REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

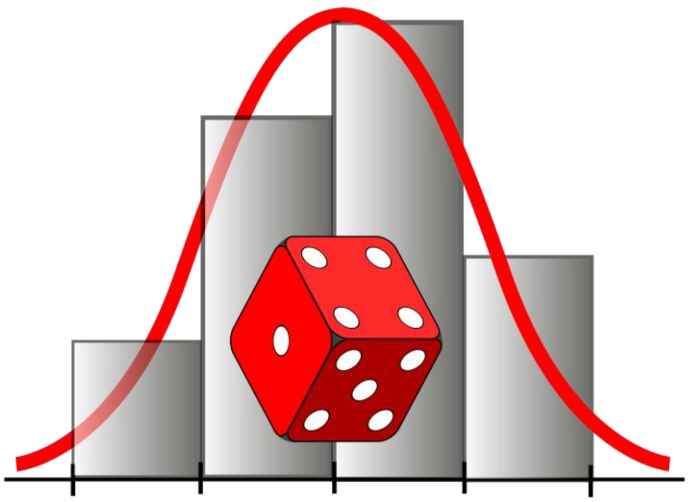
Ministère de l'enseignement supérieur et de la recherche scientifique

Université Mohamed Seddik Benyahia-Jijel

Faculté des Sciences Exactes et Informatique

Département de Mathématiques

Probabilités & Statistique



Ce cours est destiné aux étudiants en deuxième année licence, faculté des sciences et de la technologie, département de génie civil

Préparé par :

Dr. Ahlam GUIATNI

Année académique 2023 - 2024

Table des matières

Liste des tableaux

Liste des figures

Introduction

Le but de ce polycopié est de présenter une introduction aux probabilités et statistique, pouvant intéresser tous ceux qui marquent de l'intérêt pour cette discipline, qu'elle que soit leur spécialisation. Ce cours sera une référence pour les étudiants de deuxième année génie civil, génie mécanique, génie de procédés, automatique et électronique.
La statistique joue un rôle de plus longtemps très important dans tous les domaines notamment agriculture, médicine, biologie, physique et d'autres branches de sciences de la technologie.
Les probabilités est parmi les modules qui sont tenues par des nombreux chercheurs. La plupart des études dans les différentes sciences dépendent de la théorie des probabilités. Par exemple, les résultats sont souvent basés sur les lois de probabilités.

Ce cours est organisé autour de deux parties. La première partie est consacrée à la statistique, nous avons présenté les définitions de base de la statistique, séries statistiques à une variable et séries statistiques à deux variables. Ensuite le but de la deuxième partie est de présenter l'analyse combinatoire, les calculs de probabilités, les variables aléatoires et les lois de probabilité usuelles. Ce module permet aux étudiants de traiter les notions essentielles de la probabilité et de la statistique, à savoir : les séries statistiques à une et à deux variables, la probabilité sur un univers fini et les variables aléatoires.

Objectifs

Les objectifs du cours sont :

- ✓ Comprendre les notions de base : Population, Echantillon, Variable…
- ✓ Connaître les principaux paramètres de position et de dispersion et savoir les calculer.
- ✓ Analyser la distribution des valeurs des variables et le lien éventuel entre elles.
- ✓ Maitriser des notions élémentaires de probabilité.
- ✓ Connaître et savoir utiliser les formules des probabilités équiprobables.
- ✓ Découvrir les variables aléatoires.
- ✓ Apprendre à manipuler les variables aléatoires.
- ✓ Familiariser avec quelques lois usuelles.

Prérequis

Le cours «probabilités et statistique» est destiné aux étudiants en deuxième année licence, faculté des sciences et technologie, université de Jijel, premier semestre, matière commune aux spécialités : génie civil, hydraulique, génie mécanique et génie des procédés.

Pour pouvoir suivre ce cours, il est recommandé aux apprenants de connaître :

- ➢ Fondamentaux des mathématiques (Mathématiques 1 et 2).
- ➢ Notion des ensembles et sous-ensembles.
- ➢ Les calculs d'intégrales.

Partie A : Statistique

Objectifs

L'objectif de la partie A est de :

- ✓ Comprendre les notions de base : Population, Echantillon, Variable…
- ✓ Savoir résumer les données d'observation des variables sous forme de tableaux de distribution et de représentations graphiques adaptées.
- ✓ Connaître les principaux paramètres de position et de dispersion et savoir les calculer.
- ✓ Savoir commenter et interpréter les résultats.
- ✓ Analyser la distribution des valeurs des variables et le lien éventuel entre elles.

Prérequis

Il est recommandé aux apprenants de connaître.

- ➢ Notions des ensembles et sous-ensembles.
- ➢ Fonctions à plusieurs variables.

Mots clés

Population, Echantillon, Variable aléatoire, Paramètres de position, Paramètres de dispersion, Corrélation, Régression linéaire.

Chapitre 1 : Séries statistiques à une variable

Introduction

La statistique est un ensemble de méthodes qui servent à organiser les épreuves fournissant des observations, à analyser et à interpréter les résultats. L'analyse statistique se subdivise en deux grandes parties : Statistique descriptive et Statistique inférentielle. La statistique descriptive à pour but de décrire c'est à dire de résumer et de représenter les données par des représentations graphiques, paramètres de position, paramètres de dispersion et de relation et paramètres de forme.

1. Définitions de base

1.1. Notions de Population, d'échantillon, variables, modalités

Population : La collection d'objets ou de personnes étudiées (élèves, habitants, voitures…).

Individu : Elément de la population étudiée. (un élève, un habitant, une voiture,…).

Echantillon : Partie de la population étudiée. Nombre d'individus dans un échantillon noté n est appelé taille de l'échantillon

Variable : (Caractère noté X) Propriété commune aux individus de la population, que l'on veut étudier.
Nous allons distinguer deux groupes de variables aléatoires VA: Les VA qualitatives et les VA quantitatives.

Modalité : l'une des formes particulières d'un caractère X_i. La couleur des yeux est un caractère, ses modalités sont : bleu, vert, marron,…

1.2. Différents types des variables statistiques

Variable qualitative : La variable est dite qualitative quand les modalités sont des catégories.

- Variable qualitative nominale : La variable est dite qualitative nominale quand les modalités ne peuvent pas être ordonnées.
- Variable qualitative ordinale : Les modalités de la variable possèdent la propriété d'ordre.

Variable quantitative : Une variable est dite quantitative si toutes ses valeurs possibles sont numériques.

- Variable quantitative discrète : Une variable est dite discrète, si l'ensemble des valeurs possibles est dénombrable (des nombres entiers).
- Variable quantitative continue : Une variable est dite continue si elle peut prendre toutes les valeurs dans un intervalle donné (borné ou non borné).

2. Effectif et fréquence

2.1. Effectif ou fréquence absolue :

(noté n_i) nombre d'apparitions de la valeur associé à un caractère dans un échantillon.
($n = \sum_{i=1}^{k} n_i$)

- Effectif cumulé croissant : (noté $N_i \nearrow$) somme cumulée du effectif avec tous ses effectifs précédents. $N_i \nearrow = \sum_{j=1}^{i} n_j \; ; \; j \leq i$
- Effectif cumulé décroissant : (noté $N_i \searrow$) est donné par : $N_i \searrow = n - \sum_{j=1}^{i-1} n_j$

2.2. Fréquence relative :

(noté f_i) est le rapport de cet effectif à l'effectif total de la population.
$f_i = \frac{n_i}{n}$; $i = 1, 2, \dots, k$

- Fréquence cumulé croissante : (noté $F_i \nearrow$) $F_i \nearrow = \sum_{j=1}^{i} f_j \; ; \; j \leq i$
- Fréquence cumulé croissante : (noté $F_i \searrow$) $F_i \searrow = 1 - \sum_{j=1}^{i-1} f_j$

Remarque:

- $0 \leq f_i \leq 1$
- f_i peut être exprimée en pourcentage %. ($p_i = f_i \times 100$)
- $\sum_{i=1}^{k} f_i = 1$

On présente ces notions dans le tableau suivant, en écrivant les modalités numériques en ordre croissant.

X_i	n_i	$N_i \nearrow$	$N_i \searrow$	f_i	$F_i \nearrow$	$F_i \searrow$
X_1	n_1	n_1	N	f_1	f_1	1
X_2	n_2	$n_1 + n_2$	$N - n_1$	f_2	$f_1 + f_2$	$1 - f_1$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
X_k	n_k	N	n_k	f_k	1	f_k
Total	N			1		

Tableau 1 : Tableau statistique

Exemple 1 :

En notant X le nombre d'enfants de 20 familles, on a obtenu le tableau

0	2	2	3	1	3	4	2	2	3
1	0	1	4	2	2	1	3	1	2

ou la série ordonnée

0 0 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 4 4;

qui donne le tableau statistique suivant

X_i	n_i	N_i ↗	N_i ↘	f_i	F_i ↗	F_i ↘
0	2	2	20	0.10	0.10	1
1	5	7	18	0.25	0.35	0.90
2	7	14	13	0.35	0.70	0.65
3	4	18	6	0.20	0.90	0.30
4	2	20	2	0.10	1	0.1
Total	20			1		

Tableau 2 : Tableau statistique d'exemple 1

- La population : les 20 famille
- Le Caractère étudier X : Le nombre d'enfants
- Le type de caractère : Quantitatif discret
- Les modalité du caractère : $X_i \in \{0, 1, 2, 3, 4\}$

3. Représentations tabulaires et graphiques

3.1. Caractère qualitatif :

X_i	n_i	f_i	θ_i	p_i
X_1	n_1	f_1	$f_1 \times 360$	$f_1 \times 100$
X_2	n_2	f_2	$f_2 \times 360$	$f_2 \times 100$
. . .	. . .	. . .	. . .	. . .
X_k	n_k	f_k	$f_k \times 360$	$f_k \times 100$
Total	N	1	360	100

Tableau 3 : Tableau statistique « Caractère qualitatif »

Lorsque le caractère étudié est qualitatif on utilise un diagramme à bandes ou diagramme à secteurs.

- **Diagramme à bandes :** C'est un diagramme qui à chaque modalité de la variable associé un rectangle de base constante et dont la hauteur est proportionnelle à l'effectif.

- **Diagramme à secteurs :** Ce type de graphique est formé d'un cercle divisé en secteurs, chaque secteur représente une catégorie particulière. La surface de chacun des secteurs est donné en degré par : $\theta_i = f_i .360$

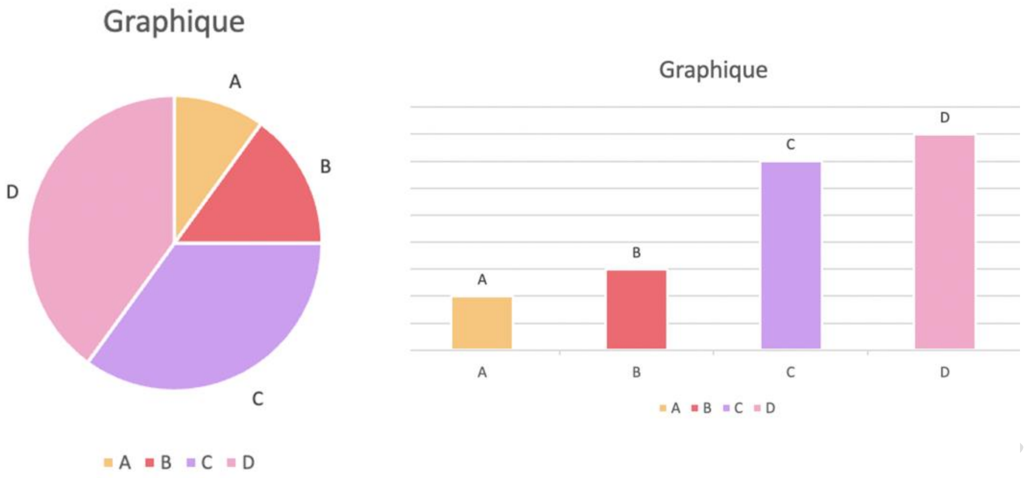


Figure 1: Diagramme à secteurs et diagramme à bandes

3.2.Caractère quantitatif discrète :

X_i	n_i	$N_i \nearrow$	$N_i \searrow$	f_i	$F_i \nearrow$	$F_i \searrow$
X_1	n_1	n_1	N	f_1	f_1	1
X_2	n_2	$n_1 + n_2$	$N - n_1$	f_2	$f_1 + f_2$	$1 - f_1$
. . .	. . .	. . .	. . .	. . .	. . .	. . .
X_k	n_k	N	n_k	f_k	1	f_k
Total	N			1		

Tableau 4 : Tableau statistique « Caractère quantitatif discret »

Lorsque le caractère étudié est quantitatif discret on utilise un diagramme en bâton.

