n(x) - F(x) \leq F(x + \varepsilon) - F(x) + P(|X_n - X| > \varepsilon) < \eta/2 + \eta/2 = \eta$$

L'inégalité (b) implique que : $\forall\, n > n_0$

$$F_n(x) - F(x) \geq F(x - \varepsilon) - F(x) - P(|X_n - X| > \varepsilon) > -\eta/2 - \eta/2 = -\eta$$

Finalement nous avons démontré que :
$\forall\, \eta > 0$, $\exists\, n_0$, $\forall\, n > n_0$, $|F_n(x) - F(x)| < \eta$ et ceci est valable pour chaque x ou la fonction F est continue.

La réciproque de cette propriété est fausse, comme le montre l'exemple suivant :
Les variables aléatoires X_n suivent une loi de Bernoulli B(1/2) : $P(X_n = 0) = P(X_n = 1) = 1/2$ et on a la relation $X = 1 - X_n$
On déduit alors que $P(X = 1) = P(X= 0) = 1/2$
L'écart entre X_n et X est constant est égale à 1, c'est pour cela que X_n ne converge pas en probabilité vers X.
La loi de X_n est la même que celle de X et $\lim_{n \to +\infty} P(X_n = i) = P(X = i)$, i = 1,2, ce qui implique la convergence en loi de X_n vers X

La convergence presque sûre
Définition : soit X_n, $n \geq 1$, une suite de variables aléatoires, on dit que X_n converge presque sûrement vers X si et seulement si $P\left(\lim_{n \to +\infty} X_n = X\right) = 1$

On note $X_n \underset{n \to +\infty}{\overset{p.s}{\to}} X$

Théorème

$X_n \underset{n\to+\infty}{\overset{p.s}{\to}} X \Leftrightarrow \forall\varepsilon>0,\ P\left(\limsup_{n\to+\infty}|X_n - X| > \varepsilon\right) = 0$

Démonstration :

$X_n \underset{n\to+\infty}{\overset{p.s}{\to}} X$

$\Leftrightarrow\ \forall\ \varepsilon > 0,\ \exists\ \mathrm{N} > 0,\ \forall\ \mathrm{n} > \mathrm{N},\ P(|X_n - X| < \varepsilon) = 1$

$\Leftrightarrow\ \forall\ \varepsilon > 0$, il existe alors une infinité d'indices n tels que $P(|X_n - X| < \varepsilon) = 1$

$\Leftrightarrow \forall\varepsilon>0,\ P\left(\limsup_{n\to+\infty}|X_n - X| > \varepsilon\right) = 0$

Ce théorème fournit une condition nécessaire et suffisante pour l'obtention de la convergence presque sûre d'une suite de variables aléatoires.

Théorème

A_n est une suite d'événements, on a : $\sum_{n=1}^{+\infty} P(A_n) < \infty \Rightarrow P\left(\limsup_{n\to+\infty} A_n\right) = 0$

Démonstration :

$P\left(\limsup_{n\to+\infty} A_n\right) = P\left(\bigcap_{n=1}^{+\infty}\bigcup_{p=n}^{+\infty} A_p\right) = P\left(\bigcap_{n=1}^{+\infty} B_n\right) = \lim_{n\to+\infty} P(B_n)$ car B_n étant une suite décroissante au sens de l'inclusion : $B_n \subset B_{n-1} \subset ... \subset B_1$.

$P(B_n) = P\left(\bigcup_{p=n}^{+\infty} A_p\right) \le \sum_{p=n}^{+\infty} P(A_p)$

Si la série de terme général $P(A_n)$, est convergente alors la somme partielle :

$$\sum_{p=n}^{+\infty} P(A_p) \xrightarrow[n\to+\infty]{} 0$$

Ce théorème nous montre qu'une condition suffisante pour que $P\left(\limsup_{n\to+\infty} A_n\right) = 0$, est que la série de terme général $P(A_n)$ soit convergente

Propriété : $X_n \underset{n\to+\infty}{\overset{p.s}{\to}} X \Leftrightarrow \forall\varepsilon>0, P\left(\sup_{p>n}|X_p - X| > \varepsilon\right) \xrightarrow[n\to+\infty]{} 0$

En effet : $X_n \underset{n\to+\infty}{\overset{p.s}{\to}} X \Leftrightarrow \forall\varepsilon>0,\ P\left(\limsup_{n\to+\infty}|X_n - X| > \varepsilon\right) = 0$

