

« La probabilité de réussir la mise sur orbite d'une fusée est d'une chance sur un million. Dépêchons-nous de rater 999 999 lancements » (Jacques Rouxel)

Cours de Probabilités

Licence – 2ème année
(Agronomie)

Enseignant : M. Walguen OSCAR

Table des matières

Chapitre 1.- Éléments de Probabilités	4
1.1.- Les vocabulaires de bases des probabilités	4
1.1.1.- Vocabulaire des évènements	4
1.1.2.- Évènement impossible, évènement certain.....	5
1.1.3.- Évènement contraire	5
1.1.4.- Opérations sur les évènements.....	6
1.1.5.- Parallélisme entre la terminologie ensembliste et la terminologie probabiliste	6
1.1.6.- Espaces probabilisables.....	7
1.2.- Modéliser une expérience aléatoire	7
1.2.1.- Définition d'une probabilité - Espace probabilisé.....	8
1.2.2.- Calcul d'une probabilité	8
1.2.3.- Différentes conceptions des probabilités.....	8
1.2.4.- Définition axiomatique d'une probabilité sur un univers Ω	10
1.2.5.- Probabilités sur un univers fini	11
1.2.6.- Équiprobabilité sur un univers fini.....	12
1.2.7.- Probabilité sur un univers infini dénombrable	12
1.2.8.- Espace probabilisé.....	13
1.3.- Système complet d'évènements.....	13
1.4.- Rappels sur l'Analyse Combinatoire	13
1.4.1. Indications générales	13
1.4.2. Situations sans répétition.....	15
1.4.3. Situations avec répétition.....	18
1.4.4. Tableau récapitulatif des formules de dénombrement.....	19
1.4.5. Triangle de Pascal - Binôme de Newton	19
1.4.6. Exemples de probabilités combinatoires.....	20
1.5.- Probabilités conditionnelles.....	21
1.6.- Théorème de la multiplication.....	22
1.7.- Diagramme en arbre.....	23
1.8.- Probabilités composées et Théorème de Bayes.....	24
1.9.- Évènements indépendants pour une probabilité P	28
Chapitre 2.- Généralités sur les variables aléatoires	31

2.1.	Définition.....	31
2.2.	Les types de variables aléatoires.....	31
2.3.	Variables aléatoires discrètes.....	32
2.3.1.	Loi de probabilité ou fonction de distribution.....	33
2.3.2.	Fonction de répartition.....	34
2.3.3.	Variables aléatoires discrètes indépendantes.....	37
2.3.4.	Espérance - variance - écart-type d'une variable aléatoire discrète - covariance de deux variables aléatoires discrètes.....	37
2.3.5.	Variable centrée réduite associée à X	40
2.4.	Variables aléatoires continues.....	41
2.4.1.	Fonction de répartition.....	41
2.4.2.	Loi ou densité de probabilité d'une variable aléatoire continue.....	42
2.4.3.	Variables aléatoires continues indépendantes.....	43
2.4.4.	Espérance - variance - écart-type d'une variable aléatoire continue.....	43
2.5.	Valeurs caractéristiques.....	44
Chapitre 3 : Lois discrètes et Lois continues usuelles.....		45
3.1.-	Loi uniforme.....	45
3.2.-	Loi de Bernoulli $\mathcal{B}(1, p)$	46
3.3.-	La loi Binomiale $\mathcal{B}(n, p)$ - Loi des tirages avec remise.....	46
3.2.1.-	La distribution d'une proportion.....	51
3.4.-	La loi hypergéométrique \mathcal{HN}, n, p ou Loi du tirage exhaustif.....	51
3.5.-	La loi de Poisson $\mathcal{P}\lambda$	53
3.6.-	Approximation des lois (Lois limites).....	57
3.5.1.-	Approximation d'une loi hypergéométrique par une loi binomiale.....	57
3.5.2.-	Approximation d'une loi binomiale par une loi de Poisson.....	57
3.7.-	Lois continues usuelles.....	62
3.7.3.-	La loi normale (Laplace-Gauss) : $\mathcal{N}\mu, \sigma$	62
3.7.3.-	Lois dérivées de la loi normale.....	76
Chapitre 4 : Inégalités et Notions de convergence de variables aléatoires.....		80
Annexe : Rappels sur les ensembles.....		85
5.-	Références Bibliographiques.....	87

Chapitre 1.- Éléments de Probabilités

Introduction¹

À l'origine, le mot « statistique » désignait les données indispensables à l'État pour gouverner. Ces données concernaient les ressources en hommes (en démographie), les productions agricoles et artisanales, les ressources financières. La statistique descriptive est née pour décrire ces ensembles de données. Très rapidement, on a recherché des modèles abstraits pour représenter ces données, ainsi est née une théorie axiomatique : **le calcul des probabilités.**

La théorie des probabilités est une science mathématique dont le but est d'étudier les lois qui régissent les phénomènes ou les expériences aléatoires.

Une expérience est aléatoire lorsqu'il n'est pas possible de prévoir quel sera le résultat de cette expérience. La théorie et le calcul des probabilités sont utilisés dans de nombreux domaines, en particulier en statistique inductive ou inférentielle.

La statistique inférentielle consiste à induire les caractéristiques inconnues d'une population à partir d'un échantillon issu de cette population. Les caractéristiques de l'échantillon, une fois connues, reflètent avec une certaine marge d'erreur possible celles de la population.