- **Diagramme en bâton :** On associe un segment vertical dont la hauteur est proportionnelle à la valeur (effectif ou fréquence) connue.

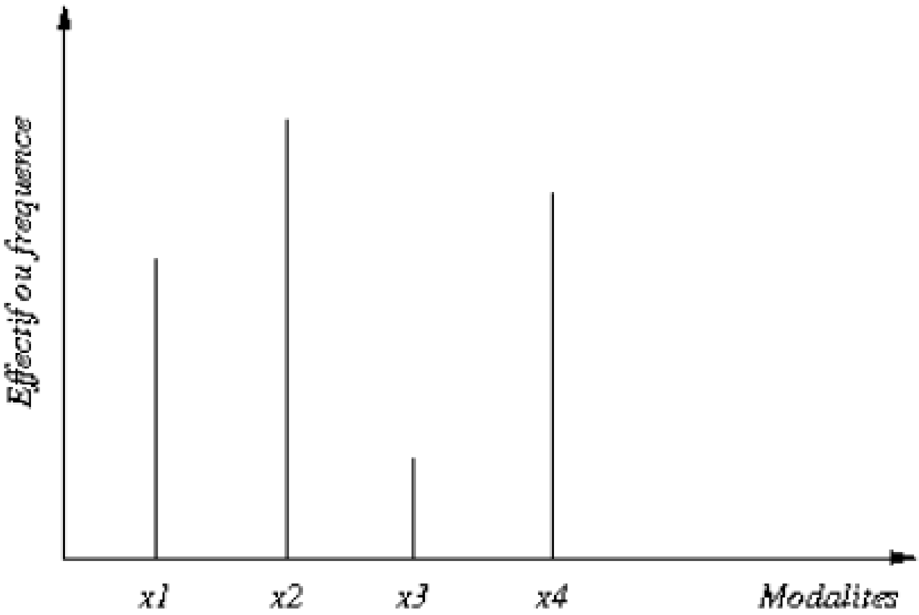


Figure 2: Diagramme en bâtons

3.3.Caractère quantitatif continu :

Classes $[e_i , e_{i+1}[$	Centre c_i	n_i	N_i ↗	N_i ↘	f_i	F_i ↗	F_i ↘
$[e_1 , e_2[$	c_1	n_1	n_1	N	f_1	f_1	1
$[e_2 , e_3[$	c_2	n_2	$n_1 + n_2$	$N - n_1$	f_2	$f_1 + f_2$	$1 - f_1$
. . .		. . .	. . .	. . .	. . .	. . .	. . .
$[e_{k-1}, e_k[$	c_k	n_k	N	n_k	f_k	1	f_k
Total		$\sum_{i=1}^{k} n_i = N$			$\sum_{i=1}^{k} f_i = 1$		

Tableau 5: Tableau statistique « Caractère quantitatif continu »

- Une classe est un intervalle de type $[e_i, e_{i+1}[$
- Le centre de classe c_i est : $\frac{e_i + e_{i+1}}{2}$
- L'amplitude d'une classe est : $a_i = e_{i+1} - e_i$

Lorsque le caractère étudié est quantitatif continu on utilise un histogramme.

- **Histogramme :** C'est un diagramme composé des rectangles adjacents, chaque rectangle associé à chaque classe ayant une surface proportionnelle à l'effectif ou à la fréquence de cette classe.

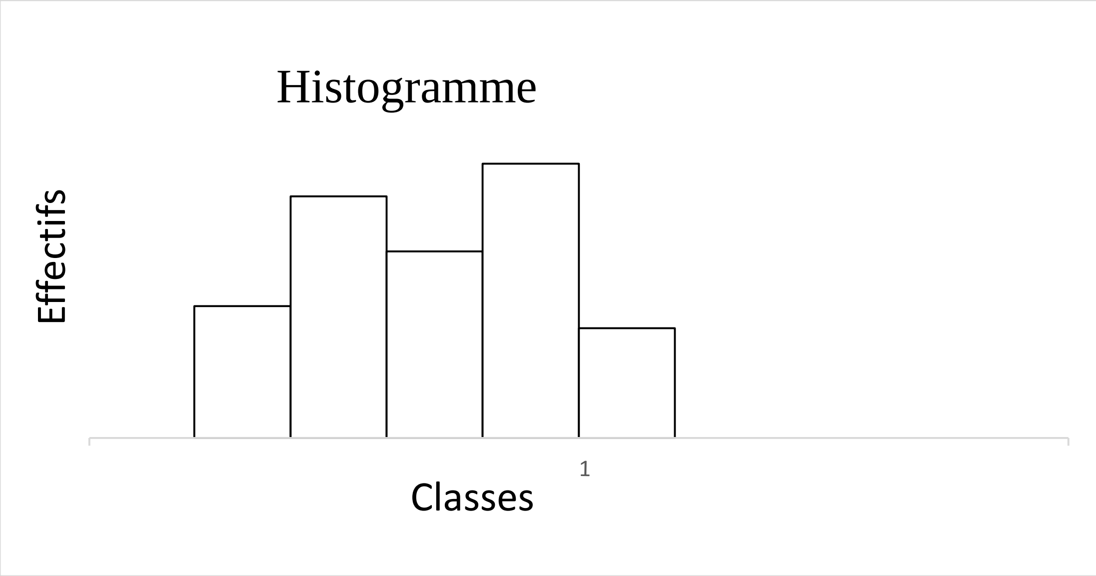


Figure 3 : Histogramme

- **Propriétés importantes sur l'histogramme :** Lorsque les classes ont toutes, la même amplitude, les hauteurs des rectangles sont proportionnelles à leurs surfaces ; par conséquent les hauteurs des rectangles sont proportionnelles aux fréquences et aux effectifs. Dans le cas où les classes sont d'amplitudes inégales, la hauteur de rectangle correspondant à la i-ème classe sera $h_i = \frac{f_i}{a_i}$ (c'est-à-dire la fréquence par unité d'amplitude) ou encore $H_i = \frac{n_i}{a_i}$ (c'est-à-dire l'effectif par unité d'amplitude).

Remarque :

Diagramme cumulatif est le diagramme représentatif de la fréquence cumulée ou d'effectif cumulé.

- Diagramme cumulatif pour une variable discrète : formé à l'aide d'un diagramme en escalier.

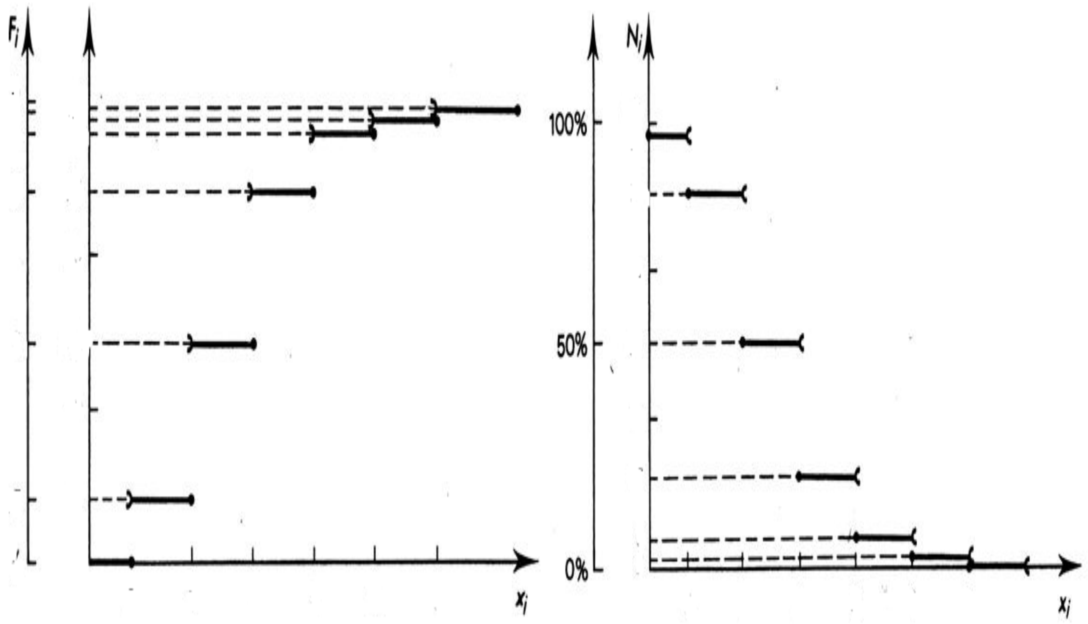


Figure 4: Diagramme cumulatif

- Diagramme cumulatif pour une variable continue : formé à l'aide de segment de droite.

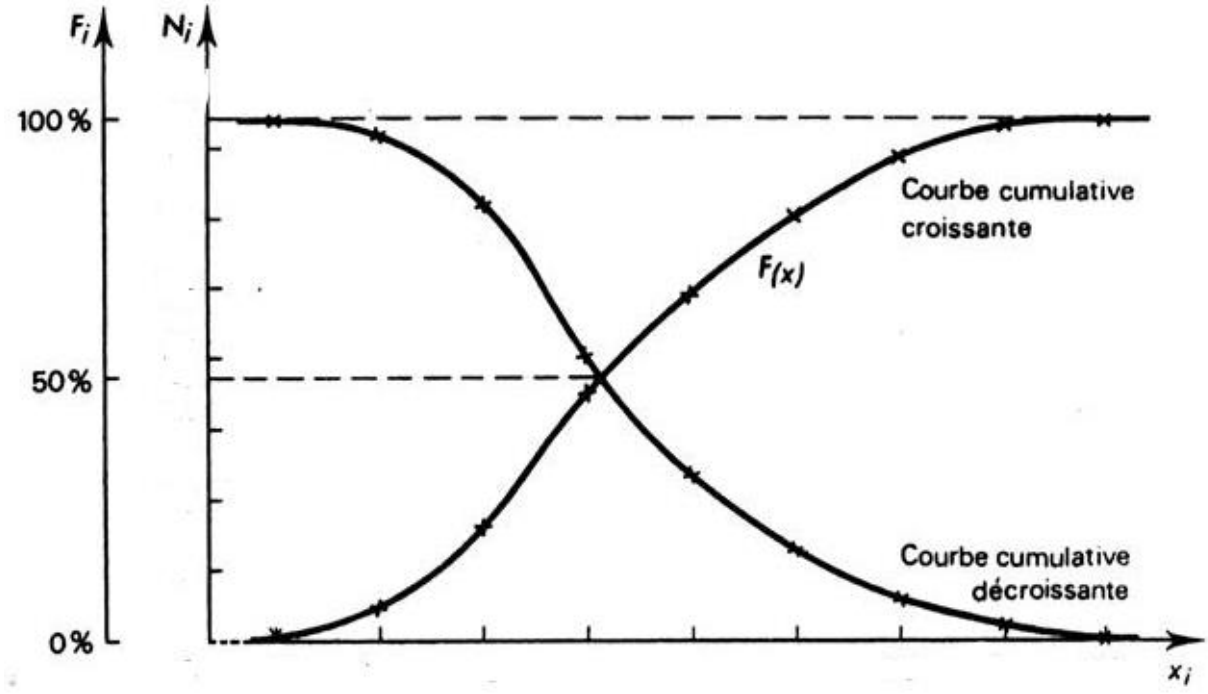


Figure 5 : Polygone cumulatif

4. Caractéristiques de position :

Les paramètres de position ou « mesures de tendance centrale » sont des grandeurs susceptibles de représenter au mieux un ensemble de données. L'appellation «tendance centrale » vient du

fait que ces paramètres donnent une idée de ce qui se passe au centre d'une distribution d'un ensemble de données. Les paramètres de position permettent de savoir autour de quelles valeurs se situent les valeurs d'une variable statistique.
On distingue trois mesures de tendance centrale :

4.1. Moyenne

La moyenne constitue l'un des paramètres fondamentaux de tendance centrale, la mesure la plus calculée et la plus utilisée lors de la description de séries statistiques mais non suffisant pour caractériser une distribution. Soit $\{x_1, x_2, \ldots, x_k\}$ une série statistique et $\{n_1, n_2, \ldots, n_k\}$ les effectifs correspondants.