$\Leftrightarrow \forall\varepsilon>0,\ P\left(\bigcap_{n=1}^{+\infty}\bigcup_{p=n}^{+\infty}|X_p - X| > \varepsilon\right) = 0$

$\Leftrightarrow \forall\varepsilon>0,\ \lim_{n\to+\infty} P\left(\bigcup_{p=n}^{+\infty}|X_p - X| > \varepsilon\right) = 0$

Or l'événement $\bigcup_{p=n}^{+\infty} |X_p - X| > \varepsilon \;\equiv\; \sup_{p\geq n} |X_p - X| > \varepsilon$, pour cela, il suffit de vérifier que si le premier événement est vrai alors le deuxième l'est également et si le deuxième est vrai, le premier l'est également.

Donc $X_n \underset{n\to+\infty}{\overset{p.s}{\rightarrow}} X \Leftrightarrow \forall\varepsilon > 0,\ \lim_{n\to+\infty} P\left(\sup_{p\geq n} |X_p - X| > \varepsilon\right) = 0$

D'après cette propriété, nous avons l'équivalence suivante :

$$X_n \underset{n\to+\infty}{\overset{p.s}{\rightarrow}} X \Leftrightarrow \forall\varepsilon > 0,\ \lim_{n\to+\infty} P\left(\bigcup_{p=n}^{+\infty} |X_p - X| > \varepsilon\right) = 0$$

On pose $u_n = P\left(\bigcup_{p=n}^{+\infty} |X_p - X| > \varepsilon\right)$ et $v_n = P(|X_n - X| > \varepsilon)$

On a $v_n \leq u_n$, et les deux suites sont positives, pour tout $n \geq 1$, donc la convergence de u_n vers 0, implique celle de v_n vers 0.

Nous avons ainsi démontré que : $X_n \underset{n\to+\infty}{\overset{p.s}{\rightarrow}} X \Rightarrow X_n \underset{n\to+\infty}{\overset{P}{\rightarrow}} X$

Exemple de convergence presque sûre

On considère la suite X_n $n \geq 1$ définie à partir d'une V A, U de loi uniforme sur [0, 1] :

$$X_n = \begin{cases} \sqrt{n} & \text{si U} \leq \dfrac{1}{\text{n}} \\ 0 & \text{sinon} \end{cases}$$

Montrons que $X_n \underset{n\to+\infty}{\overset{p.s}{\rightarrow}} 0$

On a $\forall\, \varepsilon > 0$, l'événement $X_n > \varepsilon \equiv U \in [0, 1/n]$

$$\forall\, \varepsilon > 0,\ \bigcup_{p=n}^{+\infty} X_P > \varepsilon \equiv \bigcup_{p=n}^{+\infty} U \in \left[0, \frac{1}{p}\right] = U \in \left[0, \frac{1}{n}\right]$$

$$\forall\, \varepsilon > 0,\ P\left(\bigcup_{p=n}^{+\infty} X_P > \varepsilon\right) = P\left(U \in \left[0, \frac{1}{n}\right]\right) = \frac{1}{n} \xrightarrow[n\to+\infty]{} 0$$

Nous avons dans cet exemple la convergence presque sûre de X_n vers 0, alors que la série de terme général $P(X_n > \varepsilon) = \dfrac{1}{n}$ est divergente.

Lois des grands nombres

On considère une suite de V A, $(X_i)_{i\geq 1}$ indépendantes et identiquement distribuées (iid). Les lois des grands nombres portent sur le comportement de la moyenne empirique $\overline{X}_n = \dfrac{\sum_{i=1}^{n} X_i}{n}$ lorsque $n \to +\infty$.

Loi faible des grands nombres

On suppose que la variance de chaque X_i existe. Les V A X_i étant de même loi, elles admettent alors la même moyenne μ et la même variance σ^2

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

$$Var(\overline{X}) = Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{Var\left(\sum_{i=1}^{n} X_i\right)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$ car les variables aléatoires X_i sont indépendantes.

