



ÉCOLE CENTRALE PARIS

1^{ère} ANNÉE d'ÉTUDES

PROBABILITES ET STATISTIQUE

Thérèse PHAN



ÉCOLE CENTRALE PARIS

1^{ère} ANNÉE d'ÉTUDES

PROBABILITES

Thérèse PHAN

2004 - 2005

Réservé uniquement aux Enseignants, Elèves et Anciens Elèves de l'Ecole Centrale Paris

Reproduction interdite

Table des matières

Ch. I	Le modèle probabiliste.....	5
	I- 1) Introduction	5
	I- 2) Espace probabilisable	5
	I- 3) Espace probabilisé.....	6
	I- 4) Lois conditionnelles. Indépendance	8
Ch. II	Variables aléatoires réelles	11
	II- 1) Variable aléatoire réelle.....	11
	II- 2) Fonction de répartition	11
	II- 3) Variables continues	12
	II- 4) Variable Y fonction d'une variable X	12
	II- 5) Moments d'une variable aléatoire	13
Ch. III	Lois de probabilités discrètes.....	19
	III- 1) Introduction	19
	III- 2) Définitions, propriétés	19
	III- 3) Loi de Dirac.....	20
	III- 4) Loi uniforme.....	21
	III- 5) Loi de Bernoulli	21
	III- 6) Loi binomiale $B(n, p)$	22
	III- 7) Loi hypergéométrique	24
	III- 8) Loi de Poisson.....	26
	III- 9) Loi géométrique, loi de Pascal.....	27
	III- 10) Approximations	28
Ch. IV	Lois de probabilité continues.....	31
	IV- 1) Introduction	31
	IV- 2) Loi uniforme sur $[0, a]$	31
	IV- 3) Loi exponentielle	33
	IV- 4) Loi Gamma.....	34
	IV- 5) Loi de Laplace-Gauss ou loi normale.....	36
	IV- 6) Variables fonctions de variables normales.....	39
	IV- 7) Autres lois.....	40
Ch. V	Convolution, fonctions caractéristiques	43
	V- 1) Convolution	43
	V- 2) Fonctions caractéristiques.....	45
Ch. VI	Convergences de suites de variables	49
	VI- 1) Convergence en probabilité	49
	VI- 2) Convergence presque sûre	49
	VI- 3) Convergence en moyenne d'ordre p	50
	VI- 4) Convergence en loi.....	50
	VI- 5) Hiérarchie des convergences.....	51
	VI- 6) Théorème de De Moivre – Laplace	51
	VI- 7) Convergence de la loi de Poisson.....	53
	VI- 8) Lois des grands nombres	53
	VI- 9) Théorème de la limite centrale.....	54
	VI- 10) Résumé des approximations	55
Ch. VII	Couple de variables aléatoires	57
	VII- 1) Couples de variables aléatoires discrètes.....	57
	VII- 2) Variable continue et variable discrète.....	60
	VII- 3) Couple de variables aléatoires continues	61
	VII- 4) Coefficient et rapport de corrélation.....	62

Ch. VIII	Vecteurs aléatoires.....	67
	VIII- 1) Généralités	67
	VIII- 2) Espérance, Matrice de Covariance.....	68
	VIII- 3) Vecteurs gaussiens. Loi multivariée	69
	VIII- 4) Loi normale à deux dimensions.....	70
	VIII- 5) Loi du χ^2 (Chi-deux).....	71
	VIII- 6) Loi de Fisher-Snedecor	73
	VIII- 7) Loi de Student	74
Ch. IX	Processus aléatoires	77
	IX- 1) Généralités sur les processus aléatoires.....	77
	IX- 2) Processus ponctuel de Poisson	79
	IX- 3) Applications aux phénomènes d'attente.....	82

Introduction

Les Statistiques interviennent aujourd'hui dans tous les secteurs d'activité d'un ingénieur et doivent donc faire partie de ses connaissances de base. Un Centralien les utilise aussi bien en contrôle-qualité, qu'en fiabilité, qu'en étude de marché mais aussi en recherche médicale, en finance ou en politique.

« Le mot statistique désigne à la fois un ensemble de données d'observation et l'activité qui consiste dans leur recueil, leur traitement et leur interprétation » (E. Universalis).

Dans un exemple, nous étudions des relevés concernant 10 000 vols d'une grande compagnie aérienne : retards éventuels au départ ou à l'arrivée, durée de roulement avant le départ, durée de vol... Ce relevé constitue **une statistique**.

Faire de la statistique sur ces données consiste, par exemple, à :

- ✓ calculer la durée moyenne de vol (facile),
- ✓ relever les retards éventuels,
- ✓ prévoir le retard maximal dans certaines conditions (déjà plus complexe)...

Le groupe ou l'ensemble d'objets sur lequel on fait de la statistique s'appelle **la population** (dénomination héritée des premières applications de la statistique qui concernaient la démographie).

Les éléments de la population s'appellent **les individus**. Dans l'exemple que nous avons choisi, les individus sont les vols ayant eu lieu durant la période étudiée. L'étude de tous les individus d'une population finie s'appelle un **recensement**. Lorsqu'on ne s'intéresse qu'à une partie de la population, on parle de **sondage**, la partie étudiée s'appelant **échantillon**. On peut ainsi, dans notre exemple, ne s'intéresser qu'aux vols « Londres, New-York ».

Chaque individu est décrit par un ensemble de caractéristiques appelées **variables** ou caractères. Celles-ci peuvent être **quantitatives** (s'exprimant par des nombres réels) ou **qualitatives** (s'exprimant par l'appartenance à une catégorie ou à une modalité). La durée des vols est de nature quantitative, le type d'avion utilisé est de nature qualitative.

Les variables quantitatives sont de deux sortes :

- ✓ les **variables discrètes**, ne prenant qu'un nombre fini de valeurs,
- ✓ les **variables continues** prenant leurs valeurs dans un intervalle réel.

L'analyse statistique trouve sa justification dans la variabilité. Des individus en apparence identiques peuvent prendre des valeurs différentes : ainsi un certain type d'avion sur une ligne donnée aura des durées de vol différentes ; de même un processus industriel de fabrication ne fournira jamais des caractéristiques techniques parfaitement constantes.

L'étude de cette variabilité constitue la raison d'être essentielle de l'analyse statistique. On étudie cette variabilité (description), on cherche à la prévoir (estimation), à l'expliquer (modèles) ; on peut aussi chercher à la réduire, voire à l'augmenter.

Ce cours de Probabilités et Statistique est présenté en trois parties :

1. Calcul des probabilités

Dans cette partie, nous ferons un exposé simple de la théorie des probabilités, nous étudierons la notion de concept probabiliste, d'expérience aléatoire, de variables aléatoires, nous décrirons les principales lois de probabilité ainsi que le comportement asymptotique des variables aléatoires.

2. La statistique exploratoire

Elle consiste à recueillir des données, à les structurer et à les synthétiser sous forme de tableaux, de graphiques, de résumés numériques (statistique descriptive)... mais aussi à visualiser des données multidimensionnelles (analyse des données).

Nous ferons le point sur les diverses méthodes simples de statistique descriptive.

Nous verrons quelques exemples de méthodes factorielles appliquées à de grands tableaux de données, à l'aide de logiciels statistiques performants.

3. La statistique inférentielle

A l'aide des concepts de la théorie des probabilités, elle permet de savoir s'il est possible de :

- ✓ généraliser à toute la population des caractéristiques constatées sur un échantillon,
- ✓ confirmer (ou d'infirmer) des hypothèses formulées après une étude restreinte,
- ✓ prévoir les valeurs de certaines variables par modélisation.

Nous étudierons l'estimation ponctuelle, l'estimation par intervalle et les tests statistiques les plus usuels. A la fin de ce cours, seront abordées quelques méthodes simples de modélisation.

Ce premier polycopié est consacré à la première partie de ce cours : le Calcul des probabilités. Les deux parties suivantes, la Statistique exploratoire ainsi que la Statistique inférentielle, font l'objet d'un deuxième polycopié. Enfin un troisième polycopié réunit les tables statistiques les plus fréquemment utilisées.

Ch. I Le modèle probabiliste

I- 1) INTRODUCTION

Des phénomènes soumis à l'effet du hasard apparaissent à l'heure actuelle dans de nombreux domaines tels que les sciences sociales, la médecine, la production... Le terme de « probabilité », du latin « probare » signifiant prouver, est passé dans le langage courant et garde quelquefois une connotation liée à la chance. D'autres termes du calcul des probabilités montrent leur lien avec les jeux de hasard comme celui d'espérance mathématique qui correspond alors à l'espérance du gain (ou la perte...).

Les débuts d'une théorie des probabilités se situent au 17^{ème} siècle, en liaison avec l'étude des jeux de hasard ; les premiers résultats mathématiques sont introduits par Pascal et Fermat. Puis Huyghens, Bernoulli, de Moivre, Bayes, Laplace, Borel... apportent des concepts supplémentaires, voire différents.

Dans l'exposé de ce cours, nous choisirons l'approche axiomatique de Kolmogorov, théorie logique construite à partir d'axiomes. Elle repose sur des résultats mathématiques généraux, en particulier sur la théorie de la mesure.

I- 2) ESPACE PROBABILISABLE

I-2-1 Expérience aléatoire, événements aléatoires

Une **expérience** est qualifiée d'**aléatoire** si on ne peut pas prévoir par avance son résultat et si, répétée dans les mêmes conditions, elle peut donner des résultats différents.

Exemples :

- ✓ lancers répétés d'un dé,
- ✓ relevé des durées de vie d'un lot de 1000 lampes,
- ✓ mesure du diamètre d'une pièce mécanique...

Les résultats possibles de cette expérience constituent l'**ensemble fondamental Ω** appelé aussi univers des possibles.

Un **événement aléatoire** est une assertion relative au résultat de l'expérience.

A la réalisation d'un événement, on peut donc associer tous les résultats de l'épreuve correspondante et on identifie usuellement l'événement aléatoire et la partie de Ω pour laquelle cet événement est réalisé.

Exemple : dans le lancer simultané de deux dés, on s'intéresse aux chiffres inscrits sur les faces supérieures.

- ✓ L'ensemble fondamental est $\Omega = \{(1, 1), (1, 2), \dots, (6, 6)\}$. Il possède 36 éléments.
- ✓ « la somme des deux chiffres est inférieure ou égale à 5 » est un événement aléatoire possible ; la partie Ω' de Ω correspondante est :

$$\Omega' = \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2)\}.$$

I-2-2 Algèbre des événements

Nous allons nous référer à des propriétés ensemblistes usuelles :

- ✓ à tout événement A , on associe son contraire $\bar{A}$,
- ✓ aux événements A et B , on associe leur union $A \cup B$ ou leur intersection $A \cap B$,
- ✓ l'événement certain est représenté par Ω ,
- ✓ l'événement impossible est représenté par $\emptyset$.

C'est ainsi que nous sommes amenés aux définitions suivantes :

- On appelle **tribu de parties d'un ensemble Ω** un ensemble C de parties de Ω vérifiant :
 - ✓ $\Omega \in C$
 - ✓ $\forall A \in C, \bar{A} \in C$
 - ✓ Pour tout ensemble fini ou dénombrable de parties A_i de C , $\bigcup_{i \in I} A_i \in C$
- On appelle **espace probabilisable** un couple (Ω, C) où C est une **tribu** de parties de l'ensemble Ω .
- A et B sont deux **événements incompatibles** si la réalisation de l'un exclut celle de l'autre, c'est à dire si les parties A et B sont disjointes.
- $A_1, A_2, \dots, A_n$ forment un **système complets d'événements** si elles constituent une partition de Ω : elles sont disjointes deux à deux et leur réunion forme Ω tout entier.

I- 3) ESPACE PROBABILISE

I-3-1 Probabilité

I-3-1-1. Définitions

- On appelle **probabilité sur l'espace probabilisable (Ω, C)** une application P de C dans $[0, 1]$ telle que :
 - ✓ $P(\Omega) = 1$,
 - ✓ pour tout ensemble dénombrable $A_1, A_2, \dots, A_i, \dots$ d'événements incompatibles :

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i).$$
- Le triplet (Ω, C, P) est appelé **espace probabilisé**.

I-3-1-2. Remarque

Une loi de probabilité est simplement une mesure positive de masse totale 1 et la théorie des probabilités est partie de la théorie de la mesure.

I-3-2 Propriétés

Des axiomes on déduit :

- ✓ $P(\emptyset) = 0$
- ✓ $P(\overline{A}) = 1 - P(A)$
- ✓ $A \subset B \Rightarrow P(A) \leq P(B)$
- ✓ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- ✓ $P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i)$

I-3-3 Théorème des probabilités totales

Soit $(B_i)_i$ un système complet d'événements de Ω , alors :

$$\forall A \in \Omega \quad P(A) = \sum_i P(A \cap B_i)$$

I-3-4 Espaces probabilisés élémentaires

Considérons un espace Ω fini ou dénombrable. A tout élément a de Ω , on associe un nombre réel positif ou nul $P(a)$ tel que :

$$\sum_a P(a) = 1$$

A toute partie A de Ω on associe alors le nombre : $P(A) = \sum_{a \in A} P(a)$

On définit ainsi une probabilité sur la tribu des parties de Ω .

Soit Ω un ensemble fini. Supposons que l'on attribue le même poids à chaque événement élémentaire : $P(a) = \frac{1}{\text{card}(\Omega)}$

On définit une probabilité uniforme sur cet ensemble et la probabilité associée à une partie A quelconque de Ω est alors définie par :

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$$

I-3-5 Exemple

Reprenons le lancer de deux dés.

La **probabilité de l'événement** A décrit précédemment est : $p = \frac{9}{36} = \frac{1}{4}$.

L'événement aléatoire B : « au moins un des chiffres est égal à 5 » est **incompatible** avec l'événement A ...

I- 4) LOIS CONDITIONNELLES. INDEPENDANCE

Considérons deux événements A et B. Supposons qu'on ne s'intéresse à la réalisation de A que sous la condition que B soit déjà réalisé. Cela revient à rechercher la réalisation de $A \cap B$ par rapport à B.

I-4-1 Définition

Soit (Ω, C, P) un espace probabilisé et B un événement de probabilité non nulle.

On appelle **probabilité conditionnelle sachant B** l'application de C dans $[0, 1]$, notée $P(. / B)$, définie par :

$$\forall A \in C, \quad P(A / B) = \frac{P(A \cap B)}{P(B)}$$

Cette application définit une probabilité sur le même espace probabilisé.

I-4-2 Exemple

Dans le jeu du lancer de deux dés, on rappelle que A est l'événement élémentaire : « la somme des deux chiffres est inférieure ou égale à 5 ».

On définit l'événement C comme étant l'événement : « les deux chiffres sont pairs ».

Seul le couple (2, 2) satisfait aux deux exigences.

$$P(A \cap C) = 1/36 \quad P(A) = 9/36 \quad P(A / C) = 1/9.$$

I-4-3 Indépendance

- Définition :

L'événement A est **indépendant** de l'événement B si :

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

- Conséquences :

✓ Si A est indépendant de B, alors B est indépendant de A.

✓ A et B sont indépendants si et seulement si :

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

I-4-4 Indépendance mutuelle

Les événements $A_1, A_2, \dots, A_n$ sont dits **mutuellement indépendants** si, pour toute partie I de l'ensemble $\{1, 2, \dots, n\}$:

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

Cette propriété est plus forte que l'indépendance deux à deux qui en est une simple conséquence.

I-4-5 Formules de Bayes

La définition des probabilités conditionnelles nous permet d'écrire :

$$P(A \cap B) = P(A / B) \times P(B) = P(B / A) \times P(A)$$

On en déduit la première formule de Bayes :

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

Soit $(B_i)_i$ un système complet d'événements. Le théorème des probabilités totales s'écrit :

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A / B_i) \times P(B_i)$$

On en déduit la deuxième formule de Bayes :

$$P(B_i / A) = \frac{P(A / B_i) \times P(B_i)}{\sum_k [P(A / B_k) \times P(B_k)]}$$

Ch. II Variables aléatoires réelles

II- 1) VARIABLE ALEATOIRE REELLE

II-1-1 Définitions

Une application mesurable X de l'espace probabilisé (Ω, C, P) dans R muni de sa tribu de Borel (R, B) est appelée **Variable Aléatoire Réelle**.

On appelle **loi de probabilité de X** la mesure image de P par X ; on la note P_X . Elle est définie sur (R, B) par :

$$\forall B \in B, P_X(B) = P(\{a / X(a) \in B\}) = P[X^{-1}(B)]$$

On définit bien ainsi une probabilité sur (R, B) .

II-1-2 Exemple

Dans l'exemple du lancer des deux dés, on a vu que l'univers des possibles Ω est constitué des 36 couples. La probabilité d'obtenir un quelconque de ces couples est de $1/36$. Si on s'intéresse à la somme des deux chiffres obtenus, on définit une application X de Ω dans l'ensemble $E = \{2, 3, \dots, 12\}$. X est une variable aléatoire définie sur le triplet (Ω, C, P) où C est la tribu des parties de Ω . Pour obtenir la probabilité d'une valeur quelconque de E , il suffit de chercher la partie de C antécédent de cette valeur par X .

Ainsi : $P_X(6) = P[X^{-1}(6)] = P[\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}] = 5/36$.

II- 2) FONCTION DE REPARTITION

II-2-1 Définition :

On appelle **fonction de répartition** d'une variable aléatoire X l'application F de R dans $[0, 1]$ définie par :

$$F(x) = P(X < x)$$

II-2-2 Propriétés :

- ✓ F est monotone croissante, continue à gauche,
- ✓ Si la variable aléatoire est continue, la fonction de répartition est de plus continue à droite et dérivable,
- ✓ $F(-\infty) = 0$ et $F(+\infty) = 1$,
- ✓ $P(a \leq X < b) = F(b) - F(a)$ pour tout intervalle $[a, b]$ de R .
- ✓ Réciproquement toute fonction monotone croissante continue à gauche telle que $F(-\infty) = 0$ et $F(+\infty) = 1$ définit une loi de probabilité unique sur R .

II- 3) VARIABLES CONTINUES

On parle de variable continue dès que cette variable admet une densité de probabilité et réciproquement.

II-3-1 Densité de probabilité

Une loi de probabilité P_X admet une **densité** f si, pour tout intervalle I de $\mathbb{R}$:

$$P_X(I) = \int_I f(x) dx$$

F est alors dérivable et admet f pour dérivée. On a donc :

$$P(a \leq X < b) = \int_a^b f(x) dx = F(b) - F(a)$$

II-3-2 Propriétés

- ✓ Une densité est une fonction positive d'intégrale égale à 1 : $\int_{\mathbb{R}} f(x) dx = 1$
- ✓ $P(X = x) = 0$
- ✓ $P(x \leq X < x + dx) = f(x)dx$

II- 4) VARIABLE Y FONCTION D'UNE VARIABLE X

Soit X une variable aléatoire de densité f , de fonction de répartition F et soit Y une variable aléatoire fonction de $X : Y = \varphi(X)$.