Le calcul des probabilités permet de mesurer ces marges d'erreur à l'aide d'outils statistiques désignés par **intervalle de confiance.**

1.1.- Les vocabulaires de bases des probabilités

1.1.1.- Vocabulaire des évènements

- Une **épreuve**, ou **expérience aléatoire**, est une action dont le résultat est soumis au hasard.
- L'**univers**, ou **référentiel**, noté Ω associé à une épreuve, est l'ensemble des résultats possibles. Chacun de ces résultats possibles est appelé une éventualité.
- Un **évènement** est un sous-ensemble de l'univers. Un évènement est lié à une expérience aléatoire si lorsque cette expérience a été faite, il est possible de dire si cet évènement est réalisé ou non.
- Un évènement ne contenant qu'un seul élément est appelé **évènement élémentaire**. Le terme **issu** ou **éventualité** peut également être employé.

Exemple : On lance un dé à 6 faces numérotées de 1 à 6 et on note le chiffre obtenu.

- Cela constitue une épreuve ou une expérience aléatoire.
- L'univers correspondant est l'ensemble $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Les sous-ensembles de Ω suivants : $A = \{2, 4, 6\}$, $B = \{1, 2, 3, 4\}$, $C = \{5\}$, $D = \{3, 6\}$ sont des exemples d'évènements. Ils peuvent également être définis par une « phrase ». Par exemple : $A =$ « obtenir un chiffre pair » ; $B =$ « obtenir un résultat inférieur ou égal à 4 ».
- Les évènements élémentaires sont ici : $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}$ et $\{6\}$.

NB : Soit A un évènement donné. Si, lors d'une épreuve, on obtient le résultat ω , alors :

- Si ω est un élément de A , on dit que A est réalisé ;

¹ Maurice LETHIELLEUX, *Probabilités – Estimation Statistique*, 3^e édition ; Dunod.

➔ Si ω n'appartient pas à A , on dit que A n'est pas réalisé.

Exemple : Supposons par exemple que l'on jette le dé et que l'on obtient le chiffre 3. Avec les notations introduites ci-dessus, on peut dire alors que : les évènements A et C ne sont pas réalisés, tandis que B et D le sont.

1.1.2. - Évènement impossible, évènement certain

Parmi les sous-ensembles de Ω figurent l'ensemble vide, noté \emptyset et Ω lui-même. Ces deux sous-ensembles sont donc des évènements.

Or, quel que soit le résultat ω obtenu d'une épreuve, il va de soi que :

- ω n'appartient pas à \emptyset (puisque l'ensemble vide ne contient aucun élément), donc l'évènement \emptyset n'est jamais réalisé. On l'appelle **l'évènement impossible** ;
- ω appartient à Ω (puisque l'univers contient toutes les éventualités), donc l'évènement Ω est toujours réalisé. On l'appelle **l'évènement certain**.

\emptyset = évènement impossible

Ω = évènement certain

1.1.3. - Évènement contraire

Soit A un évènement. L'évènement dit contraire, noté \bar{A} , est l'évènement qui est réalisé si et seulement si A ne l'est pas.

\bar{A} est donc le complémentaire dans Ω de A .

Exemple : Reprenons le lancer d'un dé, et considérons à nouveau les 4 évènements :

$A = \{2, 4, 6\}$, $B = \{1, 2, 3, 4\}$, $C = \{5\}$, $D = \{3, 6\}$.

Leurs contraires sont :

$\bar{A} = \{1, 3, 5\}$, $\bar{B} = \{5, 6\}$, $\bar{C} = \{1, 2, 3, 4, 6\}$, $\bar{D} = \{1, 2, 4, 5\}$.

Exercice : On dispose de 10 livres dont 3 sont en anglais et les autres en français. On choisit au hasard 4 de ces livres et on note :

A l'évènement : « on obtient au moins 2 livres en anglais », et

B l'évènement : « on obtient exactement 2 livres en anglais ».

Déterminer \bar{A} et \bar{B} .

Réponse :

Déterminer \bar{A}

A signifie que l'on obtient 2, 3 ou 4 livres en anglais.

L'évènement contraire de A est donc :

$\bar{A} =$ « On obtient 0 ou 1 livre en anglais » = « on obtient au plus un livre en anglais »

Déterminer \bar{B}

L'évènement contraire de B est :

$\bar{B} =$ « Le nombre de livres en anglais obtenu n'est pas 2 » = « on obtient 0 ou 1 ou 3 ou 4 livres en anglais ».

1.1.4.- Opérations sur les évènements

A.- Inclusion

A et B étant deux évènements, si A est un sous-ensemble de B (ou une partie de B), on note $A \subset B$, et on dit que l'évènement A est inclus dans l'évènement B (ou que A implique B). Dans ce cas, lorsque A est réalisé, B l'est également.

Exemple : Considérons par exemple l'évènement $F = \{2, 4\}$.

On a $\underbrace{\{2, 4\}}_F \subset \underbrace{\{2, 4, 6\}}_A$, donc chaque fois que l'évènement F est réalisé, l'évènement A l'est également.

B.- Intersection d'évènements

A et B étant deux évènements, c'est-à-dire deux sous-ensembles de Ω , l'intersection de A et B, notée $A \cap B$, est l'ensemble des éléments de Ω appartenant à la fois à A et à B.

$A \cap B$ est l'évènement qui n'est réalisé que si A ET B le sont.

Lorsque $A \cap B$ est vide, cela signifie que A et B ne peuvent pas être réalisés simultanément. On dit alors que A et B sont des **évènements incompatibles**.

A et B sont incompatibles si et seulement si $A \cap B = \emptyset$.

C.- Réunion d'évènements

A et B étant deux évènements, la réunion de A et B, notée $A \cup B$, est l'ensemble des éléments de Ω appartenant soit à A, soit à B, sans exclure les éléments appartenant à la fois à A et à B s'il y en a.

$A \cup B$ est l'évènement qui est réalisé si A OU B l'est (le « ou » étant non exclusif).