Proposition

Si $(X_i)_{i\geq 1}$ est une suite de variables aléatoires, iid, d'espérance μ et de variance σ^2, alors $\overline{X}_n \underset{n\to+\infty}{\overset{P}{\to}} \mu$: $\forall \varepsilon > 0, \mathrm{P}\left(\left|\overline{\mathrm{X}}_\mathrm{n} - \mu\right| > \varepsilon\right) \xrightarrow[n\to+\infty]{} 0$

Démonstration :

On utilise l'inégalité de Tchebychev pour la variable aléatoire $\overline{X}_n$

$$\forall \varepsilon > 0, \mathrm{P}\left(\left|\overline{\mathrm{X}}_\mathrm{n} - \mu\right| > \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$$

Application : Théorème de Bernoulli

On effectue n expériences successives indépendantes où on s'intéresse à chaque fois à la réalisation d'un certain événement A. On associe donc à chaque expérience i, $1 \leq i \leq n$, une variable de Bernoulli :

X_i	0	1
$P(X_i)$	1 - p	p

La fréquence empirique, c'est-à-dire le pourcentage de réalisation de A est :

$$F_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$E(F_n) = p$ et $var(F_n) = p(1-p)/n$

L'application de la loi faible des grands nombre implique que :

$$\forall \varepsilon > 0, \mathrm{P}\left(\left|\mathrm{F}_\mathrm{n} - p\right| > \varepsilon\right) \xrightarrow[n\to+\infty]{} 0$$

Comme interprétation de la convergence en probabilité de F_n vers p, il n'existe pas de valeur de n à partir de laquelle on soit sûr que chaque valeur de F_n soit à mois de ε de p.

Loi forte des grands nombres

Proposition :

Si $(X_i)_{i\geq 1}$ est une suite de variables aléatoires, iid, d'espérance μ et intégrable ($E(|X_i|)$ existe) alors $\overline{X}_n \underset{n\to+\infty}{\overset{p.s}{\to}} \mu$:

$\forall\ \varepsilon > 0,\ \exists\ \mathrm{N} > 0,\ \forall\ \mathrm{n} > \mathrm{N},\ P\left(\left|\overline{X}_n - \mu\right| > \varepsilon\right) = 0$

Comparaison entre la loi faible des grands nombres et la loi forte des grands nombres :
Dans la première loi, nous avons supposés l'existence des moments d'ordre 1 et 2 des V A et nous avons une convergence en probabilité. Dans la deuxième loi, nous avons supposé l'existence du moment d'ordre 1 et l'intégrabilité des V A et nous avons une convergence presque sûre qui implique la convergence en probabilité

Le théorème central limite

Le théorème central limite est l'un des résultats les plus importants de la théorie des probabilités. Il montre le caractère universel de la loi normale, en effet, la somme de n'importe quelle suite de V A, iid, converge en loi vers une loi normale.

Proposition (Théorème centrale limite)
Soit $(X_n)_{n\geq 1}$ une suite de V A, iid, de moyenne μ et de variance σ^2, alors la variable aléatoire
$Z_n = \dfrac{X_1 + X_2 + ... + X_n - n\mu}{\sigma\sqrt{n}} \underset{n\to+\infty}{\overset{L}{\to}} N(0,1)$, c'est-à-dire que : $\lim\limits_{n\to+\infty} P(Z_n < z) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$
Démonstration :
On se propose de calculer la transformée en s de la variable aléatoire Z_n et de montrer que celle-ci tend vers la transformée en s d'une loi normale centrée réduite lorsque n est grand.
On a $Z_n = \sum\limits_{i=1}^{n} \dfrac{Y_i}{\sqrt{n}}$, avec $Y_i = \dfrac{X_i - \mu}{\sigma}$
Les V A, Y_i constituent une suite iid, d'espérance 0 et de variance 1
On a pour une V A, Y : $f^T_{Y/\sqrt{n}}(s) = \int\limits_R e^{-sy} f_{Y/\sqrt{n}}(x)dx = f^T_Y\left(\dfrac{s}{\sqrt{n}}\right)$, car $f_{Y/\sqrt{n}}(y) = \sqrt{n} f_Y\left(\sqrt{n}y\right)$