4.1.1. Moyenne arithmétique : C'est la plus simple et la communément utilisée et ce, pas toujours à bon escient. Elle se note la plupart du temps par $\bar{x}$.

- Pour une variable discrète : $\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$
- Pour une variable continue : $\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i c_i$

Exemple 2 : Soit la série statistique suivante : 12; 13; 4; 13; 12; 14; 15; 7; 15; 13; 12; 15; 7

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \frac{3\times12+3\times13+1\times4+1\times14+3\times15+2\times7}{13} = \frac{152}{13} = 11.69$$

Exemple 3 : Soit la série statistique suivante :

Poids	[50, 55[	[55, 60[	[60, 65[	[65, 70[
n_i	2	4	8	6

On remplace les x_i par les centres de classes c_i :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i c_i = \frac{2\times52.5+4\times57.5+8\times62.5+6\times67.5}{20} = \frac{1297.5}{20} = 64.875$$

4.1.2. Moyenne géométrique : La moyenne géométrique est un instrument permettant de calculer des taux moyens notamment des taux moyens annuels. La formule de la moyenne géométrique de cette série est donnée par : $\bar{G} = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}}$

4.1.3. Moyenne harmonique : On utilise la moyenne harmonique lorsqu'on veut déterminer un rapport moyen dans des domaines où il existe des liens de proportionnalité inverse. La formule de la moyenne harmonique de cette série est donnée par : $\bar{H} = \frac{n}{\sum_{i=1}^{k}\frac{n_i}{x_i}}$

4.1.4. Moyenne quadratique : Une moyenne qui trouve des applications lorsque l'on a affaire à des phénomènes présentant un caractère sinusoïdal avec alternance de valeurs positives et de valeurs négatives. Elle est, de ce fait, très utilisée en électricité. Elle permet notamment de calculer la grandeur d'un ensemble de nombre. A titre d'information. La formule de la moyenne quadratique de cette série est donnée par : $\bar{Q} = \sqrt{\frac{1}{n} \sum_{i=1}^{k} n_i x_i^2}$

4.2. Mode « Mo »:

- Pour une variable discrète : Le mode est la valeur de la sèrie ayant l'effectif le plus élevé $Mo = x_i|_{\max(n_i)}$. Lorsque la distribution a plus d'un mode, on parle d'une distribution «multimodale» (bimodale, trimodale, etc).
- Pour une variable continue : On cherche la classe modale abrégée [e_i, e_{i+1}[celle à effectifs le plus élevé. Pour avoir une valeur exacte, le mode se calcule de la manière suivante : $Mo = e_j + \frac{d_j}{d_{j+1} + d_j} \times a_j$

  e_j : limite inférieure de la classe modale ;
  a_j : amplitude de la classe modale ;
  d_j : écart d'effectif entre la classe modale et la classe inférieure la plus proche ;
  d_{j+1} : écart d'effectif entre la classe modale et la classe supérieure la plus proche.

Exemple 4 : Soit la série statistique suivante : 12; 13; 4; 13; 12; 14; 15; 7; 15; 12; 15; 7
$Mo_1 = 12$, $Mo_2 = 15$

Exemple 5 : Soit la série statistique suivante :

Poids	[50, 55[	[55, 60[	[60, 65[	[65, 70[	[70, 75[	[75, 80[	[80, 85[
n_i	2	5	12	16	14	8	3

La classe modale est [65, 70[

$e_j = 65$; $a_j = 5$; $d_j = 16-12$; $d_{j+1} = 16-14$

Alors, $Mo = 65 + \frac{4}{2+4} \times 5 = 68.33$

4.3. Médiane « Me »

Valeur centrale sur l'axe x divisant l'échantillon en 2 groupes égaux d'individus. Pour calculer Me, il faut d'abord ordonner la série.

- Pour une variable discrète : On désigne par n le nombre d'observations.
  - Si n est impair : Il est possible d'identifier simplement la valeur qui partage

la population en deux effectifs égaux. Le rang central étant égal à $\frac{n+1}{2}$ alors $Me = x_{\frac{n+1}{2}}$.

- o Si n est pair : La médiane est alors égale à la moyenne des valeurs encadrant le milieu de la série alors $Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$

- <u>Pour une variable continue</u> : On cherche la classe contenant le $\frac{n}{2}$ individu de l'échantillon. Cette classe est appelée la classe médiane. En supposant que tous les individus de cette classe sont uniformément répartis à l'intérieur et $Me \in [e_j\ , e_{j+1}[$ la médiane se calcul de la façon suivante par interpolation linéaire :

$$Me = e_j + \frac{\frac{n}{2} - N_{j-1}\nearrow}{n_j} \times a_j$$

e_j : limite inférieure de la classe médiane ;
a_j : amplitude de la classe médiane ;
n_j : effectif de la classe médiane ;
$N_{j-1} \nearrow$: effectif cumulé inférieur à la classe médiane ;
n : taille de l'échantillon.

Exemple 6 : Soit la série statistique d'exemple 2 :
Le nombre d'effectif n = 13 impair, alors la médiane est la valeur d'ordre $\frac{n+1}{2} = \frac{13+1}{2} = 7$
La valeur d'ordre 7 est Me = 13.

Exemple 7 : Soit la série statistique d'exemple 4 :
Le nombre d'effectif n = 12 pair, alors la médiane est la moyenne arithmétique des valeurs d'ordre $\frac{n}{2} = \frac{12}{2} = 6$ et d'ordre $\frac{n}{2} + 1 = \frac{12}{2} + 1 = 7$
La valeur d'ordre 6 est 12
La valeur d'ordre 7 est 13; alors Me = $\frac{12+13}{2}$=12.5

Exemple 8 : Soit la série statistique d'exemple 5 :
On a $\frac{n}{2}$ = 30, la classe médiane est la classe qui leur effectif cumulé croissant supérieure ou égal 30, donc la classe médiane est : [65; 70[.
$e_j = 65$, $a_j = 5$, $n_j = 16$, $N_{j-1} \nearrow = 19$
Alors, Me = $65 + \frac{30-19}{16} \times 5 = 68.44$

4.4. Quartiles, Déciles et Centiles

Les quartiles, déciles et centiles sont des caractéristiques qui correspondent au même genre de préoccupation que la médiane. Il s'agit des valeurs de la variable qui correspondent aux effectifs cumulés :
$\frac{n}{4}$, $\frac{2n}{4}$, $\frac{3n}{4}$ Pour les quartiles; le 2e quartile est la médiane.
$\frac{n}{10}$, $\frac{2n}{10}$, …, $\frac{9n}{10}$ Pour les déciles; le 5e décile est la médiane.
$\frac{n}{100}$, $\frac{2n}{100}$, …, $\frac{99n}{100}$ Pour les centiles; le 50e centile est la médiane.

On les appelle caractéristiques de position, puisqu'elles permettent de placer les valeurs de la variable.

Exemple 9 : Soit la série statistique suivante :

On relevé la taille en centimètres des joueurs d'une équipe de Basket : 203, 187, 185, 206, 180, 188, 198, 195, 200, 195, 218, 210.

On ordonne la série dans l'ordre croissant

x_i	180	185	187	188	195		198	200	203	206	210	218
n_i	1	1	1	1	2		1	1	1	1	1	1
N_i ↗	1	2	3	4	5		7	8	9	10	11	12

L'ordre de la valeur de Q_1 est : $\frac{1}{4} \times 12 = 3$; donc $Q_1 = 187$.

L'ordre de la valeur de Q_3 est : $\frac{3}{4} \times 12 = 9$; donc $Q_3 = 203$.

5. Caractéristiques de dispersion

Les paramètres de dispersion nous renseignent sur la dispersion des valeurs autour de la valeur centrale de référence.

5.1.Etendue « E » :

Etendue d'une série statistique quantitative est la différence entre la plus grande valeur de x et la plus petite valeur. $E = x_{max} - x_{min}$

5.2.Variance « V(x) » :

Il est souvent intéressant de savoir si les valeurs sont très dispersées ou si elles sont proches de la moyenne. La variance est la caractéristique de dispersion la plus utilisée.

- Pour un variable discrète : $V(x) = \frac{1}{n} \sum_{i=1}^{k} n_i (x_i - \bar{x})^2$

- Pour une variable continue : $V(x) = \frac{1}{n} \sum_{i=1}^{k} n_i (c_i - \bar{x})^2$

5.3.Ecart-type « σ(x) » :

La racine carrée de la variance : $\sigma(x) = \sqrt{V(x)}$

5.4. Coefficient de variation « Cv » :

Si on veut comparer plusieurs séries statistiques ayant des moyennes très différentes, il vaut mieux se référer au coefficient de variation plutôt que l'écart-type. $Cv = \frac{\sigma_x}{\bar{x}} \times 100$

Remarque :

- Plus la valeur du coefficient de variation est faible, plus la dispersion autour de la moyenne est petite, plus la population est homogène.

- On considère qu'une distribution de données est homogène, lorsque Cv. est égal ou inférieur à 15%.

6. Caractéristiques de forme

Il existe deux mesures de forme qui caractérisent la forme des courbes représentant les distributions :
- Coefficient d'asymétrie
- Coefficient d'aplatissement

6.1. Coefficient d'asymétrie :

On se sert d'un coefficient pour mesurer l'asymétrie d'une distribution.
- Si le coefficient est nul ($S_k = 0$), alors il s'agit d'une distribution parfaitement symétrie. Et moyenne = médiane = mode

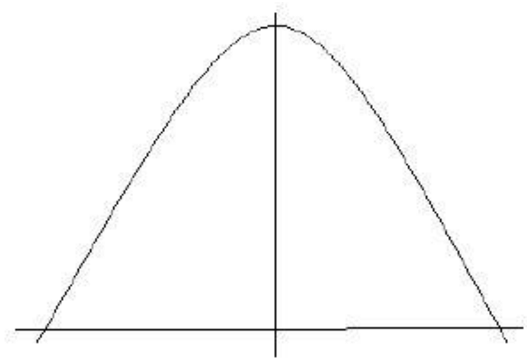


Figure 6 : Distribution parfaitement symétrie

Valeurs également distribués de part et d'autre de la valeur centrale.

- Si le coefficient est inférieur à 0, alors la distribution est du côté inférieur (étalée vers la gauche). Et moyenne < médiane < mode

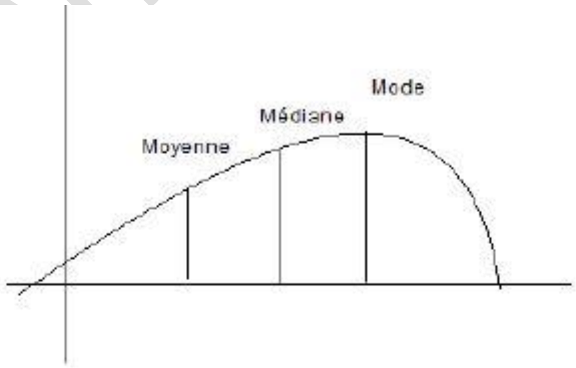


Figure 7 : Distribution étalée vers la gauche

La valeur la plus forte fréquence se situe à droite de la moyenne.