Exemple : Reprenons le lancer d'un dé, et considérons à nouveau les 4 évènements :

$A = \{2, 4, 6\}$, $B = \{1, 2, 3, 4\}$, $C = \{5\}$, $D = \{3, 6\}$.

Les évènements A et C par exemple sont incompatibles puisque $A \cap C = \emptyset$. Par ailleurs, on a : $A \cap B = \{2, 4\}$ et $A \cup B = \{1, 2, 3, 4, 6\}$.

D.- Complémentaire d'une réunion et d'une intersection

Il est bon de connaître les propriétés suivantes (Loi de De Morgan) :

$$\overline{A \cap B} = \overline{A} \cup \overline{B} \quad \text{et} \quad \overline{A \cup B} = \overline{A} \cap \overline{B}$$

On peut généraliser ces propriétés sous la forme :

$$\overline{A_1 \cap A_2 \cap \dots \cap A_n} = \overline{A_1} \cup \overline{A_2} \cup \dots \cup \overline{A_n} \quad \text{et} \quad \overline{A_1 \cup A_2 \cup \dots \cup A_n} = \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_n}$$

1.1.5.- Parallélisme entre la terminologie ensembliste et la terminologie probabiliste

Pour représenter les évènements, il est d'usage d'utiliser la théorie des ensembles (*voir annexe A*). Un évènement est donc associé à un sous-ensemble de l'univers Ω , l'ensemble des résultats de l'expérience aléatoire (\mathcal{E}) pour lesquels l'évènement est réalisé.

Terminologie ensembliste	Notation	Terminologie probabiliste
Ensemble plein	Ω	Univers des possibles
Élément de Ω	$\omega \in \Omega$	Le résultat ω est une des réalisations possibles de Ω
Sous-ensemble de Ω	$A \subset \Omega$	Évènement A
Ensemble entier	Ω	Évènement certain
Partie vide	\emptyset	Évènement impossible
Un singleton	$\{\omega\} \in \Omega$	Évènement élémentaire
Complémentaire de A dans Ω	\bar{A}	Évènement contraire de A
Réunion de A et B	$A \cup B$	Évènement A ou B
Intersection de A et B	$A \cap B$	Évènement A et B
A et B sont disjoints	$A \cap B = \emptyset$	A et B sont incompatibles
A inclus dans B	$A \subset B$	L'évènement A implique l'évènement B
Partition de Ω	$(A_i)_{i \in I}$ partition de Ω	$(A_i)_{i \in I}$ système complet d'évènements

Remarque : Toutes les opérations précédemment définies s'étendent à plus de deux événements. La classe des événements associés à une expérience aléatoire est donc une *tribu* \mathcal{A} de parties de Ω (tribu ou σ -algèbre).

En résumé :

Un espace probabilisable est un couple (Ω, \mathcal{A}) formé d'un ensemble Ω et d'une tribu \mathcal{A} de parties de Ω (qui sont les événements).

1.1.6.- Espaces probabilisables

Définition. — On appelle *espace probabilisable* tout couple (Ω, \mathcal{A}) formé d'un ensemble non vide Ω et d'une tribu \mathcal{A} de parties de Ω . Les éléments de \mathcal{A} sont appelés *événements*.

Exemples

- 1) Le couple $(\Omega, \mathcal{P}(\Omega))$ est un espace probabilisable. C'est l'espace probabilisable qu'on associe toujours à Ω , lorsque Ω est au plus dénombrable.
- 2) Le couple $(\mathbb{R}^n, \mathcal{B}^n)$ est un espace probabilisable.

1.2.- Modéliser une expérience aléatoire

En théorie des probabilités, le terme **modéliser** désigne l'opération qui consiste à associer à l'expérience aléatoire (\mathcal{E}) trois objets mathématiques, notés et appelés généralement **Ω , l'univers, \mathcal{A} , l'ensemble des événements, et P , la probabilité.**

1.2.1.- Définition d'une probabilité - Espace probabilisé

Afin de compléter le modèle probabiliste qui permettra de décrire le comportement du résultat de l'expérience aléatoire considérée, il faut maintenant trouver une façon de déterminer la **probabilité** de chacun des évènements que nous aurons à considérer.

La probabilité d'un des résultats d'une expérience aléatoire sera une mesure de l'importance de ce résultat dans l'ensemble de tous les résultats possibles de l'expérience.

Il est souvent commode de se représenter la probabilité d'un évènement comme une sorte de « poids ». Un poids total de 1 est réparti sur les éléments de l'espace fondamental (univers).

Plus le poids d'un résultat est grand, plus grande est la probabilité que l'expérience aléatoire donne ce résultat. Une fois déterminée la probabilité de chacun des *résultats*, c'est-à-dire de chaque élément de Ω , il sera aisé de calculer la probabilité d'un *évènement*, c'est-à-dire d'une partie de Ω . Notons que les probabilités négatives ou supérieures à 1 étant manifestement absurdes, on doit toujours avoir pour tout élément de Ω une probabilité comprise entre 0 et 1 (inclus).

On part maintenant d'un espace probabilisable (Ω, \mathcal{A}) et on cherche à munir les évènements, c'est-à-dire les éléments de la tribu \mathcal{A} , d'une pondération utilisant pleinement les propriétés des tribus. On obtient ainsi un triplet (Ω, \mathcal{A}, P) , appelé *espace probabilisé*. L'idée de modéliser une expérience aléatoire au moyen d'un tel triplet (non nécessairement unique !) a marqué un tournant décisif dans le développement du calcul des probabilités. Elle est due essentiellement à Kolmogorov.