Donc $f^T_Z(s) = \left(f^T_{Y/\sqrt{n}}(s)\right)^n = \left(f^T_Y\left(\dfrac{s}{\sqrt{n}}\right)\right)^n$
La V A Y, étant centrée réduite, alors :
$$\frac{df^T_Y(0)}{ds} = -E(Y) = 0 \text{ et } \frac{d^2 f^T_Y(0)}{ds^2} = Var(Y) + \left[\frac{df^T_Y(0)}{ds}\right]^2 = 1$$
La formule de Taylor à l'ordre 2 en 0, pour la fonction f^T_Y implique pour n assez grand que :
$$f^T_Y\left(\frac{s}{\sqrt{n}}\right) = f^T_Y(0) + \frac{s}{\sqrt{n}}\frac{df^T_Y(0)}{ds} + \frac{s^2}{2n}\frac{d^2 f^T_Y(0)}{ds^2} + \frac{s^2}{n}\varepsilon\left(\frac{s}{\sqrt{n}}\right), \text{ avec } \varepsilon\left(\frac{s}{\sqrt{n}}\right) \xrightarrow[\frac{s}{\sqrt{n}}\to 0]{} 0$$
$$f^T_Y(s) = 1 + \frac{s^2}{2n} + \frac{s^2}{n}\varepsilon\left(\frac{s}{\sqrt{n}}\right)$$
$f^T_Z\left(\dfrac{s}{\sqrt{n}}\right) = \left(1 + \dfrac{s^2}{2n} + \dfrac{s^2}{n}\varepsilon\left(\dfrac{s}{\sqrt{n}}\right)\right)^n = e^{n\log\left(1+\frac{s^2}{2n}+\frac{s^2}{n}\varepsilon\left(\frac{s}{\sqrt{n}}\right)\right)} \xrightarrow[n\to+\infty]{} e^{\frac{s^2}{2}}$ qui est la transformée en s d'une loi normale centrée réduite.

En pratique, le théorème central limite s'applique lorsque $n \geq 30$

Approximation de Moivre-Laplace

Cette approximation consiste à approcher une loi Binomiale B(n,p) qui est la somme de n lois de Bernoulli B(p) indépendantes, par une loi normale centrée réduite.

Proposition :
Soit S_n le nombre de succès lors de la réalisation de n épreuves indépendantes, la probabilité de réussite de chaque épreuve étant p. Alors pour tout a < b

$$\lim_{n\to+\infty} P\left(a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b \right) = \frac{1}{2\pi}\int_a^b e^{-\frac{t^2}{2}}dt$$

Cette approximation est applicable en général lorsque $\begin{cases} n \ge 30 \\ np \ge 15 \\ np(1-p) > 5 \end{cases}$

Correction de continuité : Dans cette approximation, il s'agit d'approcher S_n, une loi binomiale B(n,p), discrète par une loi normale $N\left(np, \sqrt{np(1-p)}\right)$, continue. Pour cela $P(S_n=k)$ est approchée par $P(k-0.5< S_n<k+0.5)$ si k ∈ {1, 2, …., n-1}. D'autre part, pour que $\sum_{k=0}^{n} P(S_n = k) = 1$, on approche $P(S_n = 0)$ par $P(S_n < 1/2)$ et $P(S_n = n)$ par $P(S_n > n - 1/2)$.

Exemple
Si X suit la loi B(40, 0,5), les calculs de probabilités concernant X peuvent être effectués en utilisant la loi N(20, $\sqrt{10}$)

En effet $\begin{cases} n \ge 30 \\ np = 20 \ge 15 \\ np(1-p) = 10 > 5 \end{cases}$ et $X^* = \dfrac{X-20}{\sqrt{10}}$ suit la loi N(0,1)

P(X = 20) = P(19.5<X<20.5) = $P\left(\dfrac{19.5-20}{\sqrt{10}} < X^* < \dfrac{20.5-20}{\sqrt{10}}\right) = 2\Phi\left(\dfrac{0.5}{\sqrt{10}}\right) - 1 = 0.1272$

Le résultat exact est P(X = 20) = $C_{40}^{20}(0.5)^{20}(0.5)^{20} = 0.1254$

$$P(17 \le X < 25) = P\left(\frac{16.5-20}{\sqrt{10}} < X^* < \frac{24.5-20}{\sqrt{10}}\right) = \Phi\left(\frac{4.5}{\sqrt{10}}\right) - \Phi\left(\frac{-3.5}{\sqrt{10}}\right)$$