- Si le coefficient est supérieur à 0, alors la distribution est du côté supérieur (étalée vers la droite). Et moyenne > médiane > mode

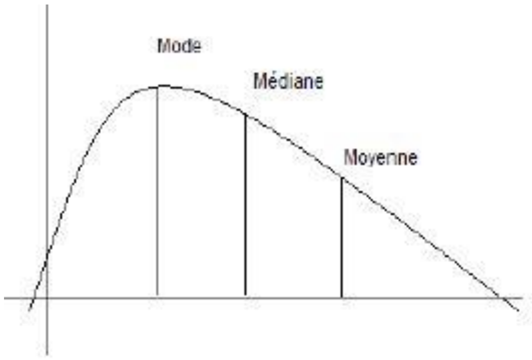


Figure 8 : Distribution étalée vers la droite

La valeur la plus forte fréquence se situe à gauche de la moyenne.

Calcul de coefficient :

- Coefficient d'asymétrie de Peason : On le note par S_p

$$S_p = \frac{3\,(\bar{x} - Me)}{\sigma_x} \text{ Ou } S_P = \frac{3\,(\bar{x} - Mo)}{\sigma_x}$$

- Si $S_P < 0$ distribution dissymétrique à gauche.
- Si $S_P = 0$ distribution symétrique.
- Si $S_P > 0$ distribution dissymétrique à droite.

- Coefficient d'asymétrie de Yule : On le note par S_Y

$$S_y = \frac{Q_3 + Q_1 - 2\,Q_2}{Q_3 - Q_1}$$

- Si $-1 \leq S_Y < 0$ distribution dissymétrique à gauche.
- Si $S_Y = 0$ distribution symétrique.
- Si $0 < S_Y \leq 1$ distribution dissymétrique à droite.

- Coefficient d'asymétrie de Fisher : On le note par S_F

$$S_F = \frac{M_3}{{\sigma_x}^3}$$

Ou $M_3 = \frac{1}{n} \sum n_i\,(x_i - \bar{x})^3$ est le moment centré d'ordre 3.
- Si $S_F < 0$ distribution dissymétrique à gauche.
- Si $S_F = 0$ distribution symétrique.
- Si $S_F > 0$ distribution dissymétrique à droite.

6.2.Coefficient d'aplatissement :

Pour mesurer l'aplatissement de la courbe, on utilise,

- Coefficient P_P de Peason : Le coefficient P_P de Peason basé sur le moment centré d'ordre 4.

$$P_P = \frac{M_4}{{\sigma_x}^4}$$

- $P_P > 3$ courbe leptocurtique ou hypernormale.
- $P_P = 0$ courbe normale.

- $P_P < 3$ courbe platicurtique ou hyponormale.

- Coefficient P_F de Fisher : Le coefficient P_F de Fisher est donné par

$$P_F = \frac{M_4}{{\sigma_x}^4} - 3$$

- $P_F > 0$ courbe leptocurtique ou hypernormale.
- $P_F = 0$ courbe normale.
- $P_F < 0$ courbe platicurtique ou hyponormale.

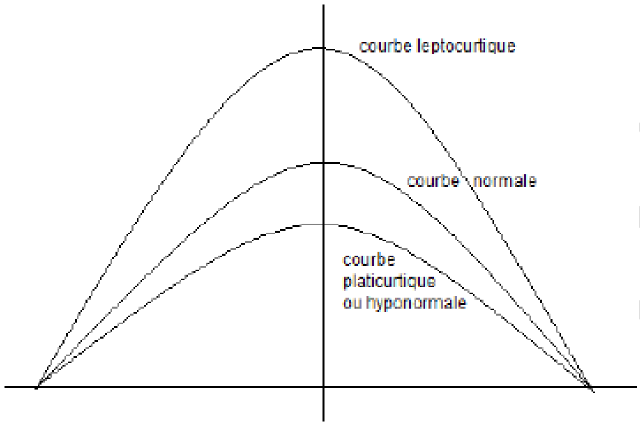


Figure 9 : Différentes courbes

Chapitre 2 : Séries statistiques à deux variables :

L'objectif est d'étudier une éventuelle relation entre des variables quantitatives x et y.

1. Nuage des points :

Un graphique qui traduise les deux séries statistiques à l'aide de diagramme à 2 dimensions.
Soit x et y deux variables statistiques numériques observées sur k individus. Dans un repère orthogonal$(o, \bar{\iota}, \bar{\jmath})$, l'ensemble des k points de coordonnées (x_i ; y_i) forme le nuage de points associé à cette série statistique.

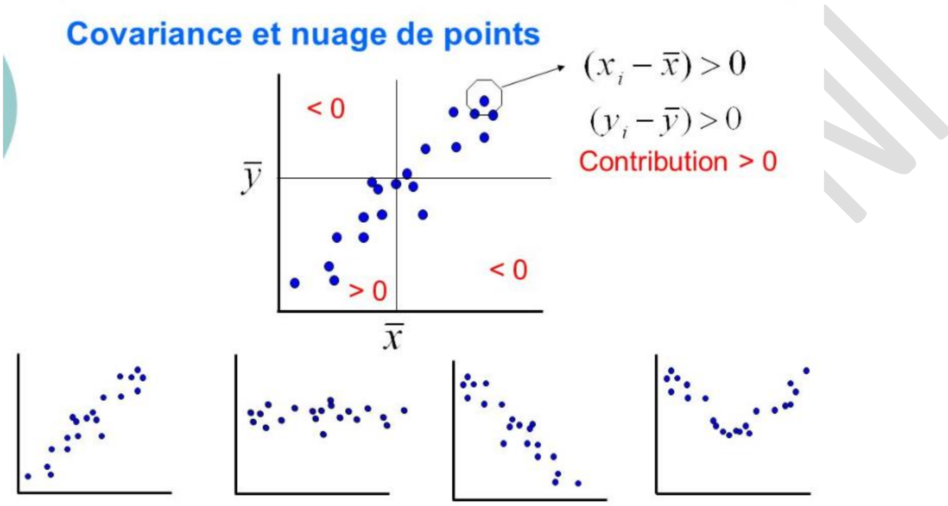


Figure 10 : Nuage de points

2. Tableaux de données (tableau de contingence)

- Tableau de contingence des effectifs :

- Variable discrètes : Paires de valeurs = (x_1, y_1) ; (x_2, y_2) ; (x_3, y_3) ; … ; (x_n, y_n)

Chaque effectif local n_{ij} correspond au nombre d'individus ayant l'abscisse x_i et l'ordonnée y_i.

X \ Y	y_1	y_2	…	y_j	…	y_q
x_1	n_{11}	n_{12}	…	n_{1j}	…	n_{1q}
x_2	n_{21}	n_{22}	…	n_{2j}	…	n_{2q}
. . .	. . .	. . .	… … …	. . .	… … …	. . .
x_i	n_{i1}	n_{i2}	…	n_{ij}	…	…
. . .	. . .	. . .	… … …	. . .	… … …	. . .
x_p	n_{p1}	n_{p2}	…	n_{pj}	…	n_{pq}

Tableau 6: Tableau de contingence « Variable discrètes »

- Variables continues : Ce sont les modalités centrales (x_1, x_2, x_3,…, x_p) et (y_1, y_2, y_3,…, y_p) des classes qui remplacent les modalités discrètes.

Chaque effectif local n_{ij} correspond au nombre d'individus dont leurs valeurs x appartiennent à la classe $[a_{i-1}, a_i[$ et leurs valeurs y appartiennent à la classe $[b_{j-1}, b_j[$.

X \ Y		$[b_0,b_1[$	$[b_1,b_2[$	…	$[b_{j-1},bj[$	…	$[b_{q-1},b_q[$
		y_1	y_2	…	y_j	…	y_q
$[a_0,a_1[$	x_1	n_{11}	n_{12}	…	n_{1j}	…	n_{1q}
$[a_1,a_2[$	x_2	n_{21}	n_{22}	…	n_{2j}	…	n_{2q}
. . .	. . .	. [2] . .	. . .	… … …	. . .	… … …	. . .
$[a_{i-1},a_i[$	x_i	n_{i1}	n_{i2}	…	n_{ij}	…	…
. . .	. . .	. . .	. . .	… … …	. . .	… … …	. . .
$[a_{p-1},a_p[$	x_p	n_{p1}	n_{p2}	…	n_{pj}	…	n_{pq}

Tableau 7: Tableau de contingence « Variable continue »

- Tableau de contingence des fréquences : On garde les mêmes tableaux précédents et on divise tous les effectifs (locaux marginaux) par l'effectif total n.

3. Distribution marginales et conditionnelles, Covariance :

3.1. Distribution marginales :

On ajoute au tableau de contingence les totaux en ligne et en colonne.

Y / X	y_1	y_2	…	y_j	…	y_q	Distribution marginale de Y n_{i*}	
x_1	n_{11}	n_{12}	…	n_{1j}	…	n_{1q}	n_{1*}	Sommes des n_{ij} de la ligne
x_2	n_{21}	n_{22}	…	n_{2j}	…	n_{2q}	n_{2*}	
. . .	. . .	. . .	… … …	. . .	… … …	. . .	. . .	
x_i	n_{i1}	n_{i2}	…	n_{ij}	…	…	n_{i*}	
. . .	. . .	. . .	… … …	. . .	… … …	. . .	. . .	
x_p	n_{p1}	n_{p2}	…	n_{pj}	…	n_{pq}	n_{p*}	
Distribution marginale de X n_{*j}	n_{*1}	n_{*2}	…	n_{*j}	…	n_{*q}	$n_{**} = n$	Effectif total
	Sommes des n_{1j} de la colonne						Effectif total	

Tableau 8: Distribution marginales « Variable discrète »

Y / X		$[b_0,b_1[$ y_1	$[b_1,b_2[$ y_2	… …	$[b_{j-1},bj[$ y_j	… …	$[b_{q-1},b_q[$ y_q	Distribution marginale de Y n_{i*}	
$[a_0,a_1[$	x_1	n_{11}	n_{12}	…	n_{1j}	…	n_{1q}	n_{1*}	Sommes des n_{ij} de la ligne
$[a_1,a_2[$	x_2	n_{21}	n_{22}	…	n_{2j}	…	n_{2q}	n_{2*}	
. . .	. . .	. . .	. . .	… … …	. . .	… … …	. . .	. . .	
$[a_{i-1},a_i[$	x_i	n_{i1}	n_{i2}	…	n_{ij}	…	…	n_{i*}	
. . .	. . .	. . .	. . .	… … …	. . .	… … …	. . .	. . .	
$[a_{p-1},a_p[$	x_p	n_{p1}	n_{p2}	…	n_{pj}	…	n_{pq}	n_{p*}	
distribution marginale de X n_{*j}		n_{*1}	n_{*2}	…	n_{*j}	…	n_{*q}	$n_{**} = n$	Effectif total
		Sommes des n_{1j} de la colonne						Effectif total	

Tableau 9 : Distribution marginales « Variable continue »

- En marge à droite (totaux en ligne) : la distribution de X : pour chaque indice i, l'effectif n_{i*} est le nombre total d'observations de la modalité x_i de X quelle que soit la modalité de Y. C'est-à-dire $n_{i*} = \sum_{j=1}^{q} n_{ij}$= Total de la ligne i.

Les p couples (x_i, n_{i*}) définissent la distribution marginale de la variable X.

- En marge en bas (totaux en colonne) : la distribution de Y : pour chaque indice j, l'effectif n_{*j} est le nombre total d'observations de la modalité y_j de Y quelle que soit la modalité de X. C'est-à-dire $n_{*j} = \sum_{i=1}^{p} n_{ij}$= Total de la colonne j.

Les q couples (y_j, n_{*j}) définissent la distribution marginale de la variable Y.

Remarque :

$$\sum_{i=1}^{p} n_{i*} = \sum_{j=1}^{q} n_{*j} = n$$

3.2. Distribution conditionnelles :

- La distribution des observations suivant les modalités de la variable Y sachant que la variable X prend la modalité x_i, est appelée distribution conditionnelle de Y pour X = x_i:
A la ligne i du tableau de contingence, on lit la distribution de la variable Y sachant que X = x_i, notée $Y|_{X=x_i}$.

- La distribution des observations suivant les modalités de la variable X sachant que la variable Y prend la modalité y_j, est appelée distribution conditionnelle de X pour Y = y_j:
A la colonne j du tableau de contingence, on lit la distribution de la variable X sachant que Y = y_j, notée $X|_{Y=y_j}$.