Nous allons chercher la densité g et la fonction de répartition G de Y , en fonction de f et F .

II-4-1 Y est fonction bijective de X

II-4-1-1. Soit φ est croissante

$$G(y) = P(Y < y) = P(\varphi(X) < y) = P[X < \varphi^{-1}(y)] = F[\varphi^{-1}(y)]$$

En dérivant, on obtient :

$$g(y) = \frac{f(x)}{\varphi'(x)} = \frac{f[\varphi^{-1}(y)]}{\varphi'[\varphi^{-1}(y)]}$$

II-4-1-2. Soit φ est décroissante

$$G(y) = 1 - F[\varphi^{-1}(y)] \quad \text{et} \quad g(y) = \frac{f(x)}{|\varphi'(x)|}$$

$$g(y) = \frac{f[\varphi^{-1}(y)]}{|\varphi'[\varphi^{-1}(y)]|}$$

Ainsi, cette dernière formule est générale pour toute fonction φ bijective.

II-4-2 Y est fonction quelconque de X

On écrit alors la fonction de répartition $G(y)$ en recherchant pour X l'antécédent de l'événement $\{Y < y = \varphi(x)\}$.

Par exemple : Soit X une variable aléatoire réelle et soit $Y = X^2$. Pour tout y positif, on a :

$$G(y) = P(Y < y) = P(X^2 < y) = P(-\sqrt{y} < X < \sqrt{y}) = F(\sqrt{y}) - F(-\sqrt{y}).$$

II- 5) MOMENTS D'UNE VARIABLE ALEATOIRE

Une loi de probabilité se caractérise par certaines valeurs particulières liées aux notions de valeur centrale, de dispersion et de forme de la distribution.

II-5-1 Espérance mathématique

II-5-1-1. Définitions

X étant une variable aléatoire définie sur $(\Omega, \mathcal{C}, P)$, l'**espérance mathématique** de X , si elle existe, est l'intégrale de X par rapport à la mesure P :

$$E(X) = \int_{\Omega} X dP$$

Cette intégrale s'écrit aussi grâce à la mesure image : $E(X) = \int_R x dP_X(x)$

✓ Si X est une variable discrète, son espérance a pour expression :

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

✓ Si X est une variable continue de densité f son espérance est obtenue par l'intégrale suivante quand elle existe :

$$E(X) = \int_R x f(x) dx$$

✓ Si φ est une fonction réelle Lebesgue-mesurable de la variable X , on définit l'espérance mathématique de la variable $\varphi(X)$ par :

$$E[\varphi(X)] = \int_R \varphi(x) dP_X(x)$$

II-5-1-2. Propriétés

Pour tout a réel et tout couple de variables aléatoires (X, Y) :

✓ $E(a) = a$

✓ $E(aX) = aE(X)$

✓ $E(X + a) = E(X) + a$

✓ $E(X + Y) = E(X) + E(Y)$

II-5-1-3. Remarque

Certaines lois n'admettent pas d'espérance, comme, par exemple, la loi de Cauchy.

II-5-1-4. Espérance du produit de deux variables aléatoires

Elle est donnée par l'intégrale suivante, quand celle-ci a un sens :

$$E(X \times Y) = \int_{R^2} x \times y dP_{XY}(x, y)$$

Si les variables aléatoires sont indépendantes : $dP_{XY}(x, y) = dP_X(x) \otimes dP_Y(y)$ et l'intégrale double se factorise :

$$E(X \times Y) = \int_R x dP_X(x) \times \int_R y dP_Y(y)$$

On en déduit la propriété suivante :

Si X et Y sont indépendantes alors $E(X \times Y) = E(X) E(Y)$

Attention :

La réciproque est fausse : l'égalité $E(X \times Y) = E(X) E(Y)$ n'entraîne pas l'indépendance des variables X et Y.

II-5-1-5. Définition d'une variable centrée

Une variable aléatoire X est **centrée** si son espérance mathématique est nulle :

$$E(X) = 0$$

Si X est une variable aléatoire d'espérance mathématique m, la variable $(X - m)$ est centrée.

II-5-2 Variance d'une variable aléatoire**II-5-2-1. Définitions**

Soit X une variable aléatoire d'espérance m.

- On appelle **variance de X**, la quantité notée **V (X)** ou **Var (X)** ou encore **σ^2** :

$$V(X) = \sigma^2 = E[(X - m)^2] = \int_R (x - m)^2 dP_X(x)$$

- σ s'appelle l'**écart-type** de X.

II-5-2-2. Remarques

- ✓ La variance est le carré de l'écart-type.
- ✓ La variance est une mesure de la dispersion de X autour de m.

II-5-2-3. Expression usuelle de la variance

Soit X une variable aléatoire réelle ; sa variance, quand elle existe, s'écrit :

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

Cette propriété est très utilisée pour le calcul d'une variance. La démonstration est très simple et repose sur les propriétés de linéarité de l'espérance :

$$E[(X - m)^2] = E(X^2) - 2mE(X) + m^2 = E(X^2) - m^2$$

II-5-2-4. Propriétés

Soit X une variable aléatoire et a et b deux constantes réelles,

- $V(a) = 0$
- $V(X + a) = V(X)$
- $V(aX + b) = a^2 V(X)$

II-5-2-5. Inégalité de Bienaymé - Tchebychev

Soit X une variable aléatoire réelle d'espérance $E(X)$ et d'écart-type σ :

$$P(|X - E(X)| > k\sigma) \leq \frac{1}{k^2}$$

démonstration :

$$\sigma^2 = \int_R (x - m)^2 dP_X(x) \geq \int_{|X - m| > k\sigma} (x - m)^2 dP_X(x) \text{ par restriction du domaine d'intégration}$$

d'une fonction positive.

$$\sigma^2 > k^2 \sigma^2 \int_{|X - m| > k\sigma} dP_X(x) \text{ par minoration de } (x - m)^2 \text{ par } k^2 \sigma^2 \text{ à l'intérieur de l'intégrale.}$$

On en déduit : $\sigma^2 \geq k^2 \sigma^2 \times P(|X - E(X)| > k\sigma)$. D'où l'inégalité proposée.

Cette inégalité a pour intérêt qu'elle ne dépend pas de la loi de X . Dans la pratique, la majoration qu'elle fournit est trop grande et il est plus judicieux de se rapporter à la loi suivie par la variable quand on la connaît.

II-5-2-6. Définition

On appelle **variable centrée réduite** une variable dont l'espérance mathématique est nulle et dont la variance est égale à 1.

Soit X une variable d'espérance m et d'écart-type σ ; on appelle **variable centrée réduite** U

associée à la variable X , la variable :

$$U = \frac{X - m}{\sigma}$$

II-5-3 Covariance de deux variables aléatoires

II-5-3-1. Définition de la Covariance

On appelle covariance des deux variables X et Y, l'expression :

$$\text{Cov}(X, Y) = E[(X - E(X)) \times (Y - E(Y))] = E(X \times Y) - E(X) \times E(Y)$$

II-5-3-2. Variance d'une somme de variables aléatoires

Si on développe l'expression de la variance de X+Y, on obtient :

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

La variance de la somme des variables X et Y s'écrit alors :

$$V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$$

II-5-4 Indépendance de deux variables aléatoires

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

$$\text{Cov}(X, Y) = 0$$

$$V(X + Y) = V(X) + V(Y)$$

On a vu que si X et Y sont indépendantes, $E(X \times Y) = E(X) \times E(Y)$ donc $\text{Cov}(X, Y) = 0$. Toutefois, la covariance peut être nulle sans que les variables soient indépendantes.

II-5-5 Autres moments

On définit les moments centrés d'ordre k, quand ils existent par :

$$\mu_k = E[(X - m)^k]$$

On peut remarquer que $\mu_1 = 0$ et $\mu_2 = V(X)$.

Les moments μ_3 et μ_4 caractérisent la forme de la distribution à l'aide des coefficients suivants :

- coefficient d'asymétrie ou **skewness** : $\gamma_1 = \frac{\mu_3}{\sigma^3}$
- coefficient d'aplatissement ou **kurtosis** : $\gamma_2 = \frac{\mu_4}{\sigma^4}$

Ces coefficients sont sans dimension. Ils sont souvent calculés par les logiciels statistiques car leurs valeurs permettent de conforter le choix d'une loi à ajuster à une distribution donnée.

II-5-6 Fractiles

On appelle fractile d'ordre α de la variable X , le nombre x_0 tel que :

$$\Pr(X < x_0) = \alpha$$

Si α est égal à $\frac{1}{2}$, le fractile correspondant est la médiane.

Pour α égal à $\frac{1}{4}$ (respectivement $\frac{3}{4}$), x_0 est le premier quartile (respectivement le troisième quartile) ; la médiane correspond au deuxième quartile.

Pour α égal à 10%, 20%,... 100%, les fractiles s'appellent des déciles (premier décile, deuxième décile...)

Pour α égal à 1%, 2%...on définit les centiles.

Ch. III Lois de probabilités discrètes

III- 1) INTRODUCTION

En Statistique, la maîtrise des lois de probabilités, dans leur diversité, est essentielle pour l'analyse d'un ensemble de données. En effet, afin d'étudier un phénomène statistique il sera nécessaire, la plupart du temps, de faire appel à une loi de probabilité. La forme de sa représentation graphique (histogramme, diagramme...), les valeurs des diverses caractéristiques numériques (moyenne, médiane, écart-type...), le nombre de paramètres dont dépend le phénomène mais aussi la nature même de celui-ci permettent d'orienter le choix d'une loi de probabilité discrète ou continue pour l'étude des données. Nous allons donc consacrer deux chapitres à l'étude de ces diverses lois, le premier étant consacré aux lois discrètes.

III- 2) DEFINITIONS, PROPRIETES

III-2-1 Définition

Une variable aléatoire X est dite discrète si elle ne prend qu'un nombre fini ou dénombrable de valeurs $x_1, x_2, \dots, x_n \dots$

La loi de probabilité d'une telle variable est dite discrète également. Elle est caractérisée par :

- ✓ l'énumération des x_i
- ✓ les probabilités associées : $p_i = P(X = x_i)$

La fonction de répartition est une fonction en escalier constant sur tout intervalle $[x_i, x_{i+1}[$ et de saut, en chaque point x_i , la probabilité $p_{i+1} = P(X = x_{i+1})$.

Les lois discrètes sont utilisées pour la modélisation des jeux de hasard, l'analyse des sondages, l'étude des processus aléatoires tels que les files d'attente mais aussi en Biologie ou en Production.

III-2-2 Espérance et Variance

L'espérance mathématique et la variance sont données par :

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

$$Var(X) = \sum_{i=1}^n x_i^2 p_i - [E(X)]^2$$

Nous allons dans ce chapitre présenter les lois discrètes les plus couramment rencontrées. Auparavant nous allons revenir sur un exemple.

III-2-3 Exemple

Dans le jeu du lancer de deux dés, on a vu qu'on pouvait définir une variable aléatoire S en considérant la somme des deux chiffres inscrits sur les faces supérieures des dés ; elle peut prendre une des valeurs de $E = \{2, 3, \dots, 12\}$. Les probabilités associées sont simples à calculer et présentées dans le tableau suivant :

S	2	3	4	5	6	7	8	9	10	11	12
P(S=s)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

On calcule alors aisément l'espérance et la variance de la variable S de la façon suivante :

$$E(S) = (2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + \dots + 11 \times \frac{2}{36} + 12 \times \frac{1}{36}) = \frac{252}{36} = 7$$

$$V(S) = \frac{1}{36} (4 \times \frac{1}{36} + 9 \times \frac{2}{36} + 16 \times \frac{3}{36} + \dots + 144 \times \frac{1}{36}) - E(S)^2 = \frac{1974}{36} - 49 = 5,83$$

L'écart-type de cette distribution est égal à $\sigma = \sqrt{5,83} = 2,41$

Sa fonction de répartition est définie simplement par :

$$F(s) = P(S < s)$$

$$\text{Ainsi, } F(5) = P(S < 5) = P(2 \leq S \leq 4) = \frac{6}{36} = \frac{1}{6}.$$

III- 3) LOI DE DIRAC

III-3-1 Définition

On appelle loi de Dirac au point a la loi de probabilité δ_a définie sur $\mathbb{R}$ par :

$$\delta_a(a) = 1 \quad \delta_a(x) = 0 \quad \forall x \neq a$$

La loi de Dirac est la loi d'un phénomène de résultat a .

III-3-2 Moments

- $E(X) = a$

$$E(X) = a\delta_a(a) + 0 = a$$

- $V(X) = 0$

$$V(X) = E(X^2) - [E(X)^2] = a^2\delta_a(a) + 0 - a^2 = 0$$

III- 4) LOI UNIFORME

III-4-1 Définition

Soit X une variable aléatoire prenant chaque valeur de l'ensemble $\{1, 2, \dots, n\}$. La loi de X est dite uniforme sur $[1, n]$ si :

$$\forall k \in [1, n] \quad P(X = k) = \frac{1}{n}$$

III-4-2 Moments

- $E(X) = \frac{n+1}{2}$

en effet, la définition de l'espérance nous permet d'écrire : $E(X) = \frac{1}{n}(1 + 2 + \dots + n) = \frac{n(n+1)}{2n}$

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

$$E(X^2) = \frac{1}{n}(1 + 2^2 + \dots + n^2) = \frac{(n+1)(2n+1)}{6}$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{(n+1)(n-1)}{12}$$

III-4-3 Domaine d'application

Cette loi intervient dans l'étude de divers jeux (pile ou face, dés, jeu de cartes, loterie...) mais aussi dans les sondages...

III- 5) LOI DE BERNOULLI

III-5-1 Définition

C'est la loi d'une variable aléatoire X ne pouvant prendre que deux valeurs notées 1 ou 0, appelées communément succès ou échec. Soit p la probabilité que la variable X soit égale à 1, alors $1-p$ est la probabilité que X soit égale à 0 :

$$P(X = 1) = p \quad P(X = 0) = 1 - p$$

X peut être considérée comme la fonction indicatrice d'un événement A de probabilité p .

III-5-2 Moments

- $E(X) = p$

$$E(X) = 1 \times P(X = 1) + 0 \times P(X = 0) = p$$

- $V(X) = p(1-p)$

$$V(X) = 1^2 P(X = 1) + 0^2 P(X = 0) - p^2 = p - p^2$$

III-5-3 Domaine d'application

Cette loi, imagée par les jeux de pile ou face, permet de modéliser les matériels « en fonction » ou « à l'arrêt ».

III- 6) LOI BINOMIALE $B(n, p)$

Cette loi s'appelle aussi loi des tirages avec remise.

III-6-1 Définitions

III-6-1-1. Définition 1

Supposons que l'on répète n fois, dans des conditions identiques, une expérience aléatoire dont l'issue se traduit par l'apparition ou la non-apparition d'un événement A de probabilité p . Le résultat de la $i^{\text{ème}}$ expérience est noté X_i et la loi de X_i est la loi de Bernoulli de paramètre p . On suppose que chaque résultat est indépendant du précédent. On appelle X la variable aléatoire égale au nombre d'apparitions de A dans les n épreuves.

On dit alors que X suit une **loi binomiale de paramètres n et p , notée $B(n, p)$.**

III-6-1-2. Définition 2

De la définition précédente, on peut déduire une deuxième définition de la loi binomiale :

Soit X la somme de n variables X_i indépendantes, suivant toutes la loi de Bernoulli de paramètre p . La loi suivie par X est, par définition, la loi binomiale de paramètres n et p .

III-6-1-3. Expression de la probabilité

Pour rechercher l'expression de $P(X = k)$, on remarque tout d'abord que :

X prend la valeur k si k variables X_i prennent la valeur 1 et les $(n - k)$ autres variables prennent la valeur 0.

Il y a C_n^k façons de répartir les k valeurs 1 et les $n - k$ valeurs 0 comme résultats des n variables X_i .

De plus les X_i sont indépendantes :

$$P(X_1 = x_1 \cap \dots \cap X_n = x_n) = \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i}$$

Or $\sum_i x_i = k$

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

III-6-1-4. Remarques

- L'expression de $P(X = k)$ justifie la dénomination de loi binomiale ; en effet pour k variant de 1 à n , on retrouve les termes du développement du binôme : $[p + (1 - p)]^n$.
- Par définition de $P(X = k)$, on a bien : $\sum_{k=1}^n P(X = k) = 1$.
- Cette loi a été tabulée (voir polycopié de tables).

III-6-2 Moments

- $E(X) = np$

Ce résultat provient de la linéarité de l'espérance : $X = \sum_{i=1}^n X_i$ donc $E(X) = \sum_{i=1}^n E(X_i) = np$

- $V(X) = np(1 - p)$

$X = \sum_{i=1}^n X_i$ et les X_i sont indépendantes, aussi $V(X) = \sum_{i=1}^n V(X_i) = np(1 - p)$.

III-6-3 Propriétés

- Si X suit la loi binomiale $B(n, p)$, la variable $Y = n - X$ suit la loi $B(n, 1 - p)$.
- Par définition de la loi binomiale :

Si $L(X) = B(n_1, p)$, $L(Y) = B(n_2, p)$ et si X et Y sont indépendantes,
alors $L(X + Y) = B(n_1 + n_2, p)$

III-6-4 Exemple

Dans une usine de fabrication d'interrupteurs, il a été prouvé par divers contrôles qu'il y a 10% de défectueux. Un nouveau contrôle a lieu à l'aide d'un échantillon de taille 20. La loi suivie par la variable X représentant le nombre d'interrupteurs défectueux dans l'échantillon est une loi binomiale de paramètres $n = 20$ et $p = 0,1$. La probabilité d'observer exactement 2 défectueux dans l'échantillon est alors :

$$P(X = 2) = C_{20}^2 \times (0,1)^2 \times (0,9)^{18}$$

III-6-5 Domaines d'application

- Cette loi décrit des phénomènes formés de répétitions d'une même épreuve ayant deux possibilités s'excluant mutuellement : « pile ou face », arrêt / marche d'un matériel...
- Elle est utilisée en contrôle qualité pour déterminer la probabilité de défaillance d'un matériel technique. Pour l'appliquer, on suppose les expériences non exhaustives : les éléments contrôlés sont remplacés dans la population.