1.2.2.- Calcul d'une probabilité

Ce qui nous intéresse, dans une expérience aléatoire, c'est de connaître les chances de réalisations d'un évènement : c'est ce que l'on appelle la **probabilité** de l'évènement. Nous verrons ici les différentes façons d'aborder le calcul d'une probabilité.

1.2.3.- Différentes conceptions des probabilités

Il existe plusieurs façons de concevoir le concept de probabilités. On peut parler de probabilité subjective, de probabilité objective ou a priori et de probabilité empirique.

A.- Probabilité subjective

On dira de la mesure d'une probabilité qu'elle est subjective lorsqu'elle ne se fonde que sur une intuition ou un pressentiment et qu'elle n'est pas appuyée par des données historiques ou autres.

Exemple : Un élève qui affirme avoir 70% de chances de réussir un examen émet une opinion subjective de nature probabiliste sur ses chances de réussite.

On comprendra aisément qu'il est impossible de construire une théorie scientifique des probabilités en se basant sur une telle conception.

B.- Probabilité objective (ou a priori)

Définition classique de la probabilité

Cette approche, proposée par Pierre Simon de Laplace au début du XIX^e siècle, repose sur le principe d'indifférence. En vertu de ce principe, on doit supposer que deux évènements sont également probables s'il n'y a pas de raison de penser qu'un des deux évènements peut être obtenu plus fréquemment que l'autre. On dit alors que les évènements sont équiprobables.

La définition classique de la probabilité énoncée par Laplace nous apprend que, si tous les cas sont équiprobables, alors la probabilité d'un évènement A , notée $P(A)$, sera :

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

Exemples :

- 1) La probabilité d'obtenir « pile » en lançant une pièce de monnaie serait de $\frac{1}{2}$, étant donné qu'il y a deux résultats possibles également vraisemblables (ou équiprobables) et qu'un seul de ces résultats consiste à obtenir « pile ». Il y a deux cas possibles et un seul cas favorable ; la probabilité vaut donc $\frac{1}{2}$.
- 2) La probabilité d'obtenir un six avec un dé non pipé (un dé honnête où toutes les faces ont les mêmes chances de se produire) est de $\frac{1}{6}$.
- 3) Lorsqu'on lance deux dés, la probabilité que la somme des points donne trois vaut $\frac{2}{36} = \frac{1}{18}$ puisqu'il y a 36 résultats possibles et 2 résultats favorables.

Remarque :

- 1) Selon la définition classique de Laplace, on parle de probabilité *a priori* parce que la probabilité d'un évènement peut être évaluée sans qu'on en fasse l'expérience ; on n'a pas besoin de lancer un seul dé ni aucune pièce de monnaie. La probabilité obtenue de cette façon est conçue à partir d'un monde idéal dans lequel les évènements sont effectivement équiprobables. Par exemple, lors du lancement d'un dé, il y a six résultats possibles. A moins que quelqu'un ne vous ait refilé un dé truqué, tous les résultats sont également probables. Par conséquent, la probabilité d'obtenir n'importe quelle face du dé est de $\frac{1}{6}$.

La probabilité a priori est une probabilité déterminée sans effectuer d'expériences.

- 2) Cette définition, bien que très valable pour le jeu de pile ou face, le jeu de dés, les jeux de cartes, ... n'est pas toujours adéquate pour les situations que l'analyste peut rencontrer. Considérons par exemple, un procédé de fabrication de lampes fluorescentes où l'on voudrait trouver la probabilité d'avoir une lampe défectueuse (évènement E). Il est pratiquement impossible de dénombrer tous les cas favorables (la lampe est défectueuse), car cela suppose qu'on vérifie toutes les lampes fabriquées. Alors, même si on trouve 150 lampes défectueuses parmi un lot de 20 000, on ne peut, d'après la définition précédente, dire que la probabilité que le processus de fabrication donne une lampe défectueuse est $P(E) = \frac{150}{20\,000}$ car cette mesure ne porte que sur une partie de la production et non sur tous les cas possibles. Nous avons alors besoin d'une autre interprétation de la probabilité.

C.- Probabilité empirique - Définition empirique de la probabilité

Dans ce cas, la probabilité est définie par une fréquence relative déduite de l'observation.

Considérons une série de N expériences où un certain évènement E s'est réalisé n fois. La fréquence relative de la réalisation de l'évènement E est $f = \frac{n}{N}$. Au fur et à mesure que N augmente le rapport $\frac{n}{N}$ tend à se stabiliser au voisinage d'une certaine valeur limite qu'on note $P(E)$ telle que :

$$\lim_{N \rightarrow \infty} \left(\frac{n}{N} \right) = P(E) \text{ où } P(E) \text{ est la probabilité de l'évènement } E.$$

Cette interprétation est la probabilité la plus réaliste pour l'analyste. En général, la probabilité $P(E)$ reste inconnue et sa valeur est estimée par la fréquence expérimentale f . Ainsi, dans l'exemple des lampes fluorescentes, la probabilité qu'une lampe quelconque soit défectueuse est estimée par la fréquence expérimentale $f = \frac{150}{20000} = 0.0075 = P(E)$ si nous considérons que 20 000 est un nombre suffisamment grand.

La probabilité empirique d'un évènement est déterminée à l'aide de l'observation et de l'expérimentation.

Pour plusieurs types de décisions, ce type de probabilité s'avère très important. La quantité d'articles que le détaillant doit maintenir en stock, le nombre d'infirmières qu'un hôpital doit recruter, les primes d'assurance qu'une compagnie doit imposer sont tous des facteurs déterminés à l'aide de probabilités empiriques.