3.3. Covariance :

On appelle covariance de la série statistique double de variables x et y le nombre réel

$$\mathrm{Cov}(x, y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}$$

Remarque :

$$\mathrm{Cov}(x, x) = V(x)$$

4. Coefficient de corrélation linéaire :

Le coefficient de corrélation linéaire est un nombre permettant de déterminer l'intensité d'un lien linéaire entre deux variables quantitatives.
Le coefficient de corrélation linéaire d'une série statistique de variables x et y est le nombre r défini par : $\rho = \frac{Cov\,(x\,,y)}{\sigma_x \sigma_y}$

Remarque:

- Le coefficient de corrélation est une valeur qui n'a pas d'unité et qui est toujours comprise entre -1 et +1
- Un coefficient de corrélation linéaire est positif indique un lien linéaire positif, alors que, si ρ est négatif, le lien linéaire entre les deux variables est négatif.
- Plus la valeur de ρ est près de -1 ou +1 plus le lien linéaire ente les deux variables est fort.

5. Droite de régression et droite de Mayer

5.1. Droite de régression

Une droite de régression est la droite qui s'ajuste le mieux à un nuage de points présentant une corrélation linéaire. La droite de régression sert à faire des prévisions. On parle de corrélation linéaire lorsque les points d'un nuage ont tendance à s'aligner. Plus la tendance est forte, plus la corrélation linéaire est forte.

La droite D d'équation y = ax + b est appelée droite de régression de y en x de la série statistique si la quantité suivante est minimale : $S = \sum_{i=1}^{n}[y_i - (ax_i + b)]^2$

Pour définir les coefficients a et b; on développe S et on considère successivement comme un trinôme en b; puis, b étant déterminé, comme un trinôme en a: On trouve :

$$a = \frac{Cov\,(x\,,y\,)}{\sigma_x^2}$$

$$b = \bar{y} - a\,\bar{x}$$

5.2. Droite de Mayer

Cet ajustement consiste à déterminer la droite passant par deux points moyens du nuage de point.

6. Courbes de régression, couloir de régression et rapport de corrélation

6.1. Courbes de régressions

Une courbe de régression permet d'analyser la relation entre deux variables (variable explicative et variable expliquée) et de mettre en avant la nature de cette relation sans faire aucune hypothèse préalable sur la forme de celle-ci. Elle fait correspondre les valeurs de la première variable avec les moyennes conditionnées de la seconde. On peut donc à partir de deux variables X et Y construire deux courbes de régression :

- La courbe de X en Y : elle met en relation les valeurs de Y (y_i) et les moyennes conditionnelles de X.
- La courbe de Y en X : elle met en relation les valeurs de X (x_i) et les moyennes conditionnelles de Y.

Mise à part le fait de donner une image, une représentation de la forme de la liaison existant entre deux variables, les courbes de régression présentent des propriétés importantes :

.elles résument le mieux la silhouette du nuage de points puisqu'elles sont en moyenne les plus proches des points de ce nuage (la somme et les moyennes des carrés des distances entre les points du nuage et les courbes de régression sont minimales)

.elles se coupent en un point qui représente le point de gravité du nuage de points

(i.e. un point ayant pour coordonnées approximatives les moyennes marginales des deux variables).

6.2. Rapport de corrélation

Pour étudier la relation entre une variable qualitative et une variable quantitative, on décompose la variation totale en variation intergroupe(ou interclasse) et en variation intragroupe(ou intraclasse). Pour mesurer l'intensité de la relation, on peut calculer un paramètre appelé rapport de corrélation.

7. Ajustement fonctionnel

Lorsque l'on veut modéliser les données, la question qui se pose ensuite est de trouver
L'équation des courbes de régression. Les courbes de régression correspondent généralement à des fonctions compliquées mais on peut chercher à les approcher par des fonctions plus simples. Le principe général est de partir d'une forme de fonction connue et de chercher les paramètres qui ajustent le mieux possible les courbes obtenues aux courbes de régression.
Par exemple, si l'on part de l'idée que la courbe de régression, si l'on mettait de côté les erreurs de mesure, serait une droite, alors elle serait caractérisée par une équation de type

$$y = ax + b$$

Il faudrait alors identi.er les paramètres a et b pour connaître l'équation de la droite.
Si l'on imagine que la courbe de régression correspond à une parabole, l'équation cherchée serait de la forme :

$$y = ax^2 + bx + c$$

Et dans ce cas il faudrait trouver les valeurs des trois paramètres, a, b, et c.

7.1.Ajustement puissance

L'ajustement puissance est basé sur la courbe représentant par l'équation de type

$$y = ax^b$$

On remarque que

$$\ln y = a \ln x + \ln b$$

On pose V = ln y et U = ln x, on détermine l'équation de la droite de régression de v en u avec la méthode de Mayer ou la méthode des moindres carrés, l'équation obtenue est de la forme v = Au + B, on en déduit l'équation de la courbe de fonction puissance : $y = ax^b$ puisque A = b et B = ln a.

7.2.Ajustement exponentielle

L'ajustement exponentiel a une courbe de fonction exponentielle d'équation :

$$y = ab^x$$

On remarque que

$$y = x \ln b + \ln a$$

On pose z = ln y, on détermine l'équation de la droite de régression de z en x avec la méthode de Mayer ou la méthode des moindres carrés, l'équation obtenue est de la forme v = Au + B on en déduit l'équation de la courbe de fonction exponentielle

$$z = Ax + B$$

On en déduit l'équation de la courbe de fonction exponentielle :
$y = ab^x$ Puisque A = ln b et B = ln a

Exercices série 1 :

Exercice 1 :

Les notes d’examens de mathématiques d’un groupe de 40 étudiants sont données par la série suivante

15 ; 9 ; 3.5 ; 7.5 ; 0 ; 0 ; 9 ; 3.5 ; 10.5 ; 15 ; 15 ; 9 ; 9 ; 9 ; 7.5 ; 7.5 ; 3.5 ; 3.5 ; 15 ; 15 ; 15 ; 15 ; 3.5 ; 3.5 ; 9 ; 10.5, 10.5, 15 ; 10.5 ; 15 ; 15 ; 9 ; 15 ; 7.5 ; 7.5 ; 3.5 ; 9 ; 3.5 ; 9 ; 9

1. Quelle est la population, la variable étudier et le type de variable.
2. Dresser le tableau statistique de cette série (effectifs, effectifs cumulés, fréquences, fréquences cumulés) et représenter graphiquement les effectifs et les effectifs cumulés.
3. Donner la moyenne et la médiane de ce tableau.
4. Déterminer l’étendue et le mode de cette série.
5. Calculer la variance et en déduire l’écart type.

Exercice 2 :

L’or d’un contrôle de police sur l’autoroute, un agent a relevé les vitesses (Km/h) de 60 voitures. On a donc le tableau suivant

Les classes	c_i	n_i	N_i ↗	N_i ↘	f_i	F_i ↗	F_i ↘
[90, 100[		4					
[100, 110[		8					
[110, 120[		11					
[120, 130[		16					
[130, 140[		10					
[140, 150[		7					
[150, 160[		4					
Total							

Tableau 10: Exercice « variable quantitative continue »

1. Quelle est la population et le caractère étudier ? Quelle est sa nature ?
2. Compléter le tableau.
3. Trouver l’étendue et la classe modale.
4. Donner la moyenne et la médiane de ce tableau.
5. Calculer la variance et en déduire l’écart type.

6. Tracer l'histogramme ainsi que le graphe des fréquences cumulées croissantes.

Exercice 3 :

Le tableau suivant donne la répartition selon le groupe sanguin de 40 individus pris au hasard dans une population.

Groupes sanguins	A	B	AB	O
L'effectif	20	10	n_3	5

Tableau 11: Exercice « variable quantitative discrète »

1. Déterminer la variable statistique et son type.
2. Déterminer l'effectif des personnes ayant un groupe sanguin AB.
3. Donner toutes les représentations graphiques possibles de cette distribution.

Exercice 4 :

Pour les données suivantes

X	1	2	7	4	6
Y	5	4	1	3	2

Tableau 12: Exercice « série à deux variables »

1. Tracer le nuage de points.
2. Calculer la moyenne et l'écart type de la variable statistique X (et de Y).
3. Calculer la covariance des variables statistiques X et Y.
4. En supposant qu'il existe une corrélation linéaire entre X et Y, déterminer cette droite de corrélation.
5. Calculer le coefficient de corrélation. Conclusion ?

Partie B : Probabilités

Chapitre 1 : Analyse Combinatoire

L'objectif du chapitre 1 est de :

- ✓ Rechercher les possibilités de valeurs d'une variable aléatoire.
- ✓ Savoir le nombre de possibilités d'objets à former à partir des éléments de base.
- ✓ Compter le nombre de possibilités pour une expérience aléatoire.

Prérequis

Il est recommandé aux apprenants de connaître.

- Méthodes du raisonnement mathématique.
- Les ensembles, les relations et les applications.

Mots clés

Arrangements, Permutations, Combinaisons, Ordre, Répétition.

Introduction

L'analyse combinatoire est un ensemble de techniques mathématiques qui servent à compter le nombre d'éléments d'un ensemble fini (souvent de cardinalité grande).

1. Principe de multiplication

Permet de compter le nombre de résultats d'expériences qui peuvent se décomposer en une succession de sous-expériences.

Principe : Suppose qu'une expérience est la succession de m sous-expériences. Si la i-ème expérience à n_i résultats possibles pour i=1, …, m alors le nombre total de résultats possibles de l'expérience globale est : $n = n_1 \, n_2 \ldots n_m$

2. Arrangement (Expérience ordonnée)

Tirer P éléments parmi une totalité de n éléments et les trier, le résultat est A_n^p possibilités d'arrangement.

2.1. Arrangement sans répétition : $A_n^p = \frac{n!}{(n-p)!}$

Exemple 1 : 15 candidats se présentent à un concours. De combien de façons différentes seront classées les trois premiers ?

Réponse : Il s'agit d'ordonner p = 3 candidats parmi n = 15 sans répétition, donc on a un arrangement sans répétition. $A_n^p = \frac{n!}{(n-p)!} = \frac{15!}{(15-3)!} = \frac{15!}{12!} = 2730$

2.2. Arrangement avec répétition : $A_n^p = n^p$

Exemple 2 : Combien de mots de 3 lettres peuvent être formés dans l'alphabet français de 26 lettres ?

Réponse : Remarquons ici que : n = 26 et p = 3.

L'ordre est important dans un mot.

On peut répéter une lettre dans un mot.

On a donc un arrangement avec répétition. $A_n^p = n^p = 26^3 = 17576$ cas possibles.

3. Permutation (Expérience ordonnée)

Pour le cas particulier p = n, l'arrangement est appelé permutation notée P_n.

3.1. Permutation sans répétition : $P_n = A_n^n = \frac{n!}{(0)!} = n!$

Exemple 3 : Quel est le nombre de manières d'ordonner les trois chiffres {1; 2; 3} ?

Réponse : Le nombre de manières est $P_n = P_3 = 3! = 6$.

3.2. Permutation avec répétition : $P_n^{n_1,\ n_2,\ \dots,\ n_k} = \frac{n!}{n_1!n_2!\dots n_k!}$ $(n=\sum_{i=1}^{k} n_i)$

Exemple 4 : Quel est le nombre de manières d'ordonner les lettres suivantes {A ; A ; B ; B ; B ; C} ?

Réponse : Le nombre de manières est $P_n^{n_1,\ n_2,\ \dots,\ n_k} = P_6^{2,\ 3,\ 1} = \frac{6!}{2!3!1!} = 60.$

4. Combinaison (Expérience non ordonnée)

Tirer aléatoirement p éléments parmi une totalité de n éléments et former des objets aléatoires sans tenir compte de l'ordre des éléments combinés.

4.1. Combinaison sans répétition : $C_n^p = \frac{n!}{p!\,(n-p)!}$ avec $p \leq n$

Exemple 5 : On tire 5 cartes dans un jeu de 32 cartes, combien y a-t-il de résultats possibles ?