III- 7) LOI HYPERGEOMETRIQUE

On a vu que la loi binomiale était la loi des tirages avec remise ou tirage non exhaustif. On suppose, dans le cas de la loi binomiale, que les éléments sélectionnés sont remis dans la population de telle sorte que p demeure constant. En fait, cette hypothèse reste vraie pour un taux de sondage faible (grande population et population sélectionnée de petite taille).

Toutefois, en pratique, de nombreux tirages se font de nature exhaustive c'est à dire que les éléments sélectionnés ne sont pas remis dans la population. La composition de la population est alors modifiée au cours du tirage.

III-7-1 Définition

Dans une population de taille N , soit p la proportion d'individus possédant une propriété P_0 particulière. On prélève un échantillon de taille n dans cette population, le tirage s'effectuant sans remise.

La variable aléatoire X égale au nombre d'individus de l'échantillon possédant la propriété P_0 suit la loi hypergéométrique notée $H(N, n, p)$ définie par :

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

Le nombre d'échantillons de taille n dans la population est C_N^n .

Dans un quelconque échantillon de taille n , il y a :

- x individus possédant la propriété P_0 , choisis parmi les Np individus de la population entière possédant P_0 .
- $(n - x)$ autres individus ne la possédant pas ; ils sont choisis parmi les $(N - Np)$ individus de la population ne possédant pas P_0 .

$\frac{n}{N}$ est appelé **taux de sondage**.

Remarques :

- cette loi dépend de trois paramètres : N , n et p .
- Elle a été tabulée (voir polycopié de tables).

III-7-2 Propriété

X est la somme de n variables de Bernoulli, non indépendantes, correspondant aux tirages successifs des n individus.

Soit X_1 la variable aléatoire correspondant au tirage du premier individu. X_1 suit une loi de Bernoulli de paramètre p .

$$E(X_1) = p \text{ et } V(X_1) = p(1 - p).$$

Soit X_2 la variable aléatoire correspondant au tirage du deuxième individu. X_2 suit donc une loi de Bernoulli de paramètre inconnu. On détermine la valeur de ce paramètre en appliquant le théorème des probabilités totales :

$$P(X_2 = 1) = P(X_2 = 1 / X_1 = 1)P(X_1 = 1) + P(X_2 = 1 / X_1 = 0)P(X_1 = 0)$$

$$P(X_2 = 1) = \frac{Np-1}{N-1} p + \frac{Np}{N-1} (1-p) = p.$$

X_2 ne prend que les deux valeurs 1 et 0, on en déduit que :

$$P(X_2 = 0) = 1 - p.$$

On procède ainsi pour toutes les variables X_i . Chaque X_i suit la loi de Bernoulli de paramètre p . En revanche, elles ne sont pas indépendantes.

III-7-3 Moments

- $E(X) = np$

Ce résultat est une conséquence de la linéarité de l'espérance.

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

Les X_i n'étant pas indépendantes, la variance de X s'écrit :

$$V(X) = \sum_i V(X_i) + 2 \sum_{i \neq j} \text{cov}(X_i, X_j)$$

Les X_i étant des variables de Bernoulli, $V(X_i) = p(1-p)$

Par définition de la covariance :

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = P(X_i X_j = 1) - p^2$$

$$\text{Cov}(X_i, X_j) = P(X_i = 1 / X_j = 1)P(X_j = 1) - p^2 = \frac{Np-1}{N-1} p - p^2$$

$$V(X) = np(1-p) + n(n-1)[p \frac{Np-1}{N-1} - p^2] = np[1-p + (n-1) \frac{p-1}{N-1}] = np(1-p) \frac{N-n}{N-1}$$

$$\frac{N-n}{N-1} \text{ s'appelle le facteur d'exhaustivité.}$$

Il tend vers 1 quand N est très grand devant n .

On peut aussi calculer $E(X)$ et $V(X)$ en utilisant l'expression de $P(X=x)$.

III-7-4 Domaine d'application

Cette loi est utilisée dans le domaine du contrôle qualité dans le cas où les éléments défectueux sont retirés au fur et à mesure.

III- 8) LOI DE POISSON

III-8-1 Définition

C'est la loi d'une variable aléatoire discrète prenant toutes les valeurs entières positives ou nulles avec les probabilités associées, où λ est un paramètre réel :

$$\forall k \in \mathbb{N} \quad P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

On peut aisément vérifier que $\sum_k P(X = k) = 1$.

Comme pour les lois précédentes, il existe des tables donnant les valeurs des probabilités élémentaires et cumulées (voir polycopié de tables).

III-8-2 Moments

- $E(X) = \lambda$

$$E(X) = \sum_k k \frac{e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \lambda \sum_k \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

- $V(X) = \lambda$

$$V(X) = E(X^2) - \lambda^2 \quad \text{or} \quad E(X^2) = \sum_k k^2 \frac{e^{-\lambda} \lambda^k}{k!} = \sum_k k(k-1) \frac{e^{-\lambda} \lambda^k}{k!} + \sum_k k \frac{e^{-\lambda} \lambda^k}{k!} = \lambda^2 + \lambda$$

III-8-3 Remarque

La loi de Poisson est la seule loi discrète vérifiant $E(X) = V(X)$.

III-8-4 Propriété

Si X et Y sont deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres λ et μ , la variable $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.

La démonstration sera faite au chapitre V.

III-8-5 Processus de Poisson

Un processus ponctuel (voir chapitre IX) est dit **processus de Poisson** s'il représente une succession d'événements aléatoires vérifiant :

- Les temps d'attente entre deux événements consécutifs sont des variables indépendantes (processus sans mémoire).
- La loi du nombre d'événements arrivant dans l'intervalle $\{t, t+T\}$ ne dépend que de T .
- Deux événements ne peuvent arriver simultanément.

La probabilité d'obtenir n événements pendant un temps T est alors :

$$P(N = n) = \exp(-cT) \frac{(cT)^n}{n!}$$

$E(N) = cT$ où c est la cadence, nombre moyen d'événements par unité de temps.

III-8-6 Domaines d'application

La loi de Poisson est la loi des petites probabilités ou loi des événements rares.

On utilise cette loi pour décrire des événements tels que : les accidents dans un atelier de grande taille ou sur la route, les suicides, le nombre de pièces défectueuses dans une livraison importante et de bonne qualité. On la retrouve aussi dans les processus des files d'attente.

III- 9) LOI GEOMETRIQUE, LOI DE PASCAL

III-9-1 Loi géométrique

C'est la loi suivie par la variable X égale au nombre d'essais nécessaires pour faire apparaître, pour la première fois, un événement de probabilité p .

$$P(X = x) = p(1 - p)^{x-1} \quad \text{pour } x \geq 1$$

Son espérance et sa variance sont :

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

III-9-2 Loi de Pascal

Le nombre X d'essais nécessaires pour observer n fois un événement de probabilité p suit la loi de Pascal définie par :

$$P(X = x) = C_{x-1}^{n-1} p^n (1 - p)^{x-n} \quad \text{pour } x \geq n$$

On peut remarquer que, si X est le nombre d'essais nécessaires pour observer n fois cet événement particulier, alors durant les $(x-1)$ essais précédents il y a eu $(n-1)$ succès et la dernière observation est un succès.

Son espérance et sa variance se calculent aisément :

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

III- 10) APPROXIMATIONS

III-10-1 Définition

Une suite de variables aléatoires discrètes $(X_n)_n$ converge en loi vers une variable aléatoire discrète X si :

$$P(X_n = x) \rightarrow P(X = x) \text{ quand } n \rightarrow \infty$$

III-10-2 Approximation de la loi hypergéométrique par la loi binomiale

Soit X une variable aléatoire suivant la loi hypergéométrique $H(N, n, p)$. Quand N tend vers $+\infty$, la loi de X tend vers la loi binomiale $B(n, p)$.

$$P(X = x) = \frac{C_{Np}^x C_{N-Np}^{n-x}}{C_N^n} = C_n^x \frac{(Np)!}{(Np-x)!} \frac{(Nq)!}{(Nq-n+x)!} \frac{(N-n)!}{N!} \quad (q = 1 - p)$$

Pour N très grand : $\frac{(Np)!}{(Np-x)!} \approx (Np)^x$ car x est alors négligeable devant Np

de même : $\frac{(Nq)!}{(Nq-n+x)!} \approx (Nq)^{n-x}$ et $\frac{N!}{(N-n)!} \approx N^n$

$$P(X = x) \approx C_n^x \frac{(Np)^x (Nq)^{n-x}}{N^n} = C_n^x p^x q^{n-x}$$

En pratique, on utilise cette approximation dès que : $\frac{n}{N} \leq 10\%$

Cette approximation est très utilisée dans la pratique des sondages : un sondage sur un échantillon de 5 000 personnes sera aussi valable dans une ville de 500 000 habitants que dans une ville 10 fois plus grande.

III-10-3 Approximation de la loi binomiale par la loi de Poisson

Soit $(X_n)_n$ une suite de variables aléatoires binomiales $B(n, p)$ telles que :

$$n \rightarrow +\infty, \quad p \rightarrow 0 \quad \text{et} \quad n \times p \rightarrow \lambda, \quad (\lambda \text{ valeur finie}).$$

Alors la suite des variables $(X_n)_n$ converge en loi vers une variable de Poisson de paramètre λ .

$$C_n^x p^x (1-p)^{n-x} = \frac{n(n-1)\dots(n-x+1)}{x!} p^x (1-p)^{n-x} = \frac{(pn)^x}{x!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{x-1}{n}\right) (1-p)^{n-x}$$

Quand $n \rightarrow +\infty$, tous les termes $\left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{x-1}{n}\right)$ tendent vers 1, ainsi que leur produit.

De même, en décomposant : $(1-p)^{n-x} = (1-p)^n (1-p)^{-x}$ le deuxième terme du produit tend vers 1 quand p tend vers 0 et $(1-p)^n \approx (1-\frac{\lambda}{n})^n$ tend vers $\exp(-\lambda)$.

Finalement :
$$C_n^x p^x (1-p)^{n-x} \rightarrow \frac{\lambda^x}{x!} e^{-\lambda}$$

En pratique, on utilise cette approximation dès que : $n > 50$, $p < 0,1$, $n \times p$ restant fini.

Attention :

La loi binomiale prend ses valeurs dans l'intervalle $[0, n]$ alors que la loi de Poisson prend ses valeurs sur l'ensemble $\mathbb{N}$ tout entier. Il sera nécessaire d'en tenir compte dans l'utilisation de cette approximation.

Ch. IV Lois de probabilité continues

IV- 1) INTRODUCTION

IV-1-1 Densité de probabilité

On appelle **variable aléatoire continue X** une variable aléatoire pour laquelle il existe une fonction **f** non négative définie sur **R** telle que, pour toute partie **A** de **R** :

$$P(X \in A) = \int_A f(x) dx$$

Cette fonction **f** s'appelle **densité de probabilité** de la variable aléatoire **X**. L'ensemble des valeurs prises par une variable aléatoire continue est infini non dénombrable.

La fonction de répartition **F** de la variable **X** est alors définie par :

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

On a vu les propriétés de **f** et **F** au chapitre II.

IV-1-2 Espérance mathématique et variance

L'espérance et la variance d'une variable aléatoire continue **X** sont définies par :

$$E(X) = \int_R x f(x) dx$$

$$V(X) = \int_R x^2 f(x) dx - [E(X)]^2$$

Dans ce chapitre, nous allons étudier les principales lois continues.

IV- 2) LOI UNIFORME SUR [0, a]

IV-2-1 Définition

Une variable aléatoire réelle **X** suit une loi uniforme sur l'intervalle **[0, a]** si sa loi de probabilité admet pour densité :

$$f(x) = \frac{1}{a} \chi_{[0,a]}$$

où $\chi_{[0,a]}$ est la fonction caractéristique du segment **[0, a]**.

La densité s'écrit aussi :

$$f(x) = \frac{1}{a} \quad \text{pour } x \in [0, a] \quad f(x) = 0 \quad \text{ailleurs}$$

On calcule simplement sa fonction de répartition :

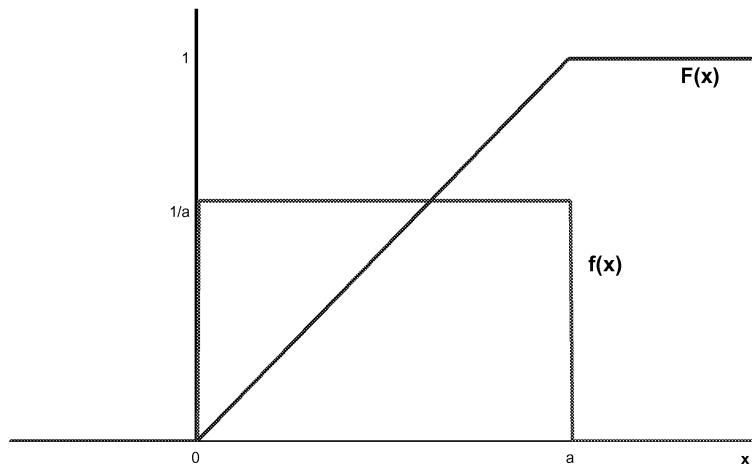
$$F(x) = \frac{x}{a} \quad x \in [0, a]$$

$$F(x) = 0 \quad x \in]-\infty, 0]$$

$$F(x) = 1 \quad x \in [a, +\infty[$$

IV-2-2 Courbes représentatives

Sur la représentation graphique suivante, figurent les courbes de f et F densité et fonction de répartition de cette variable uniforme sur $[0, a[$.



IV-2-3 Espérance et variance

- $E(X) = \frac{a}{2}$

en effet
$$E(X) = \int_0^a \frac{1}{a} x \, dx = \frac{1}{a} \frac{a^2}{2} = \frac{a}{2}$$

- $V(X) = \frac{a^2}{12}$

Par définition :
$$V(X) = \int_0^a \frac{1}{a} x^2 \, dx - \frac{a^2}{4} = \frac{a^2}{3} - \frac{a^2}{4} = \frac{a^2}{12}$$

IV-2-4 Domaine d'utilisation

La loi uniforme traduit l'hypothèse d'équirépartition sur l'intervalle.

Les tables associées sont les tables de nombres au hasard.

IV- 3) LOI EXPONENTIELLE

IV-3-1 Définition

Une variable aléatoire réelle positive suit une loi exponentielle de paramètre λ , positif, si sa densité de probabilité est définie par :

$$f(x) = \lambda \chi_{R^+}(x) \exp(-\lambda x)$$

où $\chi_{R^+}(x)$ est la fonction caractéristique de R^+ .

Sa fonction de répartition est égale à : $F(x) = \int_0^x \lambda \exp(-\lambda u) du = 1 - \exp(-\lambda x)$

$$F(x) = 1 - e^{-\lambda x} \text{ pour } x \geq 0$$

IV-3-2 Moments

- $E(X) = \frac{1}{\lambda}$

en effet $E(X) = \int_0^{+\infty} \lambda x \exp(-\lambda x) dx = \frac{1}{\lambda} \int_0^{+\infty} u \exp(-u) du = \frac{1}{\lambda}$

- $V(X) = \frac{1}{\lambda^2}$

$$V(X) = \int_{R^+} \lambda x^2 \exp(-\lambda x) dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

- Coefficient d'asymétrie $\gamma_1 = \frac{E[(X - E(X))^3]}{\sigma^3} = 2$

- Coefficient d'aplatissement $\gamma_2 = \frac{E[(X - E(X))^4]}{\sigma^4} = 9.$

- On peut remarquer que : $E(X) = \sigma = \frac{1}{\lambda}$

IV-3-3 Domaines d'application

Cette loi est utilisée en fiabilité : elle permet de modéliser la durée de vie de circuits électroniques ou de matériel subissant des défaillances brutales.

$E(X)$ est appelée MTBF (Mean Time Between Failure) et λ s'appelle le **taux de défaillance**.

IV- 4) LOI GAMMA

IV-4-1 Rappel

Γ , la fonction d'Euler, est définie par :

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

Ses principales propriétés sont :

- $\Gamma(r) = (r-1)\Gamma(r-1)$
- $\Gamma(n) = (n-1)!$ pour tout entier n positif.

IV-4-2 Définition de la loi Gamma

Une variable aléatoire réelle positive X suit une loi Gamma $\gamma(r, \lambda)$ de paramètres positifs r et λ , si sa densité de probabilité est donnée par :

$$f(x) = \frac{\lambda}{\Gamma(r)} \chi_{R^+}(x) e^{-\lambda x} (\lambda x)^{r-1}$$

On peut vérifier que f est bien une densité car elle est positive et son intégrale sur R^+ est égale à 1 par définition de la fonction Γ .

Les lois Gamma, avec r entier supérieur à 1, s'appellent aussi **lois d'Erlang**.

IV-4-3 Espérance et Variance de la loi Gamma

- $E(X) = \frac{r}{\lambda}$

$$E(X) = \frac{1}{\lambda \Gamma(r)} \int_0^{+\infty} u^r e^{-u} du = \frac{\Gamma(r+1)}{\lambda \Gamma(r)} = \frac{r}{\lambda}$$

- $V(X) = \frac{r}{\lambda^2}$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{\Gamma(r+2)}{\lambda^2 \Gamma(r)} - \frac{r^2}{\lambda^2} = \frac{r(r+1) - r^2}{\lambda^2} = \frac{r}{\lambda^2}$$

IV-4-4 Cas particulier $\lambda = 1$

Une variable aléatoire positive X suit une loi Gamma de paramètre r , notée γ_r , si sa densité est donnée par :

$$f(x) = \frac{1}{\Gamma(r)} \chi_{R^+}(x) e^{-x} x^{r-1}$$

- Si X suit une loi Gamma $\gamma(r, \lambda)$, et si λ différente de 1, la variable $Y = \lambda X$ suit la loi Gamma γ_r .

- On remarque que $E(X) = V(X)$, comme dans le cas de la loi de Poisson. Toutefois, la loi Gamma γ_r est continue et la loi de Poisson est discrète.

IV-4-5 Propriété

Si la variable aléatoire X suit une loi exponentielle de paramètre λ , la variable aléatoire λX suit la loi exponentielle de paramètre 1 qui est la loi γ_1 .