Le calcul des probabilités s'est avéré nécessaire dans un grand nombre de domaines (par exemple en économie, en physique, en médecine, en gestion, ...) où la répétition à l'identique d'une même épreuve est souvent impossible. Vers 1930, Kolmogorov a construit une théorie mathématique des probabilités, s'appuyant sur l'approche statistique de ces précurseurs. Basée sur trois axiomes fondamentaux, cette théorie confirme et généralise les propriétés établies à partir d'expériences.

1.2.4.- Définition axiomatique d'une probabilité sur un univers Ω

Une probabilité est une application P permettant d'associer un réel $P(A)$ à tout évènement et vérifiant les axiomes suivants :

- 1) Pour tout évènement A : $0 \leq P(A) \leq 1$;
- 2) $P(\Omega) = 1$;
- 3) Pour tous évènements A et B incompatibles : $P(A \cup B) = P(A) + P(B)$. On l'appelle axiome des probabilités totales ou d'additivité.

On dit que $P(A)$ est la probabilité de réaliser l'évènement A .

A.- Conséquences de la définition

P étant une probabilité définie sur Ω , on a les propriétés suivantes :

- a) La probabilité de l'évènement impossible est nulle : $P(\emptyset) = 0$;
- b) \bar{A} étant l'évènement contraire de A , la probabilité de réaliser \bar{A} vaut : $P(\bar{A}) = 1 - P(A)$;
- c) Si $A \subset B$, alors : $P(A) \leq P(B)$;

d) Pour tous évènements A et B : $P(A \cup B) + P(A \cap B) = P(A) + P(B)$

Exemple : À la sortie d'une chaîne de fabrication, des boulons sont susceptibles de présenter deux défauts. Un très grand nombre d'observations a permis d'établir que :

- La proportion de boulons fabriqués ayant le défaut a est de 5% ;
- La proportion de boulons fabriqués ayant le défaut b est de 3% ;
- La proportion de boulons fabriqués ayant les deux défauts est de 1%.

On note : A l'évènement « un boulon présente le défaut a » et B l'évènement « un boulon présente le défaut b ».

La probabilité que A se réalise vaut donc 0.05, la probabilité que B se réalise vaut 0.03, et la probabilité que A et B se réalisent simultanément vaut 0.01.

Attention : « présenter le défaut a » ne signifie pas « présenter le défaut a seulement ». Déterminer la probabilité qu'un boulon présente :

- a) Le défaut a ou le défaut b ;
- b) Le défaut a seulement ;
- c) Aucun défaut.

B.- Généralisation de l'axiome des probabilités totales

Soient $A_1, A_2, \dots, A_i, \dots, A_n$ des évènements de Ω tels que $A_i \cap A_j = \emptyset$ pour tout $i \neq j$ alors :

$$P(A_1 \cup A_2 \cup \dots \cup A_i \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_i) + \dots + P(A_n)$$

ou $P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$, propriété de σ - additivité.

1.2.5.- Probabilités sur un univers fini

On suppose dans ce paragraphe que l'univers Ω contient n éléments : $\Omega = \{a_1, a_2, a_3, \dots, a_n\}$.

Toute application P définie par la donnée des n nombres positifs ou nuls $P(a_1), P(a_2), \dots, P(a_n)$ tels que $P(a_1) + P(a_2) + \dots + P(a_n) = 1$ détermine une probabilité sur Ω .

On obtient la probabilité d'un évènement A en ajoutant les probabilités des éventualités dont la réunion est A .

Exemple 1 : Soit $\Omega = \{1, 2, 3, 4, 5, 6\}$ et P définie par : $P(1) = 0.1, P(2) = 0.4, P(3) = 0.05, P(4) = 0.15, P(5) = 0.2$ et $P(6) = 0.1$. Ces 6 nombres positifs ou nuls et leur somme vaut bien 1. D'après le théorème ci-dessus, P définit donc une probabilité sur Ω .

La probabilité d'obtenir un nombre pair, par exemple, est alors :

$$P(\{2, 4, 6\}) = 0.4 + 0.15 + 0.1 = 0.65$$

Exemple 2 : Sur le même univers $\Omega = \{1, 2, 3, 4, 5, 6\}$, considérons maintenant l'application P définie par : $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$.

$\frac{1}{6}$ est positif ou nul et $P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = 1$. P définit donc également une probabilité sur Ω .

On dit ici que les évènements élémentaires sont équiprobables car ils ont tous la même probabilité de se réaliser.

La probabilité de l'évènement A « obtenir un nombre inférieur ou égal à 4 », par exemple, est alors :

$$P(A) = P(\{1,2,3,4\}) = P(1) + P(2) + P(3) + P(4) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{4}{6} = \frac{\text{card } A}{\text{card } \Omega}$$

$$= \frac{\text{nombre de résultats "favorables"}}{\text{nombre total de résultats possibles}}$$

Cette formule, établie ici sur un cas particulier, est généralisée dans le paragraphe qui suit.

1.2.6.- Équiprobabilité sur un univers fini

On a équiprobabilité sur l'univers $\Omega = \{a_1, a_2, a_3, \dots, a_n\}$ lorsque tous les évènements élémentaires ont la même probabilité de se réaliser.

On a alors : $P(a_1) = P(a_2) = \dots = P(a_n) = \frac{1}{n}$ (puisque tous ces nombres sont égaux et que leur somme vaut 1) et pour tout évènement A :

$$P(A) = \frac{\text{card } A}{\text{card } \Omega} = \frac{\text{nombre de résultats "favorables"}}{\text{nombre total de résultats possibles}}$$

Exemple 1 : L'entreprise DUBEAU emploie 36 personnes dont 12 femmes. Un salarié de cette entreprise est tiré au sort pour répondre à un sondage. Quelle est la probabilité que ce soit un homme ?