Réponse : $C_n^p = \frac{n!}{p!\,(n-p)!} = \frac{32!}{5!\,(32-5)!}$

4.2. Combinaison avec répétition : $C_{n+p-1}^p = \frac{(n+p-1)!}{p!\,(n-1)!}$

Exemple 6 : Les combinaisons avec répétition de 2 éléments pris dans {1; 2; 3} sont:

$C_{n+p-1}^p = \frac{(3+2-1)!}{2!\,(3-1)!} = 6$

Méthode de travail

Pour toute expérience il faut connaitre :

1. L'effectif total n et le nombre de tirage p
2. Appliquer le schéma suivant

L'expérience est
- Ordonnée
  - Avec répétition
  - Sans répétition
- Non ordonnée
  - avec répétition
  - sans répétition

Exercices série 2 :

Exercice 1 :
Dans tout l'exercice, on suppose qu'il n'y a pas de répétition.

1- Combien de nombres de 4 chiffres peut-on former à l'aide des 7 chiffres 1, 2, 4, 5, 6, 8, 9 ?
2- Combien de ces nombres sont inférieurs à 5000 ?
3- Combien de ces nombres sont pairs ?
4- Combien de ces nombres sont impairs ?
5- Combien sont des multiples de 5 ?

Exercice 2 :

Combien des mots possibles à former contenant 4 lettres parmi les 26 lettres de l'alphabet latin ? Un mot peut contenir des lettres répétées.

Exercice 3 :
Un cours de probabilités & statistique est suivi par 6 hommes et 4 femmes. Un examen a lieu, puis les étudiants sont classés selon leur note. On suppose exclu que deux étudiants obtiennent la même note.

1. Combien de classement peut avoir ?
2. Si les hommes sont classés entre eux uniquement et les femmes entre elles, combien de classements globaux peut-ont avoir ?

Exercice 4 :
Combien de mots possibles peut-on écrire à partir du mot " Canada" ?

Exercice 5 :
Le comité de planification d'un collège est constitué de 3 étudiants de première année, 4 de deuxième année, 5 de troisième année et 2 de dernière année. Un sous-comité de 4 étudiants comportant un représentant de chaque classe doit être choisi.
Combien peut-on former de sous-comités ?

Exercice 6 :
Soit un lot de 7 pièces dont 4 sont bonnes et 3 défectueuses.

1- Combien d'échantillons de 3 pièces peut-on réaliser ?
2- Combien parmi ces échantillons contiennent 3 bonnes pièces ?
3- Combien au moins contiennent une pièce bonne ?

Chapitre 2 : Introduction aux probabilités

L'objectif du chapitre 2 est de :

- ✓ Explorer les espaces probabilisés.
- ✓ Comprendre les notions d'univers, expériences aléatoires, événement.
- ✓ Utiliser des notions élémentaires de probabilité.
- ✓ Exprimer la probabilité d'un événement ou issues.
- ✓ Calculer des probabilités lors d'expériences aléatoires à une ou deux épreuves.
- ✓ Connaître et savoir utiliser les formules des probabilités équiprobables.

Prérequis

Il est recommandé aux apprenants de connaître :

- ➢ La théorie des ensembles.
- ➢ Les espaces et ses propriétés.

Mots clés

Expériences aléatoires, Evénement, Probabilité.

1. Algèbre des évènements

Une épreuve est une expérience aléatoire " dont l'issue est incertaine. Les résultats éventuels d'une épreuve font généralement appel au hasard. L'ensemble des résultats éventuels

(les résultats possibles, les éventualités) s'appelle ensemble fondamental

1.1.Mesure de probabilités

1.1.1. Définition

Soit P(Ω) l'ensemble des sous-parties de Ω.

La probabilité est une application de chaque événement des parties de Ω dans ℝ. Cette application notée P s'appelle **« mesure de probabilité ».**

$P: P(\Omega) \rightarrow \mathbb{R}$

$E_i \rightarrow P(E_i)$

Le nombre $P(E_i)$s'appelle probabilité de réalisation de l'événement E_i.

1.1.2. Définition

Une fois défini l'ensemble d'événements auxquels on s'intéresse, on va essayer de traduire par un nombre leurs probabilités de réalisation. $\forall A \subset \Omega$:

$$P(A) = \frac{nombrede\ cas\ favorables}{nombre\ de\ cas\ possibles}$$

Remarques :

$\forall A \subset \Omega$

- $P(A) \geq 0$
- $P(\Omega) = 1; (\sum_{i=1}^{n} p_i = 1)$
- $0 \leqslant P(A) \leqslant 1$

Si A et B sont incompatibles ($A \cap B = \emptyset$)alors : $P(A \cup B) = P(A) + P(B)$

1.1.3. Propriétés

- $P(\emptyset) = 0$
- $P(\bar{A}) = 1 - P(A)$
- $A \subset B \Rightarrow P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

1.1.4. Sigma-algèbre

Soit Ω l'ensemble fondamental lié à une épreuve, un sous-ensemble S de P(Ω) ($S \subset P(\Omega)$)est dit σ-algébre de partie de Ω si :

- $\Omega \in S$
- $A \in S \Rightarrow \bar{A} \in S$

- $A_1, A_2, \dots, A_n \in S \Rightarrow \cup_{i=1}^{\infty} A_i \in S$

2. Théorèmes généraux de probabilités

La théorie des probabilités constitue un cadre mathématique pour la description du hasard et de la variabilité, ainsi que pour le raisonnement en univers incertain.

L'objet de la théorie des probabilités est de fournir un formalisme mathématique précis, propre à décrire des situations dans lesquelles intervient le "hasard", c'est-à-dire des situations dans lesquelles un certain nombre de conditions étant réunies (les causes), plusieurs conséquences sont possibles (les effets) sans que l'on puisse a priori savoir laquelle sera réalisée.

Probabilités = "mathématisation du hasard".

La théorie des probabilités fournit des modèles mathématiques permettant l'étude d'expériences dont le résultat ne peut être prévu avec une totale certitude. En voici quelques exemples :

Expérience	**Résultat observable**
Lancer d'un dé	Un entier $k \in \{1,2,3,4,5,6\}$
Questionnaire à 100 question binaires	Suite Ω de 100 réponses ω $\in \{oui, non\}^{100}$
Lancer d'une pièce jusqu'à la première obtention de pile	Un entier : le temps d'attente du premier succès

Tableau 13: Quelques exemples sur les expériences

3. Espaces probabilisés

3.1. Expérience aléatoire et événements

3.1.1. Expérience aléatoire : On appelle expérience aléatoire (ou événement) toute expérience dont le résultat est régi par le hasard, lorsqu' on répète l'expérience dans les même conditions.

Exemple 7 : Le jet d'un dé [2] à 6 faces et l'observation de la face supérieure est une épreuve.

3.1.2. Espace des événements : (ou ensemble fondamental)

C'est l'ensemble de tous les résultats possibles d'une épreuve, on le notera généralement par Ω.

Exemple 8 : Le jet d'un dé : $\Omega = \{1,2,3,4,5,6\}$.

3.1.3. Evénements élémentaires et composés

Un événement élémentaire est un sous-ensemble de Ω à un seul élément $A = \{6\}$.

Un événement composé est un ensemble d'événements élémentaires $B = \{2,4,6\}$

3.2. Relations et opérations sur les événements :

Le tableau suivant représente la correspondance entre deux langages : langage ensembliste et langage probabiliste.

Notations	Vocabulaire ensembliste	Vocabulaire probabiliste
Ø	Ensemble vide	Événement impossible
Ω	Ensemble plein	Événement certain
Ω	Élément de Ω	Événement élémentaire
A	Sous-ensemble de Ω	Événement
$\omega \in A$	ω appartient à A	L'événement A est réalisé
A = B	Ensembles A et B sont égaux	Les événements A et B sont identique
$A \subset B$	A inclus dans B (A implique B)	A ne peut être réalisé sans que B le soit
$A \cup B$	Réunion de A et B (A ou B)	Un au moins des deux événements est réalisé
$A \cap B$	Intersection de A et B (A et B)	Les deux événement sont réalisés
$A \cap B = \emptyset$	A et B sont disjoints	Les événements A et B sont incompatibles
$\bar{A} = \Omega - A ou A^c$	Complémentaire de A dans Ω	L'événement A n'est pas réalisé

Tableau 14 : Relations et opérations sur les événements

Remarque :

1. Les opérations logiques (et, ou, négation) sur les événements peuvent bien sur faire intervenir plus de deux événements.

Si $A_1, A_2, \ldots, A_n$ sont des événements liés à une épreuve d'ensemble fondamental Ω:

- Réalisation de l'un au moins des événements A_i; $\cup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n$.
- Réalisation dans tous les événements A_i; $\cap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n$.
- Toutes les opérations sur les ensembles restent valables sur les modèles probabilistes (commutativité, associativité, distributivité, ...)

2.

$\overline{A \cup B} = \bar{A} \cap \bar{B} \Rightarrow \overline{\cup_{i=1}^{n} A_i} = \cap_{i=1}^{n} \bar{A}_i$.

$\overline{A \cap B} = \bar{A} \cup \bar{B} \Rightarrow \overline{\cap_{i=1}^{n} A_i} = \cup_{i=1}^{n} \bar{A}_i$.

3. Deux événements A et B sont dits incompatibles si la réalisation simultanée des événements A et B est impossible ($A \cap B = \emptyset$).
4. Système complet d'événements : une partiton de Ω est un système complet d'événement, Autrement dit, des événements $(A_i)_{i \in I}$ forment un système complet si :
   - $A_i \neq \emptyset; \forall\, i = 1,2, \ldots, n$.
   - $A_i \cap A_j = \emptyset; \forall\, i \neq j (i = \overline{1,n}; j = \overline{1,n})$ ⇒incompatibles.
   - $\cup_{1=i}^{n} A_i = \Omega$.

Exercices série 3 :

Exercice 1 :

Dans un examen de médecine, l'étudiant doit choisir 3 questions parmi 80 questions.

1. Combien de possibilité d'examen on peut avoir ?
2. Un étudiant a révisé seulement 50 questions

a) Quelle est la probabilité que l'étudiant tombe sur 3 questions qu'il a révisées.
b) Quelle est la probabilité que l'étudiant tombe sur 2 questions qu'il a révisées.
c) Quelle est la probabilité que l'étudiant tombe sur une question qu'il a révisée.
d) Quelle est la probabilité que l'étudiant ne tombe sur aucune question révisée.

Exercice 2 :

Le bloc opératoire est alimenté en électricité quand l'un ou l'autre des générateurs G_1 ou G_2 fonctionne. Soient les événements.

G_1 « le générateur G_1 fonctionne »

G_2 « le générateur G_2 fonctionne »

Exprimer à l'aide de G_1, G_2 et le vocabulaire ensembliste des opérations de base les événements suivants :

1. Le bloc opératoire est alimenté en électricité.
2. Le bloc n'est pas alimenté en électricité.
3. Seul le générateur G_1 fonctionne.
4. Au moins l'un des deux générateurs fonctionne.
5. Aucun des deux générateurs ne fonctionne.

Si la probabilité d'existence de l'électricité au bloc est égale à 0.9, quelle est la probabilité que les deux générateurs tombent simultanément en panne ?

Chapitre 3 : Conditionnement et indépendance

L'objectif du chapitre 3 est de :

- ✓ Définir l'indépendance de deux événements.
- ✓ Appréhender les probabilités conditionnelles.
- ✓ Définir la formule de Bayes.

Prérequis

Il est recommandé aux apprenants de connaître :

- ➢ La théorie des ensembles.
- ➢ Les espaces et ses propriétés.

Mots clés

Indépendance, Probabilités conditionnelles, Bayes

1. Probabilités conditionnelle

Définition

Soit (Ω, S, P) un espace probabilisé, soient A, B deux événements, tels que $P(B) > 0$, La probabilités conditionnelle de A sachant B est définie par: $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A.B)}{P(B)}$.

Remarque

- $P(A/B) \neq P(B/A)$
- $P(A/A) = 1$

Exemple 9 :

On lance un dé dont les faces sont numérotées de 1 à 6, les faces 3 et 6 sont blanches, les faces 1, 2 et 4 sont rouges, la face 5 est bleue. On suppose que le dé est truqué et on a les probabilités des événements élémentaires suivantes :

P({1}) = 0.1 ; P({2}) = P({3}) = P({4}) =0.2 ; P({5}) =P({6}) = 0.15.