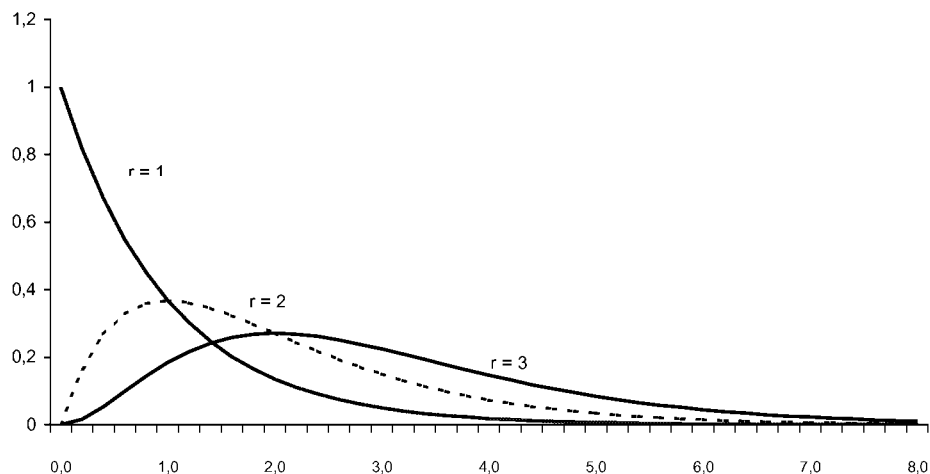
X suit une loi exponentielle de densité : $f(x) = \lambda \exp(-\lambda x)$.

Posons $Y = \lambda X$. Sa densité g vérifie $g(y)dy = f(x)dx = \exp(-y)dy$. $g(y) = \exp(-y)$

On retrouve la densité d'une loi exponentielle de paramètre 1. La loi exponentielle est un cas particulier de la loi Gamma.

IV-4-6 Courbes représentatives

Ci-dessous, figurent les courbes représentatives des densités de lois Gamma pour trois valeurs distinctes du paramètre r .



IV-4-7 Théorème

Si X et Y sont deux variables aléatoires indépendantes, suivant respectivement des lois γ_r et γ_s , alors la variable $X + Y$ suit une loi γ_{r+s} .

La démonstration sera faite dans le chapitre suivant traitant des fonctions caractéristiques.

IV-4-8 Domaines d'utilisation

- En fiabilité, la loi Gamma permet de modéliser les temps de défaillance de matériels.
- Dans la théorie des files d'attente, la loi Gamma représente la loi de probabilité d'arrivée de t événements dans un processus poissonnien.

IV- 5) LOI DE LAPLACE-GAUSS OU LOI NORMALE

Cette loi joue un rôle fondamental en Statistique : c'est un modèle très utilisé dans un grand nombre de domaines : fabrication industrielle, mesures... De plus, elle apparaît comme loi limite de certaines caractéristiques d'échantillons de grandes tailles.

IV-5-1 Définition

Une variable aléatoire réelle X suit une loi de Laplace - Gauss de paramètres m et σ si sa densité, définie sur $\mathbb{R}$, a pour expression :

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

Cette loi est notée : LG (m , σ) ou N(m , σ).

Sa fonction de répartition s'écrit alors :

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{1}{2}\left(\frac{t-m}{\sigma}\right)^2\right] dt$$

Elle n'a pas d'expression mathématique plus simple et a été tabulée pour $m = 0$ et $\sigma = 1$ (cf : poly de Tables).

IV-5-2 Espérance et variance

IV-5-2-1. Espérance et Variance d'une variable normale

Soit X une variable aléatoire suivant une loi de Laplace-Gauss LG(m , σ).

- $E(X) = m$

$$E(X) = \int_{\mathbb{R}} \frac{x}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] dx = \frac{\sigma}{\sqrt{2\pi}} \int_{\mathbb{R}} u \exp\left(-\frac{u^2}{2}\right) du + m \int_{\mathbb{R}} f(x) dx \quad \text{où } u = \frac{x-m}{\sigma}$$

La première intégrale est nulle (intégration sur $\mathbb{R}$ d'une fonction impaire) et la deuxième est égale à 1 donc on obtient bien le résultat annoncé. De la même façon on démontre que :

- $V(X) = \sigma^2$

Ces résultats expliquent le choix des paramètres caractérisant la loi.

IV-5-2-2. Espérance et variance d'une variable centrée réduite

On a vu que si une variable aléatoire suivait une loi d'espérance m et de variance σ^2 , la variable centrée réduite associée était la variable U :

$$U = \frac{X - m}{\sigma}$$

L'espérance mathématique d'une variable centrée réduite est égale à 0 et sa variance est égale à 1.

La fonction $U : X \rightarrow U(X)$ étant monotone croissante, on obtient la densité de la loi de U :

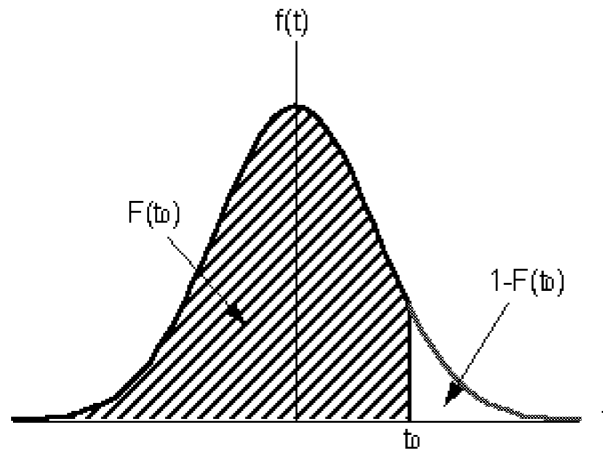
$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$$

La variable aléatoire U , centrée réduite, suit une loi de Laplace-Gauss LG (0, 1).

IV-5-3 Courbes représentatives

IV-5-3-1. Courbe représentative de la densité

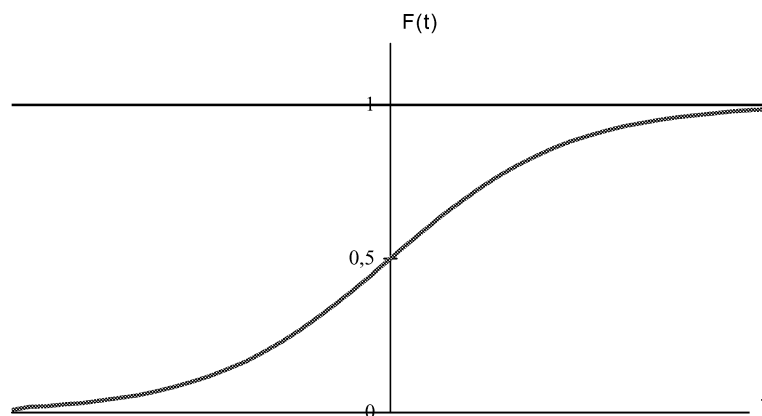
La courbe représentative de la densité f de la loi normale LG (0, 1) est :



De même, la courbe représentative de la densité de la loi de Laplace-Gauss LG (m , σ) présente une symétrie autour de l'axe $x = m$ et son étalement dépend de σ .

$F(t_0)$ représente, par définition de la fonction de répartition, l'aire de la surface hachurée. Sa courbe représentative est la suivante :

IV-5-3-2. Courbe représentative de la fonction de répartition



IV-5-4 Valeurs remarquables

En Statistique, on utilisera fréquemment certaines valeurs particulières de la fonction de répartition de cette loi et il est nécessaire de les connaître :

$$P\left(\left|\frac{X-m}{\sigma}\right| < 1,64\right) = 90\%$$

$$P\left(\left|\frac{X-m}{\sigma}\right| < 1,96\right) = 95\%$$

$$P\left(\left|\frac{X-m}{\sigma}\right| < 3,09\right) = 99,8\%$$

IV-5-5 Moments d'ordre supérieurs à deux

On peut rappeler qu'ils sont définis par :

$$\mu_k = E[(X-m)^k]$$

Dans le cas de la loi normale, ils existent pour tout ordre et on va les calculer pour une variable centrée réduite :

- ✓ Les moments d'ordre impair $k = 2n+1$ sont nuls (la fonction intégrée est impaire).
- ✓ Calculons les moments d'ordre pair :

$$\mu_{2n} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u^{2n} \exp\left(-\frac{1}{2}u^2\right) du = \frac{2}{\sqrt{2\pi}} \int_{\mathbb{R}^+} u^{2n} \exp\left(-\frac{1}{2}u^2\right) du$$

$$\text{On pose } t = \frac{u^2}{2} \quad dt = u du = \sqrt{2t} du$$

$$\mu_{2n} = \frac{2^n}{\sqrt{\pi}} \int_{\mathbb{R}^+} y^{n-\frac{1}{2}} \exp(-t) dt = \frac{2^n}{\sqrt{\pi}} \Gamma\left(n + \frac{1}{2}\right) = \frac{2^n}{\sqrt{\pi}} \times \frac{1 \times 3 \times 5 \dots \times (2n-1)}{2^n} \sqrt{\pi}$$

$$\mu_{2n} = \frac{2n!}{2^n n!}$$

Dans le cas d'une variable normale quelconque on a :

$$\mu_{2n} = \frac{(2n)!}{2^n n!} \sigma^{2n}$$

- Le coefficient d'asymétrie (Skewness) : $\gamma_1 = \frac{\mu_3}{\sigma^3} = 0$
- Le coefficient d'aplatissement (Kurtosis) : $\gamma_2 = \frac{\mu_4}{\sigma^4} = 3$

IV-5-6 Théorème

Soit X et Y deux variables aléatoires indépendantes suivant les lois $LG(m_1, \sigma_1)$ et $LG(m_2, \sigma_2)$, alors la variable aléatoire $X+Y$ suit la loi normale :

$$LG(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2}).$$

La démonstration de ce théorème sera faite dans le chapitre suivant.

IV- 6) VARIABLES FONCTIONS DE VARIABLES NORMALES

IV-6-1 Loi de la moyenne

Considérons un échantillon de n variables aléatoires X_i indépendantes, extrait d'une population X suivant la loi de Laplace-Gauss $LG(m, \sigma)$.

Posons $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

Cette variable est une combinaison linéaire de n variables aléatoires indépendantes suivant la loi $LG(m, \sigma)$. Elle suit elle-même une loi de Laplace-Gauss dont les paramètres sont :

$$E(\bar{X}) = m \qquad V(\bar{X}) = \frac{\sigma^2}{n}.$$

La variable $\bar{X}$, moyenne de n variables indépendantes extraites d'une population suivant la loi $LG(m, \sigma)$, suit la loi normale $LG(m, \frac{\sigma}{\sqrt{n}})$.

IV-6-2 Loi du carré d'une variable normale

Si U suit une loi de Laplace-Gauss $LG(0, 1)$, la variable $T = U^2$ a pour densité :

$$g(t) = \frac{f(\sqrt{t})}{\sqrt{t}} = \frac{1}{\sqrt{2\pi}} t^{-1/2} \exp(-\frac{t}{2}) \text{ pour } t > 0$$

Si U est une variable gaussienne centrée réduite, la variable $Z = \frac{U^2}{2}$ suit une loi Gamma $\gamma_{1/2}$.

IV- 7) AUTRES LOIS

IV-7-1 Loi log –normale

IV-7-1-1. Définition

X, variable aléatoire positive suit une loi log–normale si le logarithme népérien de X suit une loi normale.

Si la loi de (LnX) est la loi LG (m, σ), par changement de variables, on obtient la densité de la loi de X :

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

IV-7-1-2. Moments

- $E(X) = \exp\left(m + \frac{\sigma^2}{2}\right)$
- $V(X) = [\exp(2m + \sigma^2)][\exp(\sigma^2) - 1]$

IV-7-2 Loi de Weibull

IV-7-2-1. Définition

X suit une loi de Weibull à 3 paramètres (α , β , γ) si la variable associée $Y = \frac{X^\beta - \gamma}{\alpha}$ suit une loi exponentielle.

X représente en général une durée de vie. Elle est très utilisée en fiabilité des matériels.

On l'applique fréquemment avec $\gamma = 0$ et $\alpha = 1$, la densité est alors :

$$f(x) = \beta x^{\beta-1} \exp(-x^\beta) \text{ pour } x > 0$$

IV-7-2-2. Taux instantané de défaillance

Il est défini par :

$$\frac{P(x < X < x + dx / X > x)}{dx} = \frac{f(x)}{1 - F(x)} = \beta x^{\beta-1}$$

On retrouve les cas suivants pour un matériel donné :

- ✓ $\beta > 1$: matériel qui se dégrade avec le temps (usure)
- ✓ $\beta < 1$: matériel présentant des défauts de jeunesse.
- ✓ $\beta = 1$: matériel sans usure (loi exponentielle)

IV-7-3 Loi de Gumbel

IV-7-3-1. Définition

La variable aléatoire X suit la loi de Gumbel si la variable $Y = \exp(-X)$ suit la loi γ_1 .

La densité et la fonction de répartition de cette loi sont données par :

$$f(x) = \exp[-x - \exp(-x)]$$

$$F(x) = \exp[-\exp(-x)]$$

IV-7-3-2. Moments

- $E(X) = 0,57722\dots$ (Constante d'Euler)
- $V(X) = \frac{\pi^2}{6}$

IV-7-3-3. Applications

Cette loi est utilisée pour les distributions de valeurs extrêmes. Elle permet de modéliser des phénomènes tels que :

- ✓ crue maximale annuelle d'une rivière,
- ✓ magnitude du plus grand tremblement de terre enregistré en un an.

Ch. V Convolution, fonctions caractéristiques

V- 1) CONVOLUTION

Etant donné deux variables aléatoires X et Y indépendantes ou non, il est fréquent d'avoir besoin de trouver la loi de la somme $Z = X + Y$.

V-1-1 Cas discret

Le théorème des probabilités totales nous permet d'écrire :

$$P(Z = z) = \sum_x [P(X = x) \cap (Y = z - x)] = \sum_x P(X = x) P(Y = z - x / X = x)$$

Si X et Y sont indépendantes, l'expression devient alors :

$$P(Z = z) = \sum_x P(X = x) P(Y = z - x)$$

Dans la sommation, X prend les seules valeurs compatibles avec $Z = z$.

Exemple

Soit $Z = X + Y$ la somme de deux variables de Poisson indépendantes de paramètres respectifs λ et μ :

$$P(Z = z) = \sum_{x=0}^z e^{-\lambda} \frac{\lambda^x}{x!} e^{-\mu} \frac{\mu^{z-x}}{(z-x)!} = \frac{e^{-\lambda-\mu}}{z!} \sum_{x=0}^z C_x^z \lambda^x \mu^{z-x}$$

$$P(Z = z) = \frac{e^{-\lambda-\mu}}{z!} (\lambda + \mu)^z$$

On montre ainsi la propriété suivante :

Si X et Y sont deux variables de Poisson de paramètres λ et μ , indépendantes, leur somme $X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$.

V-1-2 Cas général

V-1-2-1. Théorème

La loi de probabilité de la somme Z de deux variables indépendantes est la mesure image de $P_X \otimes P_Y$ par l'application :

$$(x, y) \rightarrow x + y \text{ de } \mathbb{R}^2 \text{ dans } \mathbb{R}.$$

C'est le produit de convolution des deux mesures.

Pour tout borélien B :

$$P_Z(B) = \iint_{\mathbb{R}^2} \chi_B(x + y) dP_X(x) dP_Y(y)$$

V-1-2-2. X et Y sont des variables continues

$$P_Z(B) = \iint_{R^2} \chi_B(x+y) f(x) g(y) dx dy$$

On pose $z = x + y$ et $u = x$ et on applique le théorème de Fubini :

$$k(z) = \int_{D_X} f(u) g(z-u) du = \int_{D_Y} g(y) f(z-y) dy$$

D_X et D_Y sont les domaines de valeurs des variables X et Y , compatibles avec $Z = z$.

Par intégration, on obtient :

$$P(Z < z) = \int_{D_X} f(x) G(z-x) dx = \int_{D_Y} g(y) F(z-y) dy$$

Exemple

Soit X et Y deux variables aléatoires indépendantes suivant des lois Gamma de paramètres respectifs r et s . Leurs densités de probabilité sont les suivantes :

$$f(x) = \frac{1}{\Gamma(r)} e^{-x} x^{r-1} \quad g(y) = \frac{1}{\Gamma(s)} e^{-y} y^{s-1}$$

La densité de $X + Y$ est alors donnée par :

$$k(z) = \int_0^z \frac{1}{\Gamma(r)} e^{-u} u^{r-1} \frac{1}{\Gamma(s)} e^{-z+u} (z-u)^{s-1} du = \frac{e^{-z}}{\Gamma(r)\Gamma(s)} \int_0^z u^{r-1} (z-u)^{s-1} du$$

En posant $u = t z$:

$$k(z) = \frac{e^{-z} z^{r+s-1}}{\Gamma(r)\Gamma(s)} \int_0^1 t^{r-1} (1-t)^{s-1} dt$$

$$k(z) = C e^{-z} z^{r+s-1} \quad \text{où } C = \frac{1}{\Gamma(r)\Gamma(s)} \int_0^1 t^{r-1} (1-t)^{s-1} dt .$$

On sait que k est une densité. On reconnaît que k est une densité de loi Gamma et donc :

$$C = \frac{1}{\Gamma(r+s)}$$

On en déduit le théorème suivant :

La somme de deux variables aléatoires indépendantes, suivant des lois Gamma de paramètres respectifs r et s , suit une loi Gamma de paramètre $r + s$.

Remarque :

On a démontré au passage la formule suivante :

$$\int_0^1 t^{r-1} (1-t)^{s-1} dt = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}$$

V- 2) FONCTIONS CARACTERISTIQUES

V-2-1 Définition

On définit la fonction caractéristique φ_X d'une variable aléatoire réelle X comme étant la transformée de Fourier de sa loi de probabilité.

$$\varphi_X(t) = E[e^{itx}] = \int_R e^{itx} dP_X(x)$$

Cette fonction existe toujours car P_X est une mesure bornée et $|\exp(itX)| = 1$.

V-2-2 Fonction caractéristique d'une variable continue

Lorsque la variable X possède une densité f , sa fonction caractéristique s'écrit simplement :

$$\varphi_X(t) = \int_R e^{itx} f(x) dx$$

V-2-3 Propriétés

La fonction caractéristique φ_X d'une variable réelle X possède les propriétés d'une transformée de Fourier. Ainsi :

- φ_X est une **fonction continue**.
- $\varphi_{\lambda X}(t) = \varphi_X(\lambda t)$
- $\varphi_{X+a}(t) = e^{ita} \varphi_X(t)$
- $\varphi_X(0) = 1$

V-2-4 Application de la formule de Mac -Laurin

- quand les dérivées jusqu'à l'ordre k de la fonction caractéristique existent, on montre que :

$$\varphi_X^{(k)}(0) = i^k E(X^k)$$

- Si la fonction caractéristique est indéfiniment dérivable, la formule de Mac-Laurin donne :

$$\varphi_X(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} i^k E(X^k)$$

V-2-5 Fonction caractéristique d'une somme de variables

La fonction caractéristique de la variable somme $X + Y$ de deux variables aléatoires indépendantes est le produit de leurs fonctions caractéristiques.