Exemple 2 : Un libraire reçoit un carton de 50 exemplaires d'un livre dont 3 sont dédiés par l'auteur. On tire au hasard 10 livres du carton. Déterminer la probabilité des évènements :

A : « on obtient les 3 livres dédiés »

B : « on obtient un seul livre dédié »

C : « on obtient au moins un livre dédié »

1.2.7.- Probabilité sur un univers infini dénombrable

Ω est dénombrable si on peut numéroter ses éléments. Notons $\Omega = \{a_1, a_2, a_3, \dots, a_n, \dots\}$. La donnée de n nombres positifs ou nuls $P(a_1), P(a_2), \dots, P(a_n), \dots$ tels que $\sum_{i=1}^{+\infty} P(a_i) = 1$ définit une probabilité sur Ω .

Remarque :

a) On obtient la probabilité d'un évènement A en additionnant les probabilités des éventualités dont la réunion est A ;

b) La somme $\sum_{i=1}^{+\infty} P(a_i)$ est par définition la limite quand n tend vers $+\infty$ de $\sum_{i=1}^n P(a_i)$.

Exemple : Notons \mathbb{N}^* l'ensemble des entiers naturels non nuls, et considérons l'application P définie sur \mathbb{N}^* par : $P(k) = \left(\frac{1}{2}\right)^k$. Montrer que P définit une probabilité sur \mathbb{N}^* , puis calculer la probabilité de l'évènement : $A = \{2, 4, 6, 8\}$.

1.2.8.- Espace probabilisé

Définition. — On appelle *espace probabilisé* associé à une expérience aléatoire \mathcal{E} tout triplet (Ω, \mathcal{A}, P) , où (Ω, \mathcal{A}) est un espace probabilisable et où P est une mesure de probabilité sur \mathcal{A} (ou, de façon plus précise, sur (Ω, \mathcal{A})).

Exercice : Soit n un entier naturel non nul. Nous effectuons n lancers d'un dé cubique bien équilibré dont les faces sont numérotées de 1 à 6.

1) Définissez l'espace probabilisé associé aux n lancers ainsi que la probabilité associée.

2) Calculez la probabilité des deux événements A_n et B_n définis par :

- a) A_n : « Nous obtenons 1 pour la première fois au $n^{\text{ième}}$ lancer ».
- b) B_n : « Nous n'obtenons aucun 1 lors des n lancers ».

1.3.- Système complet d'évènements

Soient $A_1, A_2, \dots, A_i, \dots, A_n$ des évènements où les A_i sont inclus dans l'ensemble fondamental Ω . Alors, l'ensemble $\{A_1, A_2, \dots, A_n\}$ est un système complet d'évènements ou une partition de Ω si et seulement si :

- Aucun des évènements n'est impossible, c'est-à-dire : $\forall i, A_i \neq \emptyset$;
- Les évènements sont incompatibles deux à deux, c'est-à-dire : $\forall i \neq j, A_i \cap A_j = \emptyset$;
- L'union des évènements est l'ensemble fondamental : $\bigcup_i A_i = \Omega$.

Exemple : Pour le triple lancer d'une pièce de monnaie, on peut par exemple définir le système complet d'évènements suivant :

$$A_1 = \{PPF, PFP, PFF\}; A_2 = \{FPP, FPF, FFP\} \text{ et } A_3 = \{FFF, PPP\}$$

1.4.- Rappels sur l'Analyse Combinatoire

Le dénombrement est un outil de base pour le calcul de probabilité. Les méthodes de dénombrement n'appartiennent pas à la théorie des probabilités, mais à l'analyse combinatoire. Historiquement, les notions de permutation et de combinaison et leur développement sont l'œuvre de B. Pascal et de P. de Fermat dans le courant du XVII^{ème} siècle. Cependant certains problèmes Combinatoires ont été posés en Chine vers l'an 1100 avant J.C. Alors que le perse Rabbi Ben Erza (vers 1140) paraît avoir déjà établi beaucoup de formules attribuées à des chercheurs européens de la Renaissance. Les coefficients du binôme ont été étudiés par un arithméticien indien Bhaska vers le XII^{ème} siècle, et le fameux "triangle de Pascal" a été établi par le philosophe perse Nassir-Ad-Din en l'an 1265. Toutes les techniques combinatoires sont basées sur le principe fondamental (*principe de la multiplication*).

1.4.1. Indications générales

1) Questions à se poser

Pour dénombrer des situations, il est commode de se poser les questions :

- Quel est le nombre n d'objets de référence ?
- Quel est le nombre p d'objets concernés par une situation ?

- Les p objets sont-ils considérés sans ordre (en vrac ; tirage simultané) ou avec ordre (c'est-à-dire que la situation est différente si les mêmes p objets sont classés de façon différente) ?
- Les répétitions sont-elles impossibles (les p objets sont tous distincts ; tirage sans remise) ou possibles (tirage avec remise) ?

2) Opérations à effectuer

Quand une situation comporte plusieurs choix :

- On effectue un produit quand on doit faire un choix, *puis* un autre, ... (**principe de la multiplication**)
- On effectue une somme quand on considère un cas *ou bien* un autre, ... (**principe d'addition**)

Exemples sur le principe de la multiplication

Exemple 1 :

Les localités A et B sont reliées par $n_1 = 3$ routes différentes et les localités B et C par $n_2 = 2$ routes différentes ; alors il y a $N = 3 \times 2$ manières différentes de se rendre par la route de la localité A à la localité C.

Exemple 2 :

Jacques arrive au restaurant. Il désire prendre un repas complet (c'est-à-dire un potage, un plat de résistance, un légume, un dessert et une boisson). On lui présente un menu à la carte offrant un choix de 6 potages, 4 plats de résistance, 3 légumes, 5 desserts et 8 boissons. Combien de repas complets différents Jacques peut-il composer ?