Quelle est la probabilité d'obtenir une face avec un numéro pair sachant que la face est blanche?

Notons les événements :

N : « numéro pair »

B : « face blanche »

On cherche la probabilité P(N/B).

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

On a : $P(B) = P(\{3\} \cup \{6\}) = P(\{3\}) + P(\{6\}) = 0.2 + 0.15 = 0.35$

$P(N \cap B) = P(\{6\}) = 0.15$

D'où $P(N/B) = \frac{0.15}{0.35} = 0.42$

2. Evénements indépendants

Définition

Dans un espace probabilisé (Ω, S, P), deux événements A et B sont dits indépendants si :

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

Complément

A et B indépendants $\Leftrightarrow A et \bar{B} indépendants.$

$\Leftrightarrow \bar{A} et B indépendants.$

$\Leftrightarrow \bar{A} et \bar{B} indépendants.$

3. Formule des probabilités totales

Soit $\{A_i; i \in I\}$un système complet d'événements, tous de probabilité non nulle. Soit B un événement. Alors :

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

Cette formule permet de calculer la probabilité d'un événement B en le décomposant suivant un système complet d'événements.

4. Formule de Bayes

Soit (Ω, S, P) un espace probabilisé et $A_1, A_2, \ldots, A_n$une partition de Ω (système complet de Ω), si B est un autre événement alors :

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

Exemple 10 :

Un laboratoire d'analyse chimique reçoit un lot de tube à essai, Ces tubes sont fournis par trois sociétés différentes $A_1, A_2 et A_3$dans les proportions suivantes: 50%, 30% et 20%.

2% des tubes fabriqués par A_1, 3% de ceux fabriqués par A_2et 4% de ceux fabriqués par A_3présentent des défauts. On choisit au hasard un tube à essai dans le lot reçu.

1. Quelle est la probabilité qu'il soit défectueux ?
2. Sachant que le tube choisi est défectueux, quelle est la probabilité qu'il provienne de la société A_1?

A_1, A_2, A_3 événements provient de l'usine A_1, A_2, A_3et D événement tube défectueux alors :

$$P(A_1) = 0.5; P(A_2) = 0.3; P(A_3) = 0.2$$
$$P(D/A_1) = 0.002; P(D/A_2) = 0.003; P(D/A_3) = 0.004$$

1. La probabilité qu'il soit défectueux est : (probabilité totale)

$$P(D) = \sum_{i=1}^{3} P(A_i).P(D/A_i) = 0.5(0.002) + 0.003(0.3) + 0.004(0.2) = 0.0027$$

2. La probabilité qu'il provienne de la société A sachant que le tube choisi est défectueux est : (la formule de Bayes)

$$P(A_1/D) = \frac{P(A_1).P(D/A_1)}{\sum_{i=1}^{3} P(A_i).P(D/A_i)} = \frac{0.5(0.002)}{0.0027} = 0.37$$

Exercices série 4 :

Exercice 1 :

Une classe composé de 30 élèves tel que :

10 élèves sont fort en maths, 12 élèves sont fort en physique, 6 élèves sont fort en maths et physique.

Parmi les élèves qui sont fort en physique, calculer la probabilité de trouver un élève fort en maths.

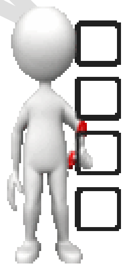
Exercice 2 :

Un appareil peut être monté avec des pièces de haute qualité (40% des cas) et avec des pièces ordinaires (60% des cas). Dans le premier cas, sa fiabilité (probabilité de fonctionnement) sur une durée t est égale à 0.95 ; dans le deuxième cas, elle est de 0.7.
L'appareil a été soumis à un essai et s'est avéré fiable (fonctionnement sans défaillance sur la durée de référence).
Déterminer la probabilité que l'appareil soit composé de pièces de haute qualité.

Chapitre 4 : Variables aléatoires

L'objectif du chapitre 4 est de :

- ✓ Découvrir les variables aléatoires.
- ✓ Apprendre à manipuler les variables aléatoires.
- ✓ Définir et calculer dans des cas simples la fonction de distribution d'une variable aléatoire discrète.
- ✓ Appréhender les Variables Aléatoires Continues (VAC).
- ✓ Définir la fonction de densité de probabilité d'une variable aléatoire continue.
- ✓ Calculer l'espérance, variance et l'écart type.



Prérequis

Il est recommandé aux apprenants de connaître :

- ➢ Les calculs d'intégrales.
- ➢ Fonctions : limites et continuité.

Mots clés

Variables aléatoires, Densité, Espérance, Variance, Ecart type.

1. Définitions et propriétés

1.1.Introduction

Une variable aléatoire VA est une fonction définie sur l'ensemble des résultats possibles d'une expérience aléatoire, telle qu'il soit possible de déterminer la probabilité pour qu'elle prenne une valeur donnée ou qu'elle prenne une valeur dans un intervalle donné. À l'origine, une variable était une fonction de gain, qui représentait le gain obtenu à l'issue du résultat d'un jeu. Par exemple, lorsqu'on lance deux dés, on s'intéresse à la somme des chiffres égale à 7. les couples (1,6), (6,1), (5,2), (2,5), (3,4), (4,3) font que la somme est égale à 7. Comme cette somme dépend des valeurs aléatoires, il s'agit d'**une variable aléatoire**.

1.2.Définition

Soit (Ω, S, P) un espace de probabilités, une variable aléatoire X sur un ensemble fondamental Ω est une fonction de Ω sur ℝ

$X: \Omega \to \mathbb{R}$
$\omega \to X(\omega)$

telle que l'image inverse de chaque intervalle de ℝ par X soit un événement de S.

$$I = [a, b] \subset \mathbb{R}$$

$X^{-1}([a, b]) = \{\omega \in \Omega / X(\omega) \in [a, b]\}$

X est appelée **variable aléatoire réelle**.

Exemple 11 :

Donnons un exemple simple du lancer de deux dés, ce qui est équivalent à lancer deux fois un dé. Une première variable aléatoire X_1donne le résultat du premier lancer, une deuxième X_2donne le résultat du deuxième lancer, c'est-à-dire $X_1(\omega) \in \{1,2,3,4,5,6\}$et $X_2(\omega) \in \{1,2,3,4,5,6\}$ que l'on note plus simplement $X_1 \in \{1,2,3,4,5,6\}$ et $X_1 \in \{1,2,3,4,5,6\}$.

Il est possible de s'intéresser à la somme des deux résultats, qui peut être notée par une variable aléatoire $S \in \{2,3,4,5,6,7,8,9,10,11,12\}$.

1.3.Types de VA

Nous allons distinguer deux groupes de variables aléatoires : Les VA quantitatives et les VA qualitatives.

- Variables quantitatives :

Ici, nous retrouverons toutes les VA numériques. On distingue deux types de variables quantitatives :

- Variables discrètes VAD: Les valeurs sont discrètes, ce sont des nombres entiers (fini ou infini dénombrable).

$X \in D_x = \{x_1, x_2, \dots, x_n\}$(VAD finie)
$X \in D_x = \{x_1, x_2, \dots\}$(VAD infinie)

- Variables continues VAC: Toutes les valeurs comprises dans un intervalle défini sont possibles. $X \in [a, b] \subset \mathbb{R}$

- Variable qualitatives :
- Variables nominales ou lexicales : Les différentes modalités de la variable ne peuvent être ordonnées.
- Variables ordinales : Les modalités de la variable possèdent la propriété d'ordre.

1.4. Probabilités d'une VA

Soit X une VA définie par une application : $\Omega \to \mathbb{R}$pouvant prendre les valeurs dans $D_x = \{x_1, x_2, \ldots, x_n\}$.

Par définition la probabilité pour que X = x est la probabilité des éléments de Ω ayant pour image la valeur x dans l'application. $X = \{x_1, x_2, \ldots, x_n\}$; $P = \{p_1, p_2, \ldots, p_n\}$,

avec $p_i = P(X = x_i): i = \overline{1, n}$,

on a $p_i = P(X = x_i) = P(\{\omega \in \Omega / X\,(\omega) = x_i\})$.

2. Loi de probabilité d'une VA

2.1. Loi de probabilité d'une VAD

Soit x une VAD (finie). On appelle loi de probabilité de x la donnée :

- De l'ensemble des valeurs possibles : $D_x = \{x_1, x_2, \ldots, x_n\}$.
- De la probabilité de chaque valeur : $P = \{p_1, p_2, \ldots, p_n\} / p_i = P(X = x_i)$.

On désignera par f la loi de probabilité de x, ce qui s'écrit : $\forall x_i, f(x_i) = P(X = x_i)$.

La fonction f s'appelle fonction de distribution de la VAD x.

Remarque

$\forall i: f(x_i) = P(X = x_i) \geq 0$

- $\sum_{i=1}^{n(\infty)} f(x_i) = \sum_{i=1}^{n(\infty)} P(X = x_i) = 1$
- $P(a \leqslant x \leqslant b) = \sum_{x_i \in [a,b]} f(x_i)$

Exemple 12 :

Soit x la VA donnée par le tableau suivant :

x_i	$P(x_i)$
0	1/8
1	3/8
2	3/8
3	1/8

Tableau 15 : Exemple sur les probabilités

$P(0 \leqslant x \leqslant 2) = \sum_{x \in [0,2]} f(x_i) = f(0) + f(1) + f(2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$.

2.2.Loi de probabilité d'une VAC

Soit x une VAC. On appelle loi de probabilité de x ou densité de probabilité la fonction f telle que :

- $\forall x \in \mathbb{R}; f(x) \geq 0$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

Remarque

Soit X VAC de densité f :

$$P(a \leqslant x \leqslant b) = \int_{a}^{b} f(x)dx$$

C'est à dire l'aire sous la courbe de f entre a et b.

Exemple 13 :

Soit f une fonction densité d'une VAC x définie par : $f(x) = kx, Si 0 \leq x \leq 1$
$0 Sinon$

1. Déterminer la valeur de k pour que f soit une densité.
2. Calculer $P(\frac{1}{2} \leqslant x \leqslant 1)$.

3. Fonction de répartition

3.1.Fonction de répartition d'une VAD

Définition

Soit x une VAD (finie ou infinie), On appelle fonction de répartition de x définie sur (Ω, S, P) la fonction $F_x$définie sur $\mathbb{R}$par :

$$F_x : \mathbb{R} \to [0,1]$$
$$x \to F_x(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i)$$

3.2.Fonction de répartition d'une VAC

Définition

Soit x une VAC de densité f, On appelle fonction de répartition de x la fonction $F_x$définie par :

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

Complément

- Le graphe de F_x est une fonction continue.
- $P(a \leqslant x \leqslant b) = F_x(b) - F_x(a)$(cas ou $F_x$est continue)

Remarque

Si $F_x$est continue sur $\mathbb{R}$et si f est continue, $\forall x \in \mathbb{R}$, alors $F_x$est dérivable et on a : $F_X'(x) = f(x)$.

4. Espérance mathématique (moyenne)

Soit x une VA de fonction de distribution f, l'ensemble mathématique de x que l'on note par E(x) est définie par :

- Pour VAD : $E(x) = \sum_{i=1}^{n(\infty)} x_i f(x_i) = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_n f(x_n)$.
- Pour VAC : $E(x) = \int_{-\infty}^{+\infty} x f(x) dx$.

5. Variance et moments

5.1.Variance

$Var(x) = E(x^2) - (E(x))^2$

- Pour VAD :$E(x^2) = \sum_{i=1}^{n} x_i^2 f(x_i)$.
- Pour VAC : $E(x^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$.

5.2.Moment non centré d'ordre k

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

5.3.Moment centré d'ordre k

$$M_k = \mathrm{E}[(x - m_1)]^k = \int_{-\infty}^{+\infty} (x - m_1)^k f(x) dx$$

Exercices série 5 :

Exercice 1 :

Soit x une variable aléatoire de probabilité définie pour $\lambda > 0$ par :

$$P(x = i) = \frac{C\, \lambda^i}{i!}, \forall\, i = 0, 1, 2, \ldots$$

Calculer P(x = 0) et P(x > 2).