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

La fonction caractéristique de $X + Y$ est : $\varphi_{X+Y}(t) = E(e^{it(X+Y)}) = E(e^{itX} e^{itY})$

Les variables X et Y sont indépendantes donc les variables e^{itX} et e^{itY} sont elles aussi indépendantes et l'espérance du produit est alors égale au produit des espérances.

$$\varphi_{X+Y}(t) = E(e^{itX} e^{itY}) = E(e^{itX})E(e^{itY}) = \varphi_X(t)\varphi_Y(t)$$

V-2-6 Unicité et inversion de la transformée de Fourier

V-2-6-1. Unicité

D'après les propriétés des transformées de Fourier, si deux variables aléatoires X et Y ont même fonction caractéristique, alors elles ont même loi de probabilité.

La fonction caractéristique détermine une unique distribution de probabilité, comme son nom le laissait supposer.

V-2-6-2. Théorème d'inversion

Le résultat suivant est obtenu en appliquant une formule d'inversion de la transformée de Fourier :

Si $\int_R |\varphi_X(t)| dt < +\infty$, alors X admet une densité f continue donnée par :

$$f(x) = \frac{1}{2\pi} \int_R \varphi_X(t) e^{-itx} dt$$

V-2-7 Fonctions caractéristiques des lois discrètes usuelles

V-2-7-1. Loi de Bernoulli de paramètre p

$$\varphi_X(t) = pe^{it} + 1 - p$$

La fonction caractéristique est définie par : $\varphi_X(t) = E(e^{itX})$. Or la variable X ne prend que deux valeurs 1 et 0 donc : $\varphi_X(t) = p \times e^{it1} + (1 - p) \times e^{it0}$.

V-2-7-2. Loi binomiale $B(n, p)$

$$\varphi_X(t) = (pe^{it} + 1 - p)^n$$

La loi binomiale est la somme de n variables indépendantes de Bernoulli de paramètres p . On applique le résultat sur la fonction caractéristique d'une somme de variables indépendantes.

V-2-7-3. Loi de Poisson de paramètre λ

$$\varphi_X(t) = \exp[\lambda(e^{it} - 1)]$$

$$\varphi_X(t) = E[\exp(itX)] = \sum_{k=0}^{+\infty} \exp(itk) e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{[\lambda \exp(it)]^k}{k!} = e^{-\lambda} e^{\lambda \exp(it)}$$

V-2-8 Fonctions caractéristiques des lois continues usuelles

V-2-8-1. Loi uniforme sur $[-a, a]$

$$\varphi_X(t) = \frac{\sin(at)}{at}$$

Ce résultat a été démontré en séances d'exercices de Mathématiques.

V-2-8-2. Lois Gamma

- $n = 1$ $f(x) = \exp(-x)$

$$\varphi_{\gamma_1}(t) = \frac{1}{1 - it}$$

- n entier quelconque : la loi γ_n est la somme de n lois γ_1 indépendantes :

$$\varphi_{\gamma_n}(t) = \frac{1}{(1 - it)^n}$$

- pour r quelconque, la formule ci-dessus se généralise :

$$\varphi_{\gamma_r}(t) = \frac{1}{(1 - it)^r}$$

La démonstration repose sur une intégration dans le champ complexe.

V-2-8-3. Loi de Laplace-Gauss

- Variable normale centrée réduite : LG (0, 1)

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

On obtient le résultat, de façon simple, en appliquant la formule de Mac –Laurin :

$$E(U^{2k+1}) = 0, \quad E(U^{2k}) = \frac{(2k)!}{2^k k!},$$

$$\varphi_U(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} i^k E(U^k) = \sum_{k=0}^{+\infty} \frac{t^{2k}}{2k!} i^{2k} \frac{(2k)!}{2^k k!} = \sum_{k=0}^{+\infty} (-1)^k \frac{t^{2k}}{2^k k!} = \exp\left(-\frac{t^2}{2}\right)$$

- Variable normale de paramètres m et σ : LG (m, σ)

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

La démonstration utilise simplement le résultat précédent et les propriétés de linéarité de la fonction caractéristique.

V-2-9 Somme de deux variables normales indépendantes

La somme de deux variables indépendantes suivant les lois normales $LG(m_1, \sigma_1)$ et $LG(m_2, \sigma_2)$ suit la loi normale :

$$LG(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$$

Soit X et Y deux variables indépendantes de lois respectives : $LG(m_1, \sigma_1)$ et $LG(m_2, \sigma_2)$.

La fonction caractéristique de la somme $X + Y$ est :

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

Or, les fonctions caractéristiques de X et Y sont données par la formule du paragraphe précédent et on obtient donc la fonction caractéristique de la somme :

$$\varphi_{X+Y}(t) = e^{it(m_1+m_2)} \exp\left[-\frac{t^2}{2}(\sigma_1^2 + \sigma_2^2)\right]$$

On reconnaît la fonction caractéristique d'une variable suivant la loi normale $LG(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

Ch. VI Convergences de suites de variables

Dans ce chapitre, on va étudier les différents types de convergence que l'on rencontre en Probabilités.

VI- 1) CONVERGENCE EN PROBABILITE

VI-1-1 Définitions

- La suite de variables aléatoires $(X_n)_n$ converge en probabilité vers la constante a si :

$\forall \varepsilon$ et η arbitrairement petits, il existe n_0 tel que :

$$n > n_0 \Rightarrow P(|X_n - a| > \varepsilon) < \eta$$

- La suite de variables aléatoires $(X_n)_n$ converge en probabilité vers la variable X si la suite de variables $(X_n - X)_n$ converge vers 0.

VI-1-2 Remarque

D'après l'inégalité de Bienaymé Tchebychev,

$$P(|X_n - E(X_n)| > \varepsilon) < \frac{V(X_n)}{\varepsilon^2}$$

Lorsque $E(X_n) \rightarrow a$, il suffit de montrer que $V(X_n) \rightarrow 0$ pour établir la convergence de la suite $(X_n)_n$ vers a quand n tend vers l'infini.

VI- 2) CONVERGENCE PRESQUE SURE

VI-2-1 Définitions

- Deux variables aléatoires X et Y sont presque sûrement égales si :

$$P(\{w / X(w) \neq Y(w)\}) = 0$$

- La suite de variables aléatoires $(X_n)_n$ converge presque sûrement vers la variable X si :

$$P(\{w / \lim X_n(w) \neq X(w)\}) = 0$$

VI-2-2 Remarques

- ✓ La limite presque sûre d'une suite de variables aléatoires n'est pas unique mais deux limites éventuelles sont presque sûrement égales.
- ✓ La convergence presque sûre implique la convergence en probabilité.

VI- 3) CONVERGENCE EN MOYENNE D'ORDRE p

VI-3-1 Définition

La suite de variables aléatoires $(X_n)_n$ converge vers la variable X en moyenne d'ordre p si :

$$E[|X_n - X|^p] \rightarrow 0 \quad \text{quand } n \rightarrow +\infty$$

VI-3-2 Remarques

- La convergence en moyenne d'ordre p implique la convergence en probabilité.
- La convergence en moyenne quadratique d'ordre 2 est la plus couramment utilisée.

VI- 4) CONVERGENCE EN LOI

VI-4-1 Définitions

- La suite de variables aléatoires $(X_n)_n$ de fonctions de répartition F_n , converge en loi vers la variable X de fonction de répartition F quand n tend vers l'infini si, en tout point de continuité de F , la suite des $(F_n)_n$ converge vers F .
- Une suite de variables discrètes $(X_n)_n$ convergent vers la variable discrète X si :

$$\forall x, P(X_n = x) \rightarrow P(X = x) \quad \text{quand } n \rightarrow +\infty$$

VI-4-2 Propriétés

- La convergence en loi implique la convergence simple des densités respectives.
- La convergence en probabilité entraîne la convergence en loi.

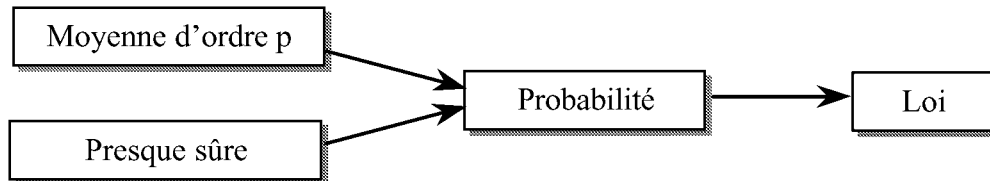
VI-4-3 Théorème de Lévy - Cramer – Dugué (admis)

La convergence en loi de la suite de variables aléatoires $(X_n)_n$ vers la variable X implique la convergence de la suite de fonctions caractéristiques des X_n vers la fonction caractéristique de X , la convergence étant uniforme sur tout intervalle fermé de $\mathbb{R}$.

Si les fonctions caractéristiques d'une suite de variables aléatoires $(X_n)_n$ convergent, quand n tend vers l'infini, vers une fonction φ dont la partie réelle est continue à l'origine, alors φ est la fonction caractéristique d'une variable X telle que la suite des variables aléatoires $(X_n)_n$ converge en loi vers X .

VI- 5) HIERARCHIE DES CONVERGENCES

Les diverses convergences en Probabilités vérifient les implications suivantes :



VI- 6) THEOREME DE DE MOIVRE – LAPLACE

VI-6-1 Théorème

$(X_n)_n$ étant une suite de variables aléatoires binomiales de paramètres n et p , la variable aléatoire : $U = \frac{X_n - np}{\sqrt{np(1-p)}}$ converge en loi vers une variable LG $(0, 1)$.

La démonstration repose sur les fonctions caractéristiques :

$$\varphi_{X_n}(t) = (p \exp(it) + 1 - p)^n$$

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

$$\text{Ln}[\varphi_U(t)] = n \text{Ln}\left[1 + p\left(\exp\left(\frac{it}{\sqrt{np(1-p)}}\right) - 1\right)\right] - \frac{itnp}{\sqrt{np(1-p)}}$$

$$\text{Ln}[\varphi_U(t)] \cong n \text{Ln}\left[1 + \frac{itp}{\sqrt{np(1-p)}} - \frac{pt^2}{np(1-p)}\right] - \frac{itnp}{\sqrt{np(1-p)}}$$

$$\text{Ln}[\varphi_U(t)] \cong n\left(\frac{itp}{\sqrt{np(1-p)}} + \frac{(p^2 - p)t^2}{2np(1-p)}\right) - \frac{itnp}{\sqrt{np(1-p)}} \cong -\frac{t^2}{2} \quad \varphi_U(t) \cong \exp\left(-\frac{t^2}{2}\right)$$

A la limite, on retrouve bien la fonction caractéristique de la loi normale.

VI-6-2 Approximation d'une loi binomiale par une loi normale, correction de continuité

Pour n assez grand, on peut considérer qu'une variable aléatoire suivant une loi binomiale de paramètres n et p , suit une loi de Laplace–Gauss $LG[np, \sqrt{np(1-p)}]$.

On considère que les conditions nécessaires pour cette approximation sont :

$$np > 5 \text{ et } n(1-p) > 5$$

Cette approximation est une conséquence du théorème précédent. Toutefois elle demande quelques précautions, en effet : la loi binomiale est une loi **discrète**, prenant ses valeurs dans

$[0, n]$ alors que la loi de Laplace-Gauss est une loi **continue** prenant ses valeurs dans $\mathbb{R}$. Dans le cas d'une loi discrète, un point a une probabilité non nulle, ce qui est faux dans le cas d'une loi continue.

La représentation de la densité de la loi normale est une courbe en cloche tandis que celle de la loi binomiale est un diagramme en bâtons.

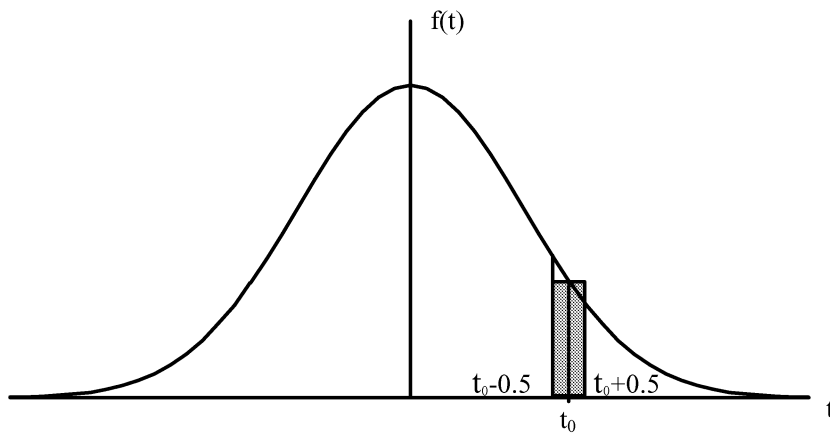
On est donc amené à effectuer la correction de continuité suivante :

Une valeur approchée de $P(T = t_0)$ est donnée par l'aire comprise entre la courbe en cloche et les droites $x = t_0 - 0,5$ et $x = t_0 + 0,5$.

$$P(T = t_0) = P\left(\frac{t_0 - 0,5 - np}{\sqrt{np(1-p)}} < U < \frac{t_0 + 0,5 - np}{\sqrt{np(1-p)}}\right)$$

et de même, si on cherche la probabilité que T soit inférieur ou égal à une valeur donnée t_0 :

$$P(T \leq t_0) = P\left(U < \frac{t_0 + 0,5 - np}{\sqrt{np(1-p)}}\right)$$



Exemple :

Soit X une variable aléatoire suivant la loi binomiale de paramètres $n = 40$ et $p = 0,3$.

On remarque que : $np = 12$ et $n(1-p) = 28$. On peut donc considérer que X suit approximativement une loi normale ; de plus $np(1-p) = 8,4$ donc cette loi normale est $LG(12, \sqrt{8,4})$.

$$P(X = 11) = C_{12}^{11}(0,3)^{11}(0,7)^1 = 0,1319$$

$P(10,5 < X < 11,5) = 0,131$. L'approximation est très satisfaisante.

VI- 7) CONVERGENCE DE LA LOI DE POISSON

VI-7-1 Théorème

Soit $(X_\lambda)_\lambda$ une famille de variables aléatoires telle que chaque variable X_λ suit la loi de Poisson $P(\lambda)$; la suite de variables $\frac{X_\lambda - \lambda}{\sqrt{\lambda}}$ converge en loi vers une variable aléatoire U suivant la loi LG $(0, 1)$ quand $\lambda \rightarrow +\infty$.

Démonstration :

$$\varphi_{X_\lambda}(t) = \exp[\lambda(e^{it} - 1)]$$

$$\varphi_{\frac{X_\lambda - \lambda}{\sqrt{\lambda}}}(t) = \exp\left[\lambda\left(\exp\left(\frac{it}{\sqrt{\lambda}}\right) - 1\right)\right] \exp\left(\frac{-it\lambda}{\sqrt{\lambda}}\right) = \exp\left[\lambda\left(\exp\left(\frac{it}{\sqrt{\lambda}}\right) - 1 - it\sqrt{\lambda}\right)\right]$$

$$\varphi_{\frac{X_\lambda - \lambda}{\sqrt{\lambda}}}(t) \cong \exp\left[\lambda + it\sqrt{\lambda} - \frac{t^2}{2} - \lambda - it\sqrt{\lambda}\right] \cong \exp\left(-\frac{t^2}{2}\right)$$

On retrouve là encore la densité de la loi normale.

VI-7-2 Approximation de la loi de Poisson par la loi normale

Soit X une variable suivant une loi de Poisson de paramètre λ . Pour λ assez grand, on peut considérer que X suit une loi de Laplace–Gauss : $LG(\lambda, \sqrt{\lambda})$. Cette approximation est très satisfaisante pour $\lambda > 18$.

Il est nécessaire, là encore, de faire une correction de continuité.

VI- 8) LOIS DES GRANDS NOMBRES

VI-8-1 Loi faible des grands nombres

Soient $(X_n)_n$ une suite de variables aléatoires indépendantes centrées telles que les variances $V(X_n)$ existent et vérifient : $\frac{1}{n^2} \sum_{i=1}^n V(X_i) \rightarrow 0$ quand $n \rightarrow +\infty$.

La suite des moyennes $\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i$ converge en probabilité vers 0 quand n tend vers l'infini.

Pour démontrer ce résultat, il suffit d'appliquer l'inégalité de Bienaymé-Tchebychev.

VI-8-2 Loi forte des grands nombres

Soient $(X_n)_n$ une suite de variables aléatoires indépendantes centrées telles que les variances $V(X_n)$ existent et vérifient :
$$\sum_{i \geq 1} \frac{V(X_i)}{i^2} < \infty.$$

La suite des moyennes $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ converge presque sûrement vers 0 quand n tend vers l'infini.

La démonstration repose sur l'inégalité de Kolmogorov que l'on admettra :

Si (X_n) est une suite de variables aléatoires indépendantes, centrées et de variances finies,

alors :

$$P\left(\sup_{k \leq n} \left| \sum_{i=1}^k X_i \right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sum_{i=1}^n V(X_i)$$

VI- 9) THEOREME DE LA LIMITE CENTRALE

La distribution normale fut introduite par le mathématicien De Moivre en 1733 comme approximation de la loi binomiale avec n suffisamment grand. Ce résultat fut généralisé par Laplace et d'autres pour devenir le « théorème central-limite » puis « théorème de la limite centrale » donnant les conditions dans lesquelles une variable aléatoire tend vers une variable normale.

VI-9-1 Théorème 1

Soit $(X_n)_n$ une suite de variables aléatoires indépendantes suivant toutes la même loi inconnue d'espérance m et d'écart -type σ .

La variable :
$$\frac{X_1 + X_2 + \dots + X_n - nm}{\sigma\sqrt{n}}$$
 converge en loi, quand $n \rightarrow +\infty$, vers une variable normale centrée réduite.

Démonstration :

$$\frac{X_1 + X_2 + \dots + X_n - nm}{\sigma\sqrt{n}} = \sum_{i=1}^n \frac{X_i - m}{\sigma\sqrt{n}}$$

Chaque variable $\frac{X_i - m}{\sigma\sqrt{n}}$ a pour espérance 0 et pour variance $1/n$. La fonction caractéristique

φ de cette variable a donc pour premiers termes : $1 - \frac{t^2}{2n}$, les termes suivants étant des

infinitement petits d'ordre $1/n^2$. La fonction caractéristique φ^n de la variable $\sum_{i=1}^n \frac{X_i - m}{\sigma\sqrt{n}}$ est

donc équivalente à $(1 - \frac{t^2}{2n})^n$ et tend vers $\exp(-\frac{t^2}{2})$ quand $n \rightarrow +\infty$.