Ici, la composition d'un repas complet suppose un choix de potage **avec** un choix de plat de résistance **avec** un choix de légumes **avec** un choix de dessert **avec**, enfin, un choix de boisson.

Pour calculer le nombre de repas qu'il est ainsi possible de composer, on utilise donc la règle de **multiplication**, ce qui donne :

$$6 \times 4 \times 3 \times 5 \times 8 = 2\,880 \text{ repas possibles}$$

Exemple 3 :

Supposons qu'une plaque d'immatriculation contient deux lettres distinctes suivies de trois chiffres dont le premier est différent de zéro. Combien de plaques différentes peut-on imprimer ?

Il y a 26 façons différentes d'imprimer la première lettre, 25 façons différentes d'imprimer la seconde lettre (puisque la première lettre imprimée ne peut plus être choisie pour la deuxième), 9 façons différentes d'imprimer le premier chiffre et 10 façons différentes d'imprimer les deux autres chiffres. On en déduit que l'on peut imprimer :

$$26 \times 25 \times 9 \times 10 \times 10 = 585\,000 \text{ plaques différentes}$$

Exemple 4 : Le comité de planification d'un collège est constitué de 3 étudiants de première année, 4 de deuxième, 5 de troisième 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 ?

Réponse : Nous pouvons considérer le choix d'un sous-comité comme le résultat combiné de 4 expériences distinctes, chacune consistant à choisir un unique représentant dans l'une des classes. Par conséquent, en application du principe fondamental généralisé, il y a : $3 \times 4 \times 5 \times 2 = 120$ sous-comités possibles.

Exemples sur le principe d'addition

Exemple 1 :

Jeanne vient au restaurant pour prendre une collation (c'est-à-dire ou bien un potage, ou bien un sandwich, ou bien un dessert). On lui présente un menu offrant un choix de 5 potages, 7 sandwiches et 4 desserts. Combien de collations différentes peut-elle choisir ?

Dans ce cas-ci, comme Jeanne doit effectuer son choix de la façon suivante : potage **ou bien** sandwich **ou bien** dessert ; on doit faire appel à la règle d'**addition** pour calculer qu'elle a $5 + 7 + 4 = 16$ possibilités de collations différentes.

Exemple 2 :

Dans un jeu de 52 cartes, combien en compte-t-on qui soient une carte rouge ou un as ?

Dans ce cas,

$$\begin{aligned} \#(\text{rouge ou as}) &= \#(\text{rouge}) + \#(\text{as}) - \#(\text{as rouge}) \\ &= 26 + 4 - 2 = 28 \end{aligned}$$

1.4.2. Situations sans répétition

1) Avec ordre (Arrangement)

Dans un ensemble à n éléments, il s'agit de choisir p éléments tous distincts (ce qui nécessite $p \leq n$) et avec ordre. Une telle situation est un **arrangement** de n éléments pris p à p . Leur nombre (qu'on peut noter A_n^p) est :

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

Par exemple ; $(a, c, d), (a, d, c), (b, c, e)$ sont des arrangements distincts, d'ordre 3 de (a, b, c, d, e) .

Remarque : Chacun de ces arrangements est constitué de trois lettres *différentes* prises parmi cinq lettres *différentes*.

Deux de ces choix diffèrent :

- Soit par leur *composition* (le premier et le troisième choix par exemple, ou le deuxième et le troisième) ;
- Soit, quand ils ont même composition (et c'est le cas du premier et du deuxième choix), par *l'ordre* dans lequel ont été disposés les éléments qui composent ce choix.

Exemple :

Les arrangements de deux lettres prises parmi 4 lettres $\{a, b, c, d\}$ sont au nombre de :

$$A_4^2 = \frac{4!}{2!} = 12$$

Ce sont : (a, b) ; (a, c) ; (a, d) ; (b, a) ; (b, c) ; (b, d) ; (c, a) ; (c, b) ; (c, d) ; (d, a) ; (d, b) ; (d, c) .

Remarques :

1. Lorsque $p = n$, il s'agit d'ordonner n objets entre eux, c'est-à-dire d'effectuer une permutation de ces n objets sans répétition.
2. Une permutation de n éléments est une disposition ordonnée de ces n éléments.
3. Les permutations de n éléments sont au nombre de $A_n^n \equiv P_n = n!$.
4. Avec la convention $0! = 1$, on a donc : $A_n^0 = 1$ et $A_n^n = 1$.

Notion de factorielle

Définition et notation : Le produit des entiers positifs de 1 à n inclus intervient très souvent en mathématiques. On le note par le symbole spécial $n!$ (lire « *factorielle n* ») et définit par :

$$n! = n(n-1) \cdots \cdots \cdots$$

..

$m=1$

De même »

$$(n+1)! = (n+1)n!$$

Si $m < n$, alors

$$\frac{n!}{m!} = \prod_{p=1}^{n-m} (m+p)$$

Par convention, $0! = 1$.

D'où, pour tout $n \in \mathbb{N}$;

$$n! = \begin{cases} 1 \times 2 \times 3 \times \cdots & \text{si } n \neq 0 \\ 1 & \text{si } n = 0 \end{cases}$$

Exemple : Une urne contient 9 boules numérotées de 1 à 9. On y choisit 5 boules l'une après l'autre et on inscrit les numéros ainsi obtenus dans l'ordre.

Combien de dispositions obtient-on si :

1. On ne remet pas la boule après chaque tirage ?
2. On remet la boule après chaque tirage ?

2) Sans ordre (Combinaison)

Dans un ensemble à n éléments, il s'agit de choisir une partie à p éléments (ce qui nécessite $p \leq n$).