N-B : C est une constante à déterminer. Pour cela, on utilise le développement en série de la fonction : $x \rightarrow e^x = \sum_{i=0}^{+\infty} \frac{x^i}{i!}$

Exercice 2 :

Soit f la fonction définie pour $\lambda > 0$ par :

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

1. La fonction f est-elle une densité de probabilité d'une variable aléatoire x ?
2. Calculer $P(x > 1)$.
3. Si $\lambda = 1$, déterminer la fonction de répartition F de la variable aléatoire, calculer E(x), V(x), et σ_x.

Exercice 3 :

La fonction de répartition d'une VA x est donnée par :

$$F(x) = \begin{cases} 0, & x < 2 \\ (x-2)^2, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

1. Calculer $P(1 \leq x \leq 5/2)$; $P(5/2 \leq x \leq 7/2)$; $P(5/2 \leq x \leq 3)$.
2. Déterminer la densité de probabilité de la VA x.

Exercice 4 :

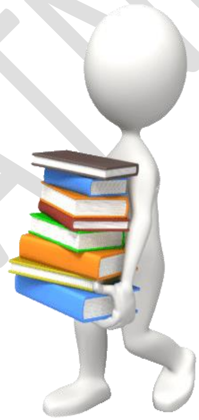
On jette deux dés symétriques. Soient x_1 et x_2 les VAs correspondant à la somme et le produit des points obtenus.

1. Construire les suites de répartitions de x_1 et x_2.
2. Calculer $E(x_1)$ et $E(x_2)$.

Chapitre 5 : Lois de probabilité discrètes et continues usuelles

L'objectif du chapitre 5 est de :

- ✓ Décrire les variables aléatoires sous la forme d'une expérience type
- ✓ Analyser cette expérience en détail pour pouvoir déduire les principales caractéristiques de toutes les expériences aléatoires
- ✓ Apprendre à utiliser quelques lois usuelles discrètes.
- ✓ Familiariser avec quelques lois usuelles continues.
- ✓ Calculer et démontrer l'espérance et la variance de chaque loi



Prérequis

Il est recommandé aux apprenants de connaître :

- ➢ Les calculs d'intégrales.
- ➢ Fonctions : limites et continuité.

Mots clés

Bernouli, Binomiale, Poisson, Uniforme, Exponentielle, Normale

1. Lois de probabilités usuelles discrètes

1.1. Loi de Bernoulli:

Définition : On dit que la variable aléatoire réelle X suit la loi de Bernoulli de paramètre p si ne prenne que les valeurs 0 et 1 tel que P (x = 1) = p; P (x = 0) = 1-p = q.
On note par: X ~ B (p)

Caractéristiques:
E (x) = p; V(x) = pq; σ_x = $\sqrt{pq}$

Preuve : On a $E(X) = \sum_{i=1}^{2} x_i P_i = 1 \times P(X = 1) + 0 \times P(X = 0) = p$
Et $E[(X - E(X))^k] = E[(X - p)^k] = (-p)^k q + (1 - p)^k p$
Soit $E[(X - E(X))^k] = (-p)^k q + (1 - p)^k p$
En particulier $V(X) = E[(X - E(X))^2] = (-p)^2 q + (1 - p)^2 p$
$= p^2q + (1-p)^2 p = pq (p+q) = pq.$

1.2. Loi Binomiale:

Définition: Une variable aléatoire X suit une loi de Binomiale de paramètres n et p si elle admet pour densité de probabilité suivante:

$$k = 0, 1, 2, \dots, n; P(X = k) = C_n^k p^k q^{n-k}, \mathrm{q} = 1 - \mathrm{p}$$

On note par: X ~ B (n, p) :
Une variable de Bernoulli est un cas particulier d.une variable binomiale.

X ~ B (1, p)

Caractéristiques:
E (x) = np; V(x) = np(1-p) = npq; σ_x = $\sqrt{npq}$

Preuve : On a

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

Ce qui donne, en utilisant la formule du binôme de Neweton

$$\sum_{k=0}^{n} f(x) = (p + (1 - p))^n = 1$$

Ainsi, si on prend, pour tout k entier inférieur ou égal à n;

P (x = k) = f (x)

On définit une variable aléatoire discrète et sa loi de probabilité.
La définition de E (x) donne

$$E(X) = \sum_{k=0}^{n} kC_n^k p^k (1 - p)^{n-k} = \sum_{k=1}^{n} kC_n^k p^k (1 - p)^{n-k}$$

Comme on a $C_n^k = \frac{n}{k} C_{n-1}^{k-1}$, il vient

$$E(X) = np\sum_{k=1}^{n} C_{n-1}^{k-1}p^{k-1}(1-p)^{n-k}$$

On peut donc écrire en utilisant de nouveau la formule de binôme

$$E(X) = np\,[p + (1-p)]^{n-1} = np$$

On a de même

$$E[X(X-1)] = \sum_{k=0}^{n} k(k-1)\,C_n^k p^k (1-p)^{n-k}$$

Comme on a $C_n^k = \frac{n(n-1)}{k(k-1)}\,C_{n-2}^{k-2}$

Il vient

$$\begin{aligned} E[X(X-1)] &= n(n-1)p^2\sum_{k=0}^{n} C_{n-2}^{k-2}p^{k-2}(1-p)^{n-k} \\ &= n(n-1)p^2[p + (1-p)]^{n-2} \\ &= n(n-1)p^2 \end{aligned}$$

D'où

$$E(X^2) = E[X(X-1)] + E[X] = n(n-1)p^2 + np$$

Ensuite

$$V(X) = E(X^2) - [E(X)]^2 = np(1-p) = npq$$

Exemple 14 : On jette 10 fois de suite une pièce de monnaie équilibrée. Quelle est la probabilité d'avoir un total de 8 piles? Soit X une v.a. qui associe à ces 10 lancers de pièce le nombre de pile.

On a X ~ B $(10, \frac{1}{2})$

$$P(X = 8) = C_{10}^{8}\left(\frac{1}{2}\right)^{8}\left(\frac{1}{2}\right)^{10-8} = 0.0439$$

1.3.Loi de Poisson

Définition : Une variable aléatoire X suit une loi de Poisson de paramètres λ $(\lambda > 0)$ si elle admet pour densité de probabilité suivante:

$$P(k,\lambda) = P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, k = 0, 1, 2, \ldots$$

Caractéristiques:

E (X) = λ; V ar (X) = λ; $\sigma_X = \sqrt{\lambda}$

Preuve : On a

$$\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}\,e^{\lambda} = 1$$

Alors la relation P (X = k) = f (x) définit une variable aléatoire discrète X et sa loi de probabilité.

La définition de E (X) donne

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

Comme $\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{\lambda}$

il vient E (X) = λ

On a de même

$$E[(X-1)] = \sum_{k=0}^{\infty} k(k-1) \frac{e^{-\lambda}\lambda^k}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}$$

Comme $\sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = e^{\lambda}$, on obtient $E[(X-1)] = \lambda^2$

Par conséquent $E(X^2) = E[X(X-1)] + E[X] = \lambda^2 + \lambda$

IL en résulte que $V(X) = E(X^2) - [E(X)]^2 = \lambda$

2. Lois de probabilités usuelles continues

2.1.Loi uniforme:

Définition : Une variable aléatoire continue X suit une loi uniforme sur [a, b] si elle admet pour densité de probabilité la fonction f définie par:

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

On note par: $X \sim U_{[a,b]}$

Caractéristiques:

$$E(X) = \frac{a+b}{2}, V(X) = \frac{(b-a)^2}{12}, \sigma_X = \sqrt{\frac{(b-a)^2}{12}}$$

Preuve : Par définition de E (X)

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

De même

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_a^b \frac{x^2}{b-a} dx = \frac{b^2+ab+a^2}{3}$$

Comme $V(X) = E(X^2) - [E(X)]^2 = \frac{(b-a)^2}{12}$

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

2.2.Loi exponentielle:

Définition : Une variable aléatoire continue X suit une loi exponentielle de paramètre λ si elle admet pour densité de probabilité la fonction f définie par :

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

Caractéristiques:

$$E(X) = \frac{1}{\lambda}, V(X) = \frac{1}{\lambda^2}, \sigma_X = \frac{1}{\lambda}$$

Preuve : Par définition de E (X)

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^{+\infty} \lambda x e^{-\lambda x} dx = \lim_{x\to+\infty} \int_0^x \lambda t e^{-\lambda t} dt$$

L'intégration par parties donne

$$\int_0^x \lambda t e^{-\lambda t} dt = -x e^{-\lambda x} + \frac{1}{\lambda}(1 - e^{-\lambda x})$$

D'où $E(X) = \frac{1}{\lambda}$

2.3.Loi normale (Loi normale centrée réduite ou loi de Gauss):

Définition : Une variable aléatoire continue X suit une loi normale centrée réduite notée par N (0, 1), si elle admet pour densité de probabilité la fonction f définie sur R par: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

f étant une densité, la surface comprise entre l'axe des abscisses et la courbe y = f(x) a une aire finie égal à 1:

La fonction de répartition de X: F(x) = $F(x) = \int_{-\infty}^{x} f(t)dt$

Caractéristiques:

E (X) = 0; V(X) = 1; $\sigma_x = 1$

Preuve : D'après la définition de

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

La fonction $xe^{-\frac{x^2}{2}}$ est impair, on a, alors E (X) = 0.

De même

$$V(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx$$

La fonction $x^2 e^{-\frac{x^2}{2}}$ est paire, on a, alors

$$V(X) = 2\int_0^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{\pi}} \int_0^{+\infty} x^2 e^{-x^2} dx$$

Faisons le changement de variables par les coordonnées polaires, on obtient facilement

$$V(X) = \frac{2}{\sqrt{\pi}} \times \frac{\sqrt{\pi}}{2} = 1$$

Exercices série 6

Exercice 1:

1. Montrer que pour n assez grand, p assez petit et np=cte, on peut faire l'approximation de la loi binomial par la loi de Poisson.
2. La probabilités qu'une pièce extraite d'un lot de 200 pièces, soit défectueuse est 0.02. On désigne par X la variable aléatoire qui compte le nombre de pièces défectueuses sur le lot.

Déterminer la loi exacte de X et calculer la probabilité que sur les 200 pièces, il y ait :

- Pièces défectueuses.
- Au plus 5 pièces défectueuses.

Peut-on approximer la loi exacte de X, par une autre loi ? (Justifier). Si oui, calculer alors les probabilités : $P(X = 5)$ et $P(X \leq 5)$.

Exercice 2:

Une central téléphonique reçoit en moyenne 300 appels par heure. On suppose que le nombre d'appels pendant un intervalle de temps suit une loi de Poisson.

Calculer la probabilité que durant deux minutes la centrale reçoit :

- Trois appels.
- Au moins un appel.
- Au plus deux appels.

Exercice 3:

Une variable aléatoire T suit une loi exponentielle de paramètre $\lambda > 0$.

- Trouver le paramètre de cette loi sachant que $P(T \leq 70) = 0.05$.
- Calculer $P(T > 30)$.

Bibliographie

[1] D. Dacunha-Castelle et M. Duflo. *Probabilités et statistiques: Problèmes à temps fixe.* Masson, 1982.

[2] J.-F. Delmas. *Introduction au calcul des probabilités et à la statistique.* Polycopié ENSTA. 2008.

[3] F. Dominique et A. Fuchs, *Calcul de probabilités, cours et exercices corrigés.* deuxième édition. 1998.

[4] W. Feller. *An Introduction to Probability Theory and its Applications.* Volume 1.Wiley & Sons, Inc. 3rd edition. 1968.

[5] G. Grimmett, D. Stirzaker. *Probability and Random Processes.* Oxford University Press. 2nd edition. 1992.

[6] J. Jacod and P. Protter. *Probability Essentials.* Springer. 2000. 78

[7] A. Montfort. *Cours de statistique mathématique.* Economica. 1988.

[8] A. Montfort. *Introduction à la statistique.* Ecole Polytechnique. 1991.

[9] A. Ouaoua. *Probabilités et Statistiques : Cours et exercices corrigés.* Université de Skikda. 2020.