VI-9-2 Théorème 2 (admis)

Ce théorème est dû à Lindeberg :

Soit $(X_i)_i$ une suite de variables aléatoires indépendantes d'espérances mathématiques m_i et de variances $V_i = V(X_i)$. Soient F_i les fonctions de répartition des variables $X_i - m_i$ et soit : $S_n^2 = \sum_{i=1}^n V_i$

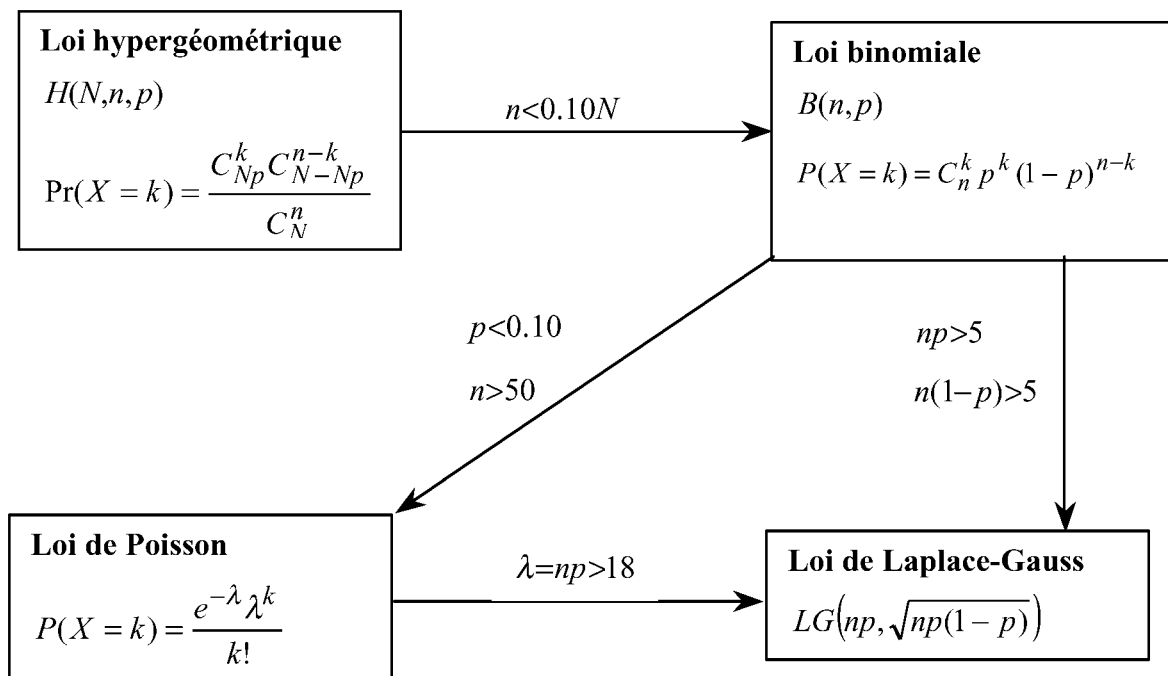
Si la limite de : $\frac{1}{S_n^2} \sum_{i=1}^n \int_{|x|>S_n} x^2 dF_i(x)$ est nulle quand n tend vers $+\infty$, alors la

variable aléatoire $\frac{1}{S_n} \sum_{i=1}^n (X_i - m_i)$ converge en loi vers une variable aléatoire U normale centrée réduite.

On peut remarquer que les variables X_i n'ont pas nécessairement la même loi.

VI- 10) RESUME DES APPROXIMATIONS

Le tableau ci-dessous permet de résumer les diverses approximations avec les conditions minima à respecter.



Ch. VII Couple de variables aléatoires

Dans le chapitre I, nous avons étudié le conditionnement d'un événement par un autre événement. Plus généralement, beaucoup d'études statistiques mettent en jeu plusieurs variables aléatoires et nous nous intéressons à leurs lois conjointes, conditionnelles...

Dans ce chapitre, nous allons exposer les généralités concernant un couple de variables aléatoires. Nous donnerons ensuite quelques résultats essentiels concernant les vecteurs aléatoires.

Usuellement, la loi de probabilité d'un couple (X, Y) de variables aléatoires réelles est une application de l'espace probabilisé (Ω, C, P) dans $\mathbb{R}^2$ muni de sa tribu de Borel. Toutefois dans la première partie de ce chapitre, nous allons nous intéresser à deux variables prenant leurs valeurs dans des ensembles finis ou dénombrables.

VII- 1) COUPLES DE VARIABLES ALEATOIRES DISCRETES

Soit X et Y deux variables aléatoires prenant leurs valeurs respectives x_i et y_j dans deux ensembles au plus dénombrables.

Y	y_1	...	y_j	...	y_q	
X						
x_1						
...			...			
x_i		...	p_{ij}	...		$p_{i.}$
...						
x_p			...			
			$p_{.j}$			

VII-1-1 Loi conjointe

VII-1-1-1. Définition

La loi du couple (X, Y) , appelée **loi conjointe**, est définie par la donnée des nombres p_{ij} :

$$p_{ij} = P(x_i, y_j) = p[(X = x_i) \cap (Y = y_j)]$$

VII-1-1-2. Propriétés

$$\checkmark \quad p_{ij} \in [0, 1]$$

$$\checkmark \quad \sum_i \sum_j p_{ij} = 1$$

VII-1-2 Lois marginales

On appelle loi marginale de X, la loi de X prise séparément :

$$P(X = x_i) = \sum_{j=1}^q p_{ij} = p_{i.}$$

On définit, de même, la loi marginale de Y.

VII-1-3 Lois conditionnelles

Considérons deux événements $\{X = x_i\}$ et $\{Y = y_j\}$ de probabilités non nulles. On peut alors définir deux lois conditionnelles en rappelant que X et Y peuvent être des variables qualitatives (cf chapitre I) :

VII-1-3-1. Loi conditionnelle de X sachant $Y = y_j$

$$P(X = x_i / Y = y_j) = \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{.j}}$$

VII-1-3-2. Loi conditionnelle de Y sachant $X = x_i$

$$P(Y = y_j / X = x_i) = \frac{P(X = x_i \cap Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i.}}$$

VII-1-3-3. Remarques

- Il est fréquent d'utiliser la connaissance que l'on a d'une loi conditionnelle et de la loi marginale correspondante pour déterminer la loi du couple, ceci à l'aide du théorème des probabilités totales :

$$P(X = x_i \cap Y = y_j) = P(X = x_i / Y = y_j) \times P(Y = y_j) = P(Y = y_j / X = x_i) P(X = x_i)$$

- Les formules de Bayes permettent d'exprimer une loi conditionnelle en fonction de l'autre, ainsi par exemple :

$$P(X = x_i / Y = y_j) = \frac{P(Y = y_j / X = x_i) P(X = x_i)}{\sum_{i=1}^p P(Y = y_j / X = x_i) P(X = x_i)}$$

VII-1-4 Indépendance

Les variables X et Y sont indépendantes si :

$$\forall (i, j) \quad p_{ij} = p_{i.} \times p_{.j}$$

Les p lois conditionnelles de Y à $X = x_i$ fixé, où i varie, sont donc toutes identiques à la loi de Y et de même les q lois de X à $Y = y_j$ fixé, où j varie, sont toutes égales à la loi de X.

VII-1-5 Espérance conditionnelle

Supposons que Y est une variable réelle. On va définir des moments conditionnels sous réserve qu'ils existent.

VII-1-5-1. Définition de l'espérance conditionnelle

On définit l'espérance conditionnelle de Y sachant $X = x$, notée $E(Y / X = x)$, comme l'espérance de Y prise par rapport à sa loi conditionnelle :

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

On remarque que c'est une fonction de x appelée aussi **fonction de régression de Y en X** . Son ensemble image est l'ensemble des moyennes conditionnelles de Y sachant X .

- On définit alors une **variable aléatoire appelée espérance conditionnelle**, comme la variable prenant les valeurs $E(Y / X = x)$ avec les probabilités respectives $P(X = x)$.

Cette variable est linéaire :

$$E(Y_1 + Y_2 / X) = E(Y_1 / X) + E(Y_2 / X)$$

VII-1-5-2. Théorème de l'espérance totale

$$\boxed{E[E(Y / X)] = E(Y)}$$

Revenons à la définition de l'espérance mathématique de la variable considérée c'est-à-dire la variable $E(Y / X)$:

$$E[E(Y / X)] = \sum_x E(Y / X = x)P(X = x) = \sum_x [\sum_y yP(Y = y / X = x)]P(X = x)$$

$$E[E(Y / X)] = \sum_y y [\sum_x P(Y = y / X = x)P(X = x)] = \sum_y yP(Y = y) = E(Y)$$

VII-1-6 Variance conditionnelle

VII-1-6-1. Définition de la variance conditionnelle de Y sachant $X = x$

Cette variance conditionnelle de Y sachant $X = x$, notée $V(Y / X = x)$, est définie par :

$$V(Y / X = x) = E[\{Y - E(Y / X = x)\}^2 / X = x]$$

On remarque là encore que c'est une fonction de x .

On peut donc définir une **variable aléatoire appelée variance conditionnelle** comme la variable prenant les valeurs $V(Y / X = x)$ avec une probabilité égale à $P(X = x)$.

VII-1-6-2. Théorème de la variance totale

$$V(Y) = E[V(Y/X)] + V[E(Y/X)]$$

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

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

- $E([Y - E(Y/X)]^2) = E(E([Y - E(Y/X)]^2 / X))$ d'après le théorème de l'espérance totale. On reconnaît l'expression de l'espérance de la variance conditionnelle :

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

- $E([E(Y/X) - E(Y)]^2) = E([E(Y/X) - E(E[Y/X])]^2) = V[E(Y/X)]$ en appliquant le théorème de l'espérance totale.
- $E([Y - E(Y/X)][E(Y/X) - E(Y)] / X) = [E(Y/X) - E(Y)]E([Y - E(Y/X)] / X)$ car le terme $E(Y/X) - E(Y)$ est constant à X fixé (espérance conditionnelle).

De plus en développant, on montre que $E([Y - E(Y/X)] / X) = 0$ donc l'espérance conditionnelle du double produit est nulle ; on en déduit que l'espérance du produit est nulle aussi. Le théorème est démontré.

VII- 2) VARIABLE CONTINUE ET VARIABLE DISCRETE**VII-2-1 Fonction de répartition conditionnelle**

Soit X une variable aléatoire discrète et soit Y une variable aléatoire continue. On définit simplement la fonction de répartition conditionnelle par :

$$G(y/x) = P(Y < y / X = x) = \frac{P(Y < y \cap X = x)}{P(X = x)}$$

La densité conditionnelle $g(y/x)$, quand elle existe, est, par définition, la dérivée de G.

VII-2-2 Densité marginale

La densité marginale de Y est obtenue par :

$$g(y) = \sum_x g(y/x)P(X = x)$$

VII-2-3 Moments

Si l'espérance mathématique de Y existe, l'espérance mathématique de la loi conditionnelle existe aussi et :

$$E(Y / X = x) = \int_R y g(y/x) dy$$

Les théorèmes de l'espérance totale et de la variance totale sont toujours valables.

De plus, grâce à la formule de Bayes, on a :

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

En revanche, il n'est pas question d'écrire ici une quelconque égalité entre « densités ponctuelles ».

VII- 3) COUPLE DE VARIABLES ALEATOIRES CONTINUES

VII-3-1 Lois conjointes et lois marginales

Soit (X, Y) un couple de variables à valeurs dans $\mathbb{R}^2$. On définit ainsi les fonctions suivantes :

VII-3-1-1. Fonction de répartition

$$H(x, y) = P(X < x \cap Y < y)$$

VII-3-1-2. Fonctions de répartition marginales

$$F(x) = H(x, \infty) = P(X < x)$$

$$G(y) = H(\infty, y) = P(Y < y)$$

VII-3-1-3. Densité

Si le couple (X, Y) est absolument continu, c'est-à-dire s'il admet une densité $h(x, y)$:

$$h(x, y) = \frac{\partial^2 H}{\partial x \partial y}$$

les densités marginales s'obtiennent par :

$$f(x) = \int_{\mathbb{R}} h(x, y) dy \quad g(y) = \int_{\mathbb{R}} h(x, y) dx$$

Rappelons que les variables aléatoires X et Y sont indépendantes si et seulement si :

$$\forall x, \forall y \quad h(x, y) = f(x) \times g(y)$$

VII-3-2 Densités conditionnelles d'un couple de variables continues

VII-3-2-1. Définition

Si le couple de variables aléatoires réelles (X, Y) admet une densité $h(x, y)$, les densités conditionnelles respectives des variables X et Y sont données par :

$$g(y/x) = \frac{h(x, y)}{f(x)} \quad f(x/y) = \frac{h(x, y)}{g(y)}$$

La deuxième formule de Bayes nous permet d'écrire :

$$g(y/x) = \frac{f(x/y)g(y)}{\int_R f(x/y)g(y)dy}$$

VII-3-2-2. Espérance conditionnelle

On définit l'espérance conditionnelle de Y sachant X = x par :

$$E(Y/X = x) = \int_R yg(y/x)dy$$

On définit de la même façon l'espérance conditionnelle de X sachant Y = y.

Les théorèmes de l'espérance totale et de la variance totale sont également valables dans le cas des variables continues.

On peut remarquer là encore que $E(Y/X = x)$ est une fonction de x.

La **courbe de régression** de Y en X est l'ensemble des points $\{x, E(Y/X = x)\}$.

VII- 4) COEFFICIENT ET RAPPORT DE CORRELATION

VII-4-1 Espace de Hilbert de classes de variables aléatoires

L'ensemble de toutes les variables aléatoires définies sur un même domaine Ω , muni du produit scalaire : $\langle X, Y \rangle = E(XY)$, forme un espace de Hilbert V.

Dans cet espace, la norme est donc : $\|X\| = \sqrt{E(X^2)}$.

Si une variable est centrée, on remarque que sa norme est égale à son écart-type. De même, pour deux variables centrées, la covariance est leur produit scalaire.

Considérons l'ensemble des variables aléatoires constantes, c'est une droite D de l'espace de Hilbert V.

L'espérance mathématique d'une variable X est la projection de cette variable sur la droite des variables constantes.

Démonstration : La projection de X sur D réalise le minimum de $\|X - a\|$.

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

Ce minimum est atteint pour $a = E(X)$.

On en déduit que, au sens de la norme L^2 , **E(X) est la meilleure approximation de la variable X par une constante.**

Comme $Cov(X, Y) = \langle [X - E(X)], [Y - E(Y)] \rangle$, par l'inégalité de Schwarz, on montre que :

$$|Cov(X, Y)| \leq \|X - E(X)\| \cdot \|Y - E(Y)\|.$$

Soit X et Y deux variables aléatoires.

Le cosinus de l'angle formé par $X - E(X)$ et $Y - E(Y)$ est égal à $\frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$.

En effet :
$$\frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\langle [X - E(X)], [Y - E(Y)] \rangle}{\|X - E(X)\| \cdot \|Y - E(Y)\|}$$

VII-4-2 Coefficient de corrélation linéaire ρ

VII-4-2-1. Définition

Soient X et Y deux variables aléatoires, on appelle coefficient de corrélation linéaire :

$$\rho = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

VII-4-2-2. Propriétés

- ρ est une mesure de dépendance symétrique
- $|\rho| \leq 1$
- $\rho = \pm 1$ si X et Y sont en liaison linéaire.

Le cosinus de l'angle $[X - E(X), Y - E(Y)]$ étant égal à ± 1 , on a : $Y - E(Y) = a [X - E(X)]$

- Si ρ est nul, X et Y ne sont pas en relation linéaire mais cela n'exclut pas l'existence de toute autre relation entre ces deux variables (cf : exemples en exercices).

VII-4-3 Interprétation géométrique de l'espérance conditionnelle

Considérons de nouveau l'espace de Hilbert V des variables aléatoires définies sur un domaine donné ; il admet le sous espace V_X constitué des variables aléatoires ne dépendant que d'une variable X donnée comme sous-espace de Hilbert.

L'opérateur associant à toute variable aléatoire son espérance conditionnelle à X est un projecteur orthogonal. $E(Y/X)$ s'interprète comme la projection orthogonale de Y sur V_X .

$E(Y/X)$ réalise donc le minimum de $E([Y - \varphi(X)]^2)$ à φ variable et est donc la meilleure approximation de Y par une fonction de X .

On en déduit que $Y - E(Y/X)$ est non corrélé avec X et le théorème de la variance totale est simplement le théorème de Pythagore appliqué au triangle $Y, E(Y)$ et $E(Y/X)$:

$$\|Y - E(Y)\|^2 = V(Y) = \|Y - E(Y/X)\|^2 + \|E(Y/X) - E[E(Y/X)]\|^2$$

$$\|Y - E(Y)\|^2 = V(Y) = V[E(Y/X)] + E[V(Y/X)]$$

VII-4-4 Rapport de corrélation η^2

VII-4-4-1. Définition

Le rapport de corrélation de la variable Y en la variable X est la mesure de liaison, non symétrique, définie par :

$$\eta_{Y/X}^2 = \frac{V[E(Y/X)]}{V(Y)}$$

VII-4-4-2. Propriétés

- $\eta_{Y/X}^2$ est le carré du cosinus de l'angle formé par $Y - E(Y)$ et le sous-espace V_X .
- $0 \leq \eta_{Y/X}^2 \leq 1$

en effet $V(Y) = V[E(Y/X)] + E[V(Y/X)]$ et $E[V(Y/X)] \geq 0$

VII-4-4-3. Rapport de corrélation maximal

$$\eta_{Y/X}^2 = 1 \quad \text{alors} \quad Y = \varphi(X)$$

Par le théorème de la variance totale, si le rapport de corrélation est égal à 1, alors $E[V(Y/X)]$ est nulle. On en déduit que $V(Y/X)$ est presque sûrement nulle (variable positive). Si on fixe X, V(Y) est nulle, donc Y ne prend qu'une valeur. On en déduit que Y est fonction de X.

VII-4-4-4. Nullité du rapport de corrélation

Si X et Y sont indépendantes, alors le rapport de corrélation $\eta_{Y/X}^2$ est nul

La réciproque est fautive : en effet, si le rapport de corrélation est nul, il y a absence de dépendance en moyenne mais on ne peut pas conclure que les variables sont indépendantes.

Si $\eta_{Y/X}^2 = 0$, alors on dit simplement que la variable Y est non corrélée avec la variable X : il y a absence de dépendance en moyenne. $Y - E(Y)$ est alors orthogonale à l'espace V_X .