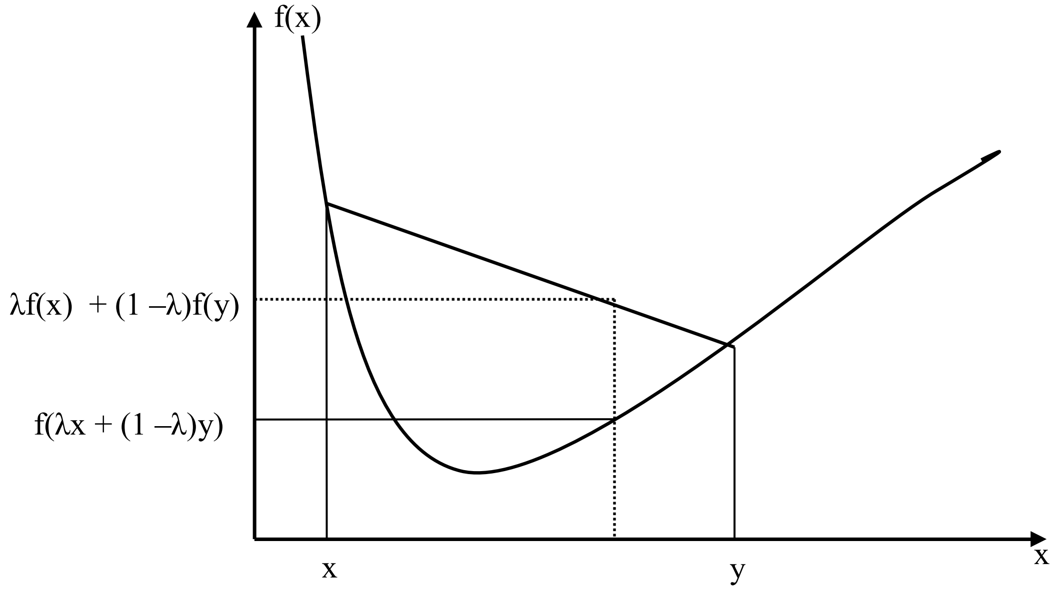
$$= \Phi\left(\frac{4.5}{\sqrt{10}}\right) + \Phi\left(\frac{3.5}{\sqrt{10}}\right) - 1 = 0.9222 + 0.8665 - 1 = 0.7887$$

Un calcul à la machine de $\sum_{k=17}^{24} C_{40}^{k} 0.5^k 0.5^{40-k} = 0{,}789$

Annexe

Propriétés d'une fonction convexe :

Définition
Soit f une fonction définie sur un intervalle I de R. On dit que f est convexe sur I si et seulement si : $\forall x, y \in I,\ \forall \lambda \in [0,1],\ f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$



Courbe représentative d'une fonction convexe

Une fonction est dite concave si son opposée est convexe.
Une fonction convexe sur un intervalle I, admet une dérivée à gauche et une dérivée à droite sur tout point x de I, avec $f_g'(x) \leq f_d'(x)$.
Pour tout $(x,y,z) \in I^3$ tel que $x < y < z$, on a :

$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} \leq \frac{f(z)-f(y)}{z-y}$$

Pour démontrer ces inégalités, on utilise la convexité de f et le fait que $y = \lambda x + (1-\lambda)z$, avec $\lambda \in]0, 1[$, en effet : $\lambda = \dfrac{z-y}{z-x}$

Si on fait une première fois tendre y vers x et une deuxième fois y vers z, on obtient :

$$f_d'(x) \leq \frac{f(z)-f(x)}{z-x} \leq f_g'(z)$$

Si la fonction f est deux fois dérivable sur I, alors f'' est positive.

La limite supérieure d'une suite d'événements

Soit $(A_n)_{n\geq 1}$ une suite d'événements, la limite supérieure de la suite A_n est définie par :

$$\limsup_{n\to+\infty} A_n = \bigcap_{n=1}^{+\infty}\bigcup_{p=n}^{+\infty} A_p = \bigcap_{n=1}^{+\infty} B_n \text{ avec } B_n = \bigcup_{p=n}^{+\infty} A_p$$

Théorème :

Pour que $\limsup_{n\to+\infty} A_n$ soit réalisé, il faut et il suffit qu'il existe une infinité d'indices p tels que A_p soit réalisé.

Preuve

$\limsup_{n\to+\infty} A_n$ est réalisé :

Donc il existe une infinité d'indices p tels que B_p soit réalisé et pour chaque B_p, il existe au moins un indice n tel que A_n soit réalisé.