VII-4-4-5. Rapport de corrélation et coefficient de corrélation

- $\eta_{Y/X}^2 \geq \rho^2$ car ρ^2 est le carré du cosinus de l'angle formé par $Y - E(Y)$ avec le sous-espace de V engendré par X et la droite D des constantes.
 - Si $\eta_{Y/X}^2 = \rho^2$ alors $E(Y / X) = aX + b$. En effet, l'égalité du coefficient et du rapport signifie que $E(Y/X)$ appartient au sous-espace engendré par X et D .
- C'est le cas de la **régression linéaire** (voir poly de statistique).

Ch. VIII Vecteurs aléatoires

VIII- 1) GENERALITES

VIII-1-1 Définition

Un vecteur aléatoire X est une application de l'espace probabilisé (Ω, C, P) dans un espace vectoriel réel, en général R^n .

En pratique, R^n est muni de sa base canonique ; on identifiera alors X au n-uplet de variables aléatoires formé par ses composantes sur cette base :

$$X = (X_1, X_2, \dots, X_n)$$

VIII-1-2 Fonction de répartition et densité

VIII-1-2-1. Fonction de répartition

La fonction de répartition F du vecteur aléatoire X est une application de R^n dans R définie par :

$$F(x_1, \dots, x_n) = P(X_1 < x_1, \dots, X_n < x_n)$$

VIII-1-2-2. Densité

La densité f d'un vecteur aléatoire continu est définie par :

$$f(x_1, \dots, x_n) = \frac{\partial^n F}{\partial x_1 \partial x_2 \dots \partial x_n}$$

VIII-1-2-3. Changement de variables dans une densité

Soit $Y = \varphi(X)$ un changement de variables biunivoque (X et Y sont des vecteurs de R^n).

Si f est la densité du vecteur X , la densité g du vecteur Y est donnée par :

$$g(y) = \frac{f[\varphi^{-1}(y)]}{|\det(J)|}$$

où $\det(J)$ est le jacobien de la transformation (déterminant des dérivées premières).

La démonstration est faite dans le cours de Mathématiques des classes préparatoires.

VIII-1-3 Fonction caractéristique

Soit a un vecteur non aléatoire de composantes $(a_1, \dots, a_n)$. On appelle fonction caractéristique du vecteur aléatoire X la fonction de a définie par :

$$\varphi_X(a) = E[\exp i(a_1 X_1 + \dots + a_n X_n)]$$

VIII-1-3-1. Théorème 1

Les composantes $X_1, X_2, \dots, X_n$ de X sont indépendantes si et seulement si la fonction caractéristique de X est égale au produit des fonctions caractéristiques de ses composantes :

$$\varphi_X(a) = \prod_{i=1}^n \varphi_{X_i}(a_i)$$

Si les X_i sont indépendantes, l'espérance du produit est égale au produit des espérances et la fonction caractéristique de X s'écrit bien comme le produit des fonctions caractéristiques des X_i .

La réciproque est admise ; elle fait appel à l'inversion d'une transformée de Fourier.

VIII-1-3-2. Théorème 2

La loi de X est entièrement déterminée par la connaissance des lois de toutes les combinaisons linéaires de ses composantes.

Ce théorème, admis ici, permet donc de déterminer la loi d'un vecteur à partir des lois des combinaisons linéaires de ses composantes. On va l'utiliser pour les vecteurs gaussiens.

VIII- 2) ESPERANCE, MATRICE DE COVARIANCE**VIII-2-1 Espérance**

Si $E(X_i)$ est l'espérance de la variable X_i , l'**espérance du vecteur X** est définie comme le vecteur colonne $E(X)$ de R^n dont les composantes sont les $E(X_i)$:

$$[E(X)]_i = E(X_i) \text{ pour tout } i$$

VIII-2-2 Matrice de Variance-Covariance

La **matrice de variance-covariance Σ de X** est définie comme la matrice carrée symétrique d'ordre p dont les éléments ε_{ij} sont :

$$\varepsilon_{ij} = \text{Cov}(X_i, X_j) \text{ pour } i \neq j$$

$$\varepsilon_{ii} = V(X_i)$$

VIII-2-3 Théorème de Cramer-Wold

Une condition nécessaire et suffisante pour qu'une matrice symétrique soit la matrice de variance-covariance d'un vecteur aléatoire est que cette matrice soit positive.

VIII- 3) VECTEURS GAUSSIENS. LOI MULTINORMALE

VIII-3-1 Définition

X vecteur de R^n est un vecteur aléatoire gaussien si toute combinaison linéaire de ses composantes est une variable aléatoire normale.

On remarquera que :

- la loi de X est parfaitement déterminée grâce au théorème de Cramer-Wold.
- il n'est pas suffisant à un vecteur aléatoire d'avoir toutes ses composantes variables gaussiennes pour qu'il soit lui-même gaussien.

VIII-3-2 Théorème

La fonction caractéristique de X, vecteur gaussien supposé centré est :

$$\varphi_X(a) = \exp\left(-\frac{1}{2}a'\Sigma a\right)$$

En effet, si on pose $Y = a'X = \sum_{i=1}^n a_i X_i$, la fonction caractéristique de Y s'écrit :

$$\varphi_Y(t) = E(e^{itY}) = E(e^{ita'X}).$$

On a alors $\varphi_Y(1) = E(e^{ia'X}) = \varphi_X(a)$

Or Y est une variable gaussienne centrée de variance : $V(Y) = a'\Sigma a$ comme combinaison linéaire gaussienne de variables gaussiennes centrées. Sa fonction caractéristique s'écrit :

$$\varphi_Y(t) = \exp\left[-\frac{t^2}{2}V(Y)\right] = \exp\left[-\frac{t^2}{2}(a'\Sigma a)\right]$$

En posant $t = 1$, on obtient l'expression voulue pour $\varphi_X(a)$.

VIII-3-3 Propriété

Les composantes d'un vecteur gaussien X sont indépendantes si et seulement si la matrice Σ est diagonale c'est-à-dire si et seulement si elles ne sont pas corrélées.

Si la matrice Σ est diagonale de termes $V(X_i)$, la variance de Y s'écrit comme la somme des

variances des X_i : $a'\Sigma a = \sum_{i=1}^n a_i^2 V(X_i)$ où a' est la transposée de a et $\varphi_X(a) = \prod_{i=1}^n \varphi_{X_i}(a_i)$.

VIII- 4) LOI NORMALE A DEUX DIMENSIONS

VIII-4-1 Définition de la densité

La matrice Σ de variance-covariance d'une loi normale à deux dimensions est la suivante :

$$\begin{vmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{vmatrix}$$

Le déterminant de la matrice Σ est :

$$\text{Det } \Sigma = \sigma_1^2 \sigma_2^2 (1 - \rho^2)$$

Soit $X = (X_1, X_2)$ un vecteur gaussien. Sa **densité** $f(x_1, x_2)$ est donnée par :

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-m_1}{\sigma_1}\right)^2 - 2\rho\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-m_2}{\sigma_2}\right)^2\right]\right\}$$

VIII-4-2 Propriétés

On suppose, ici, pour alléger les notations, que les espérances m_1 et m_2 sont nulles.

La densité du couple gaussien peut s'écrire :

$$f(x_1, x_2) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[-\left(\frac{x_1^2}{2\sigma_1^2}\right)\right] \times \frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left[-\frac{(x_2 - \rho\frac{\sigma_2}{\sigma_1}x_1)^2}{2\sigma_2^2(1-\rho^2)}\right]$$

- La distribution marginale de la variable X_1 (respectivement X_2) est la loi normale de paramètres $LG(0, \sigma_1)$ (respectivement $LG(0, \sigma_2)$).
- La distribution conditionnelle de X_2 sachant X_1 a pour loi : $LG(\rho\frac{\sigma_2}{\sigma_1}X_1, \sigma_2\sqrt{1-\rho^2})$

La courbe de régression est la droite d'équation : $E(X_2 / X_1) = \rho\frac{\sigma_2}{\sigma_1}X_1$

VIII-4-3 Théorème

Soit $X = (X_1, X_2)$ un couple gaussien à valeurs dans $\mathbb{R}^2$. Les variables X_1 et X_2 sont indépendantes si et seulement si elles ne sont pas corrélées.

Si les variables sont indépendantes, on sait qu'elles ne sont pas corrélées.

Si les variables ne sont pas corrélées, alors $\text{Cov}(X_1, X_2) = 0$ donc $\rho = 0$. La densité du vecteur X se décompose en un produit de deux densités gaussiennes $LG(m_1, \sigma_1)$ et $LG(m_2, \sigma_2)$.

VIII- 5) LOI DU χ^2 (CHI-DEUX)

VIII-5-1 Définition

$X_1, X_2, \dots, X_n$ étant n variables normales, centrées, réduites, indépendantes, on appelle loi du Chi-deux à n degrés de liberté, la loi de la variable :

$$\sum_{i=1}^n X_i^2$$

C'est donc la loi de la somme des carrés des composantes d'un vecteur gaussien centré et de matrice de variance-covariance la matrice identité.

VIII-5-2 Densité

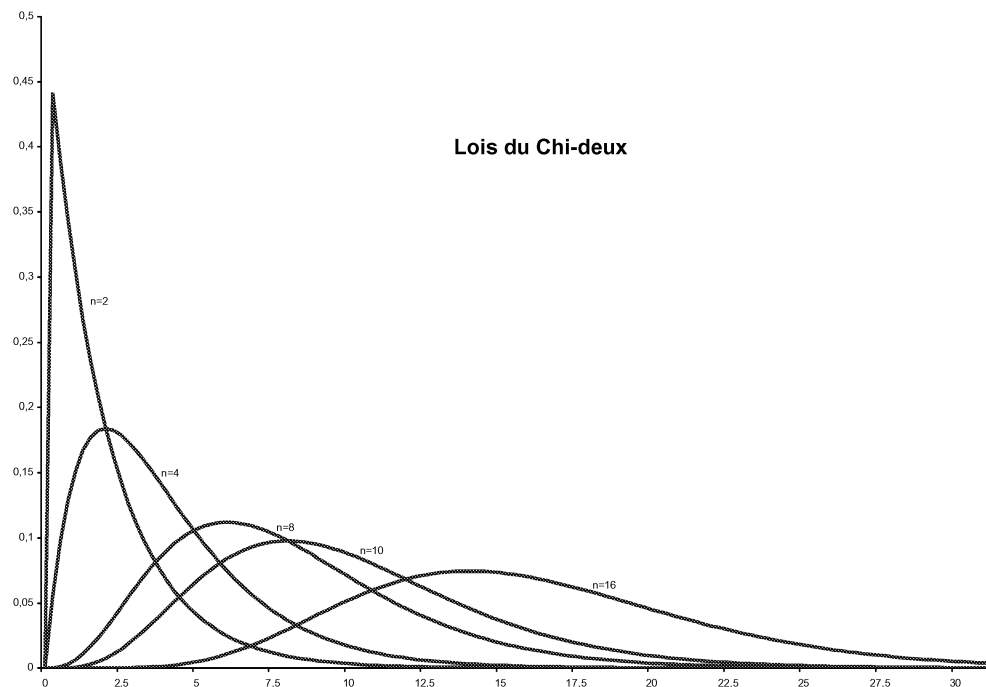
La densité de cette loi χ^2 s'écrit :

$$g(u) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{u}{2}} u^{\frac{n}{2}-1}$$

Les valeurs de sa fonction de répartition ont été tabulées.

VIII-5-3 Courbes représentatives

Dans la représentation graphique ci-dessous, figurent les courbes représentatives des densités des lois du Chi-deux pour les valeurs $n = 2$, $n = 4$, $n = 8$, $n = 10$ et $n = 16$ du paramètre.



VIII-5-4 Théorèmes

VIII-5-4-1. Somme de lois du chi-deux

Si la loi de X est la loi $\chi^2(m)$, si la loi de Y est la loi $\chi^2(n)$, et si X et Y sont indépendantes alors la loi de leur somme X+Y est la loi $\chi^2(m+n)$.

Par définition, $X + Y$ est la somme de $m+n$ variables normales centrées réduites indépendantes.

VIII-5-4-2. Loi du chi-deux et loi Gamma

Soit X une variable aléatoire suivant une loi Gamma de paramètre r. La variable 2X suit une loi du Chi-deux à 2r degrés de liberté.

$$L(X) = \gamma_r \Rightarrow L(2X) = \chi^2(2r)$$

La densité d'une variable du chi-deux à un degré de liberté est celle du carré d'une variable de

Laplace-Gauss U^2 . Elle s'écrit : $f(t) = \frac{1}{\sqrt{2\pi}} t^{-1/2} \exp(-\frac{t}{2})$

On en déduit que la variable $T = \frac{U^2}{2}$ suit une loi Gamma de paramètre $\frac{1}{2}$.

VIII-5-5 Fonction caractéristique

Elle se déduit de celle de la loi Gamma :

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

VIII-5-6 Moments

Soit X une variable aléatoire suivant une loi du Chi-deux à n degrés de liberté :

$$E(X) = n$$

$$V(X) = 2n$$

Par définition, la loi du chi-deux à n degrés de liberté est la loi de la somme de n variables carrés de variables normales centrées réduites indépendantes. Grâce aux propriétés de l'espérance et de la variance :

$$E(X_i) = 0 \quad E(X_i^2) = V(X_i) + [E(X_i)]^2 = 1 \quad E\left(\sum_{i=1}^n X_i^2\right) = n$$

Si la loi de X est une loi $\chi^2(n)$ la loi de $Y = X/2$ est une loi $\gamma_{n/2}$.

$$V(Y) = \frac{n}{2} \quad V(X = 2Y) = 4 \times \frac{n}{2} = 2n$$

VIII- 6) LOI DE FISHER-SNEDECOR

VIII-6-1 Définition

Soit X et Y deux variables aléatoires indépendantes suivant des lois du Chi-deux à, respectivement, n et p degrés de liberté ; on appelle loi de Fisher–Snedecor la loi de la variable :

$$Z = \frac{X/n}{Y/p}$$

Cette loi est très importante en Statistique en particulier pour les intervalles de confiance (voir polycopié de Statistique). Découverte en 1924 par Fisher, elle a été tabulée par Snedecor en 1934.

VIII-6-2 Densité

Elle s'écrit (de façon peu simple) :

$$g(t) = \frac{\Gamma(n)}{\Gamma^2(\frac{n}{2})} \frac{(\frac{n}{p})^{\frac{n}{2}} t^{\frac{n}{2}-1}}{(1 + \frac{n}{p} t)^{\frac{n+p}{2}}}$$

VIII-6-3 Moments

$$E(Z) = \frac{p}{p-2} \text{ pour } p > 2.$$

$$V(Z) = 2 \frac{p^2}{n} \frac{n+p-2}{(p-2)^2(p-4)} \text{ pour } p > 4.$$

VIII-6-4 Tables de la loi de Fisher

Les valeurs données par la table dépendent du seuil α et des degrés de liberté n et p . Les tables 10 et 11 du polycopié de tables statistiques donnent les valeurs k et k' telles que :

$$P(F(n, p) > k) = 0,05 \text{ et } P(F(n, p) > k') = 0,01$$

On peut remarquer si Z suit une loi de Fisher $F(n, p)$ alors $1/Z$ suit une loi de Fisher $F(p, n)$ et :

$$P(F(p, n) < \frac{1}{k}) = 0,05 \text{ et } P(F(p, n) < \frac{1}{k'}) = 0,01$$

VIII- 7) LOI DE STUDENT

VIII-7-1 Définition

Soit X une variable normale centrée réduite et Y une variable suivant une loi du Chi-deux de paramètre n , indépendante de X . On définit la loi de Student à n degrés de liberté comme la loi de la variable T définie par :

$$T = \frac{X}{\sqrt{Y/n}}$$

Une variable de Student est une variable centrée.

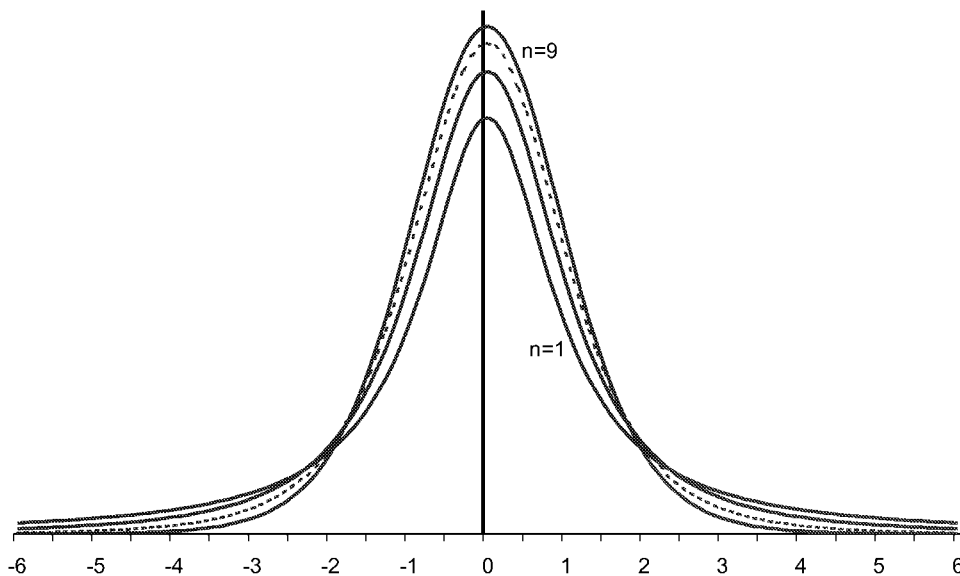
Les valeurs de sa fonction de répartition ont été tabulées.

VIII-7-2 Densité

La densité de la loi de Student à n degrés de liberté est :
$$f(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n} \Gamma(\frac{1}{2}) \Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

VIII-7-3 Représentations graphiques

Dans le graphique ci-dessous, sont représentées les densités des lois de Student pour les valeurs $n = 1$, $n = 2$, $n = 4$ et $n = 9$.



VIII-7-4 Moments de la loi de Student

La variable étant centrée, les moments d'ordre impair sont nuls :

$$E(T) = 0 \text{ pour } n > 1$$

$$V(T) = \frac{n}{n-2} \text{ pour } n > 2$$

$$\mu_3 = 0 \text{ pour } n > 3 \quad \mu_4 = \frac{3n^2}{(n-2)(n-4)} \text{ pour } n > 4$$

VIII-7-5 Relations avec les lois de Laplace-Gauss et de Fisher-Snedecor

- La loi de Student converge vers la loi normale centrée réduite LG (0, 1) quand $n \rightarrow +\infty$.
- Si T est une variable aléatoire de Student de paramètre n, alors la variable T^2 est une variable de Fisher - Snedecor F(1, n).

VIII-7-6 Cas particulier de la loi de Cauchy

Pour $n = 1$, la loi de Student correspondante est la loi du quotient de deux variables aléatoires normales indépendantes. Cette loi est appelée **loi de Cauchy** et a la caractéristique de ne posséder aucun moment. Sa densité est : $f(t) = \frac{1}{\pi(1+t^2)}$.

Ch. IX Processus aléatoires

On dit qu'un système évolue de façon aléatoire s'il prend au cours du temps une série d'états successifs dont on ne peut pas prévoir sa configuration à l'instant $n+1$ ni son évolution future. On cherche alors un modèle mathématique susceptible de décrire le comportement de tels systèmes à l'aide des observations des états passés. En particulier, on a fait appel aux probabilités pour analyser ces évolutions.

L'analyse de ces systèmes débouche sur de nombreuses applications telles que l'étude des files d'attente, des arrivées de clients dans un service ou l'étude de phénomènes plus spécifiquement physiques tels que la pression, le débit...

IX- 1) GENERALITES SUR LES PROCESSUS ALEATOIRES

On ne parlera ici que de processus aléatoires réels.

IX-1-1 Définition

On appelle **fonction aléatoire réelle** une application de l'espace probabilisé (Ω, C, P) dans un espace de fonctions de variables réelles, qui associe à tout élément a de Ω , une fonction de la variable réelle t :

$$a \rightarrow X_t(a)$$

où $X_t(a)$ est la fonction : $t \rightarrow X_t(a)(t) = X(a, t)$.

Si t représente le temps, on parle de **processus aléatoire ou stochastique**. Si t prend des valeurs discrètes équidistantes, on parle de **séries chronologiques**.

Pour a fixé, $X_t(a)$, fonction du temps, est appelée **trajectoire de a** .

A t fixé, X_t fonction de a est une **variable aléatoire réelle** sur (Ω, C, P) .

Un processus est donc décrit par une suite de variables aléatoires, indicée par le temps.

IX-1-2 Moments

Si la variable X_t admet pour tout t une espérance m_t et une variance σ_t^2 finies, celles-ci définissent des fonctions du temps.

Toutefois des processus peuvent avoir même espérance et même variance et être très différents. C'est pourquoi on introduit deux caractéristiques du processus qui permettent de résumer sa régularité :

✓ **Fonction de covariance :**

$$C(t, s) = \text{Cov}(X_t, X_s)$$

✓ **Fonction d'autocorrélation :**

$$\rho(t, s) = \frac{C(t, s)}{\sigma_t \sigma_s}$$

IX-1-3 Continuité

Un processus aléatoire $(X_t)_t$ est **continu en moyenne quadratique** si :

$$E[(X_{t+h} - X_t)^2] \rightarrow 0 \text{ dès que } h \rightarrow 0$$

La continuité en moyenne quadratique implique la continuité de l'espérance, de la variance et de la fonction de covariance. En revanche, elle ne permet pas d'en déduire la continuité des trajectoires.

IX-1-4 Processus stationnaires

IX-1-4-1. Définition

Un processus est dit **stationnaire** si sa loi de probabilité est invariante par translation :

$$P[(X_{t_1} < x_1) \cap \dots \cap (X_{t_n} < x_n)] = P[(X_{t_1+h} < x_1) \cap \dots \cap (X_{t_n+h} < x_n)] \quad \forall (t_i), \forall h$$

Les processus physiques sont pour la plupart stationnaires. En revanche, ce n'est pas le cas de la plupart des processus économiques.

IX-1-4-2. Propriétés

X_t et X_{t+h} ont alors mêmes caractéristiques :

$$\forall t \quad E(X_t) = m$$

$$\forall t \quad V(X_t) = \sigma^2$$

$$\forall (t, s) \quad C(t, s) = C$$

IX-1-4-3. Autres définitions

- Lorsque seules les propriétés ci-dessus sont vérifiées, on dit que le processus est stationnaire au sens large.
- Un processus tel que $X_{t+h} - X_t$ est stationnaire pour tout h , est dit à **accroissements stationnaires**.

IX-1-5 Processus ponctuel

On définit un processus ponctuel comme étant un processus caractérisé par une famille de variables de Bernoulli $(X_t)_t$, chaque variable X_t étant égale à 1 si un événement se produit à l'instant t . Le nombre d'événements arrivés entre 0 et t définit une fonction aléatoire.

Ces processus ponctuels sont très utilisés dans les études des phénomènes suivants :

- ✓ accidents dans une usine,
- ✓ arrivée de clients dans un service, arrivée de bateaux dans un port,
- ✓ appels téléphoniques à un standard ...

IX- 2) PROCESSUS PONCTUEL DE POISSON

IX-2-1 Définitions

IX-2-1-1. Processus de Poisson

Un processus ponctuel est dit processus de Poisson s'il représente une succession d'événements aléatoires $(E_i)_{i \in [1,n]}$ vérifiant :

- ✓ Les temps d'attente entre deux événements consécutifs sont des variables indépendantes (processus sans mémoire).
- ✓ La loi du nombre d'événements arrivant dans l'intervalle $\{t, t+T\}$ ne dépend que de T.
- ✓ Deux événements ne peuvent arriver simultanément.

IX-2-1-2. Cadence d'un tel processus

On appelle cadence du processus l'espérance de la loi du nombre d'événements arrivant dans l'intervalle $\{t, t+1\}$.

IX-2-2 Lois associées au processus de Poisson

IX-2-2-1. Etude de la probabilité qu'aucun événement ne se produise pendant une durée h

Soit $P_0(h)$ la probabilité qu'il ne se passe rien durant l'intervalle $[t, t+h[$. D'après la deuxième assertion sur les processus de Poisson, elle ne dépend que de h. On en déduit que si h et k sont deux durées, en considérant les intervalles $[t, t+h+k[$, $[t, t+h[$ et $[t+h, t+h+k[$ et d'après la troisième assertion :

$$\forall(h, k) \quad P_0(h + k) = P_0(h) \times P_0(k)$$

On en déduit le résultat suivant, où $c > 0$:

$$P_0(h) = \exp(-ch)$$

IX-2-2-2. Etude de la durée T séparant deux événements consécutifs

La probabilité que cette durée T soit supérieure à t est la probabilité qu'il ne se passe rien dans un intervalle de durée t, soit :

$$P(T > t) = \exp(-ct)$$

On en déduit la fonction de répartition :

$$P(T < t) = 1 - \exp(-ct)$$

Cette fonction de répartition est celle de la loi exponentielle de paramètre c. Son espérance est égale à :

$$E(T) = 1 / c$$

On sait que cT suit alors une loi γ_1 .

IX-2-2-3. Etude de la durée Θ séparant n événements consécutifs

Θ est la durée séparant n événements consécutifs $E_1, \dots, E_n$. Alors $T_1, \dots, T_{n-1}$ est la succession des durées séparant les événements E_i et :

$$\Theta = T_1 + \dots + T_{n-1}$$

Θ est la somme de n variables aléatoires indépendantes toutes de même loi. Chaque cT_i suit une loi γ_1 donc $c\Theta$ suit une loi γ_{n-1} . La densité de Θ est donc :

$$f_{n-1}(\theta) = \exp(-c\theta) \frac{(c\theta)^{n-2}}{(n-2)!} c$$

La durée Θ séparant n événements consécutifs suit une loi Gamma de paramètre $n-1$.

IX-2-2-4. Etude du nombre d'événements se produisant pendant une durée T

Soit N le nombre d'événements se produisant pendant une durée T . La probabilité que $N = n$ est :

$$P(N = n) = P(N \geq n) - P(N \geq n+1)$$

Soit Θ la variable aléatoire séparant n événements consécutifs.

$$P(\Theta > \theta) = \frac{1}{\Gamma(n-1)} \int_0^\theta e^{-ct} (ct)^{n-2} dt$$

$$\text{Donc } P(N > n / T) = \frac{1}{\Gamma(n-1)} \int_0^T e^{-ct} (ct)^{n-2} dt$$

$$\text{On en déduit : } P(N = n) = \frac{1}{\Gamma(n)} \int_0^T e^{-ct} (ct)^{n-1} dt - \frac{1}{\Gamma(n+1)} \int_0^T e^{-ct} (ct)^n dt$$

En intégrant par parties la première intégrale, on obtient :

$$P(N = n) = \exp(-cT) \frac{(cT)^n}{n!}$$

On en déduit les résultats suivants :

- $E(N) = cT$
- **c est la cadence c'est-à-dire l'espérance mathématique du nombre d'événements aléatoires observés par unité de temps.**

IX-2-2-5. Répartition des dates $E_1, \dots, E_n$ dans l'intervalle de temps $[0, T]$

La probabilité pour que le premier événement se passe entre t_1 et $t_1 + dt_1$ est :

$$P(t_1 < E_1 < t_1 + dt_1) = c \exp(-ct_1) dt_1$$

La probabilité conditionnelle que le deuxième événement se passe entre t_2 et $t_2 + dt_2$ sachant la date à laquelle E_1 a eu lieu est :

$$P(t_2 < E_2 < t_2 + dt_2 / E_1) = c \exp[-c(t_2 - t_1)] dt_2$$

...

La probabilité qu'aucun événement n'arrive après E_n sachant la date de E_n est :

$$\exp[-c(T - t_n)]$$

On en déduit la densité de la loi de probabilité de $(E_1, E_2, \dots, E_n, N)$:

$$f(t_1, t_2, \dots, t_n, n) = c^n \exp(-cT)$$

D'où la densité de la loi conditionnelle :

$$g(t_1, t_2, \dots, t_n / N = n) = c^n \exp(-cT) \times \frac{n!}{(cT)^n \exp(-cT)}$$

$$g(t_1, t_2, \dots, t_n / N = n) = \frac{n!}{T^n}$$

Les temps $t_1, t_2, \dots, t_n$ constituent un échantillon ordonné de la loi uniforme sur $[0, T]$. Si on s'intéresse aux dates seules et non à leur ordre, la densité est divisée par $n!$, nombre d'ordres possibles.

IX-2-2-6. Etude du processus

Le processus N_t , nombre d'arrivées pendant un temps t , suit une loi de Poisson $P(ct)$. On en déduit l'espérance et la variance : $E(N_t) = ct$ et $V(N_t) = ct$.

On en déduit que le processus n'est pas stationnaire. En revanche, il est à accroissements stationnaires : la variable $(N_{t+h} - N_t)$ suit une loi de Poisson de paramètre h

Cherchons la fonction de covariance de ce processus :

✓ Si $s > t$: $N_s = N_t + X$ où X est une variable indépendante de N_t (processus stationnaire) et :

$$C(t, s) = \text{Cov}(N_t, N_t + X) = V(N_t) + \text{Cov}(N_t, X) = V(N_t) \text{ car } X \text{ et } N_t \text{ sont indépendantes.}$$

On en déduit que : $C(t, s) = ct$

✓ Si $s < t$ on démontre, de la même façon, que $C(t, s) = cs$

On obtient enfin :

$$C(t, s) = c \times \inf(t, s)$$

Cette fonction est continue en $t = s$. Le processus est continu en moyenne quadratique.

IX- 3) APPLICATIONS AUX PHENOMENES D'ATTENTE

Ces phénomènes sont décrits pour représenter des arrivées aléatoires de clients à un guichet comportant plusieurs postes où ils resteront durant un temps lui aussi aléatoire.

Le cas le plus courant (et le plus simple) est celui pour lequel :

- ✓ Les arrivées suivent un processus de Poisson de cadence λ
- ✓ Le temps de service obéit à une loi exponentielle de paramètre μ où μ est l'espérance du nombre de clients pouvant être servis par poste et par unité de temps.

On suppose que la file d'attente est illimitée et que l'ordre d'arrivée est respecté.

On pose : $\psi = \frac{\lambda}{\mu}$.

Il est nécessaire que : $\psi < k$ où k est le nombre de guichets, sinon la file s'allongera indéfiniment car le nombre moyen d'arrivées sera supérieur au nombre moyen de départs.

IX-3-1 Probabilités élémentaires

IX-3-1-1. Probabilité d'une arrivée pendant un temps dt infiniment petit

On sait que le nombre d'arrivées X obéit à un processus de Poisson de cadence λ . Donc la probabilité d'aucune arrivée pendant dt est :

$$P(X = 0) = \exp(-\lambda dt)$$

La durée dt est un infiniment petit donc, on en déduit :

$$P(X = 0) \cong 1 - \lambda dt \quad P(X = 1) \cong \lambda dt$$

IX-3-1-2. Probabilité d'un départ pendant la durée dt

La probabilité qu'il n'y ait aucun départ est celle d'une durée de service au moins égale à dt , c'est-à-dire :

$$P(T > dt) = \exp(-\mu dt) \cong 1 - \mu dt$$

$$1 - P(T > dt) \cong \mu dt$$

IX-3-1-3. Probabilités élémentaires des événements arrivant pendant le temps dt

On a admis qu'il ne pouvait pas y avoir deux événements simultanés donc il ne peut y avoir un départ et une arrivée pendant le temps dt . La probabilité qu'il y ait une arrivée et aucun départ est le produit $\lambda dt \times (1 - \mu dt) \cong \lambda dt$, le terme en dt^2 étant considéré comme négligeable. On obtient finalement :

	Un départ	Aucun départ
Une arrivée	0	λdt
Aucune arrivée	μdt	$1 - \lambda dt - \mu dt$

Attention : s'il y a s postes de service, la probabilité d'un départ est $s\mu dt$.

IX-3-2 Equations d'état

Nous recherchons les probabilités des états du système $P_n(t)$, c'est-à-dire les probabilités qu'il y ait n clients dans le système à l'instant t . Pour cela, nous allons examiner les différentes éventualités compatibles avec la définition du processus.

IX-3-2-1. Première équation d'état

On cherche $P_0(t + dt)$. Deux possibilités apparaissent et seulement deux :

- ✓ Il n'y a aucun client au temps t et il n'en arrive pas dans l'intervalle $]t, t + dt]$
- ✓ Il y a un client au temps t et il part durant l'intervalle $]t, t + dt]$

On en déduit l'équation :

$$P_0(t + dt) = (1 - \lambda dt)P_0(t) + P_1(t)\mu dt$$

IX-3-2-2. Equations d'état pour $n < k$

Cherchons $P_n(t + dt)$ comme précédemment :

- ✓ Il y a n clients au temps t et il n'y a aucun départ et aucune arrivée
- ✓ Il y a $(n+1)$ clients au temps t et il y a un départ
- ✓ Il y a $(n-1)$ clients au temps t et il y a une arrivée et aucun départ

L'équation d'état s'écrit alors :

$$P_n(t + dt) = (1 - \lambda dt - n\mu dt)P_n(t) + P_{n+1}(t)(n+1)\mu dt + P_{n-1}(t)\lambda dt$$

IX-3-2-3. Equations d'état pour $n \geq k$

Le raisonnement est le même que ci-dessus mais la probabilité élémentaire d'un départ durant un temps dt est égale à $k\mu dt$. On obtient l'équation suivante :

$$P_n(t + dt) = (1 - \lambda dt - k\mu dt)P_n(t) + P_{n+1}(t)k\mu dt + P_{n-1}(t)\lambda dt$$

IX-3-2-4. Résolution des équations

L'évolution du système est décrit par les équations d'état ci-dessus auxquelles il faut ajouter la relation :

$$\sum_{n=0}^{+\infty} P_n(t) = 1$$

et les conditions initiales :

$$P_0(0) = 1$$

$$\forall n, P_n(0) = 0$$

La résolution de ces équations fait appel à la transformation de Laplace et aux fonctions de Bessel. Toutefois, on s'intéresse au régime stationnaire.

IX-3-3 Régime stationnaire

IX-3-3-1. Définition et étude du régime stationnaire

Sous la condition :

$$\frac{\lambda}{\mu k} < 1$$

les probabilités $P_n(t)$ peuvent être considérées comme indépendantes du temps. C'est la caractérisation du régime stationnaire.

En supposant donc, que ces probabilités ne dépendent pas du temps, on obtient :

$$\lambda P_0 = \mu P_1$$

$$\mu(n+1)P_{n+1} = (\lambda + n\mu)P_n - \lambda P_{n-1}$$

$$\mu k P_{n+1} = (\lambda + k\mu)P_n - \lambda P_{n-1}$$

En résolvant ces équations de proche en proche, on obtient :

$$P_n = \frac{1}{n!} \times \left(\frac{\lambda}{\mu}\right)^n \times P_0 \quad 1 \leq n < k$$

$$P_n = \frac{1}{k! k^{n-k}} \times \left(\frac{\lambda}{\mu}\right)^n \times P_0 \quad n \geq k$$

On obtient P_0 en écrivant que la somme des probabilités est égale à 1 :

$$P_0 = \frac{1}{1 + \sum_{i=1}^{k-1} \frac{\lambda^i}{i! \mu^i} + \frac{\lambda^k}{k! \mu^k (1 - \frac{\lambda}{\mu k})}}$$

IX-3-3-2. Cas extrêmes

- Si le centre ne possède qu'un guichet, ($k = 1$) :

$$P_n = (1 - \frac{\lambda}{\mu}) \left(\frac{\lambda}{\mu}\right)^n$$

On retrouve une distribution géométrique de fonction de répartition : $F(n) = 1 - \left(\frac{\lambda}{\mu}\right)^{n+1}$

- Si le centre possède un grand nombre de guichets, le risque de file d'attente est très faible. A la limite, le nombre d'utilisateurs dans le centre est distribué suivant une loi de Poisson :

$$P_n = \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n \exp\left(-\frac{\lambda}{\mu}\right)$$

IX-3-3-3. Caractéristiques moyennes du régime stationnaire

Appelons :

N : le nombre d'utilisateurs dans le centre

F : le nombre d'utilisateurs dans la file d'attente

$$F = N - k \text{ pour } N > k$$

$$F = 0 \text{ sinon}$$

I : le nombre de postes inoccupés

A : le temps d'attente

Posons $\phi = \frac{\lambda}{\mu}$

On obtient par calcul toutes les caractéristiques suivantes :

$$E(F) = P_0 \times \frac{\phi^{k+1}}{(k-1)!(k-\phi)^2}$$

$$E(I) = k - \phi$$

$$E(A) = P_0 \times \frac{\phi^k}{\mu(k-1)!(k-\phi)^2}$$

En régime stationnaire, on peut remarquer que

$$E(F) = \lambda E(A)$$

Enfin, dans le cas d'un guichet unique, on a les expressions très simples suivantes :

$$E(I) = P_0 = 1 - \phi$$

$$E(F) = \frac{\phi^2}{1 - \phi} = \phi E(N)$$

$$E(A) = \frac{\phi}{\mu(1 - \phi)}$$