Il existe une infinité d'indices p tels que A_p soit réalisé :

Donc $\forall$ n ≥ 1, $B_n = \bigcup_{p=n}^{+\infty} A_p$ est réalisé

Il s'en suit alors que $\bigcap_{n=1}^{+\infty} B_n = \limsup_{n\to+\infty} A_n$ est réalisé

Convergence d'une série

Soit s_n, n $\in$ N, une série de terme général u_n, n $\in$ N, s_n converge vers une limite l si et seulement si $\forall \varepsilon > 0$, $\exists$ N > 0, $\forall$ n > N, $|u_{n+1} + u_{n+2} + ...| < \varepsilon$

En effet : $s_n = \sum_{i=1}^{n} u_i$

$\lim_{n\to+\infty} s_n = l \Leftrightarrow \forall\varepsilon > 0, \exists \mathrm{N} > 0, \forall \mathrm{n} > \mathrm{N}, |\mathrm{s_n} - l| < \varepsilon$

$\Leftrightarrow \forall\varepsilon > 0, \exists \mathrm{N} > 0, \forall \mathrm{n} > \mathrm{N}, |\mathrm{u}_{n+1} + u_{n+2} + ...| < \varepsilon$

Exercices

Exercice 1

Soit X une variable aléatoire discrète, dont les valeurs appartiennent à N^*, si P(X = k) est décroissante, prouver que $P(X = k) \leq 2\frac{E(X)}{k^2}$

Soit X une variable aléatoire continue, positive et admettant une densité décroissante. Prouver que $f(x) \leq 2\frac{E(X)}{x^2}$ pour tout x > 0

Exercice 2

Soit X une variable aléatoire continue positive de moyenne 25. Que peut-on dire des espérances suivantes :

1. $E[X^3]$
2. $E\left[\sqrt{X}\right]$
3. E[logX]
4. $E[e^{-X}]$

Exercice 3

Un bruit X peut être considéré comme une variable aléatoire normale, centrée de variance σ^2, Supposons que toute valeur expérimentale de X cause une erreur dans un système de communication numérique si elle est plus grande que A.

1. Déterminer la probabilité qu'une valeur expérimentale de X cause une erreur si :
   (a) $\sigma^2 = 10^{-3}A^2$
   (b) $\sigma^2 = 10^{-1}A^2$
   (c) $\sigma^2 = A^2$
   (d) $\sigma^2 = 4A^2$
2. Pour une valeur donnée de A, quelle est la variance la plus grande telle que l'on obtienne une probabilité d'erreur pour toute valeur expérimentale de X inférieure à 10^{-3}, 10^{-6} ?

Exercice 4

Le poids d'un bretzel W (en grammes) est une variable aléatoire continue décrite par la densité : $f_W(w) = \begin{cases} 0 & w \leq 1 \\ w - 1 & 1 \leq w \leq 2 \\ 3 - w & 2 \leq w \leq 3 \\ 0 & w \geq 3 \end{cases}$

1. Quelle est la probabilité que 102 bretzels pèsent plus de 200 grammes ?
2. Si on choisit 4 bretzels indépendants, quelle est la probabilité qu'exactement 2 des 4 bretzels pèsent chacun plus de 2 grammes ?
3. Quel est le plus petit entier (non seulement ils sont immangeables, mais en plus on ne peut pas les casser) N pour lequel le poids total des N bretzels excède 200 grammes avec une probabilité de 0.99 ?

Exercice 5

Pour 3600 jets indépendants d'une pièce parfaite,

1. Déterminer un nombre n tel que la probabilité soit 0.5 pour que le nombre de piles résultant soit compris entre 1789 et n.
2. Déterminer la probabilité pour que le nombre de piles soit à 1 % de son espérance.

Exercice 6

Une pièce est lancée n fois. Chaque lancer est indépendant et suit une loi de Bernoulli de paramètre p. La variable aléatoire X est le nombre de piles obtenus. Pour chacune des expressions suivantes, trouver soit les valeurs de k et des constantes qui rendent la proposition vraie quel que soit $n \geq 1$, soit démontrer qu'aucune valeur de k existe.

1. $E\left[\frac{X}{n}\right] = Bn^k$
2. $E[(X - E[X])^2] = Cn^k$
3. $E[X^2] = Dn^k$
4. Pour n grand, $P\left[\left|\frac{X}{n} - p\right| \leq Fn^k\right] = 0.15$