Leur nombre est le nombre de **combinaison** de n éléments pris p à p (ancienne notation C_n^p). La notation actuelle est :

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

Exemple 1 : Les différentes combinaisons de l'ensemble $\{a, b, c, d\}$ pris 3 à la fois sont représentées par les sous-ensembles $\{a, b, c\}$, $\{a, b, d\}$, $\{a, c, d\}$, $\{b, c, d\}$.

Par ailleurs, les sous-ensembles $\{a, b, c\}$, $\{a, c, b\}$, $\{c, a, b\}$, $\{b, a, c\}$, $\{b, c, a\}$, $\{c, b, a\}$ représentent le même sous-ensemble $\{a, b, c\}$, c'est-à-dire la même combinaison (l'ordre n'intervient pas).

Exemple 2 : De combien de façons peut-on former un comité de trois hommes et deux femmes choisis parmi 7 hommes et 5 femmes ?

Réponse : On peut choisir 3 hommes parmi 7 de C_7^3 différentes façons et 2 femmes parmi 5 de C_5^2 différentes façons, donc d'après le principe de multiplication, le nombre total de choix du comité est : $C_7^3 \times C_5^2 = 350$.

Exemple 3 : Le cinéma Étoile emploie des équipes de 3 employés pour exploiter le comptoir alimentaire tous les soirs. Au total, 7 employés sont disponibles chaque soir pour travailler. Combien d'équipes différentes peuvent être formées pour exploiter le comptoir alimentaire ?

Réponse :

A l'aide de la formule de la combinaison, on obtient : $C_7^3 = 35$.

Avec ces 7 employés, on peut former 35 équipes différentes de 3 employés.

Propriétés :

$$\bullet \binom{n}{p} = \binom{n}{n-p} \qquad \bullet \binom{n+1}{p+1} = \binom{n}{p} + \binom{n}{p+1}$$

Cas particuliers

$$\bullet \binom{n}{0} = \binom{n}{n} = 1 \qquad \bullet \binom{n}{1} = \binom{n}{n-1} = n$$

Exemple : 7 salariés sont convoqués un matin à 8 heures pour un bilan de santé. Lorsqu'ils se présentent, tous ensemble, 2 d'entre eux doivent aller voir le cardiologue et les 5 autres doivent se rendre auprès des infirmières qui font les prises de sang.

- **De combien de façons peut-on** choisir les 2 personnes qui commenceront la visite par un électrocardiogramme ?
- **De combien de façons peut-on** choisir les 5 personnes qui commenceront la visite par une prise de sang ?

Réponse :

- Il y a C_7^2 façons de choisir 2 personnes parmi 7, or $C_7^2 = \frac{7 \times 6}{2!} = 21$, il y a donc 21 possibilités pour le choix des 2 salariés qui doivent aller voir le cardiologue.
- Il y a C_7^5 façons de choisir 5 personnes parmi 7, or $C_7^5 = \frac{7 \times 6 \times 5 \times 4 \times 3}{5!} = 21$, il y a donc également 21 possibilités pour le choix des 5 salariés qui commenceront par une prise de sang.
- Il était évident que l'on obtiendrait le même résultat : lorsque l'on choisit 2 personnes parmi 7, on en laisse 5.

Remarque : Il peut être intéressant d'utiliser cette formule pour simplifier des calculs numériques.

Exemple : Calculer C_{150}^{147}

On a : $C_{150}^{147} = C_{150}^3 = \frac{150 \times 149 \times 148}{3!} = 551\,300.$

Exemple : Reprenons l'exemple des 7 salariés convoqués pour le bilan de santé. L'un d'eux s'appelle Jean.

1) De combien de façons peut-on constituer le groupe de 5 personnes qui commencent par la prise de sang si l'on impose :

- a. Que Jean fasse partie de ce groupe ?
- b. Que Jean n'en fasse pas partie ?

2) Comment s'interprète la somme des deux nombres obtenus ?

Réponse :

1.a. Si Jean fait partie du groupe « prise de sang », il ne reste à choisir que 4 personnes parmi 6, ce qui donne C_6^4 possibilités soit 15 groupes possibles.

1.b. Si Jean ne fait pas partie de ce groupe, il faut choisir 5 salariés parmi 6, ce qui donne C_6^5 possibilités soit 6 groupes possibles.

2. On a : $15 + 6 = 21$. On retrouve le nombre total de groupes de 5 personnes choisies parmi 7.

On a bien établi que : $C_7^5 = C_6^5 + C_6^4$

Remarque : $C_{n+1}^{p+1} = C_n^{p+1} + C_n^p$

1.4.3. Situations avec répétition

1) Avec ordre (arrangement avec répétition)

Dans un ensemble à n éléments, il s'agit de choisir p éléments rangés (avec la possibilité de choisir plusieurs fois le même). Il y a n^p possibilités.

2) Sans ordre (combinaisons avec répétition)

Dans un ensemble à n éléments, il s'agit de choisir p éléments sans ordre (avec la possibilité de choisir plusieurs fois le même). $\binom{n+p-1}{p}$ est le nombre de combinaisons avec répétition. Il est

aussi noté :

$$\Gamma_n^p$$

3) Permutations avec répétition

Soit un ensemble à n éléments comportant :

- n_1 éléments d'un premier type, indiscernables entre eux ;
- n_2 éléments d'un deuxième type, indiscernables entre eux ;
- \vdots
- n_q éléments d'un $q^{\text{ème}}$ type, indiscernables entre eux.

Une permutation avec répétition de ces n éléments est une disposition ordonnée de ces éléments. Il y en a :

$$\frac{n!}{n_1! n_2! \cdots n_q!}$$

1.4.4. Tableau récapitulatif des formules de dénombrement

	Sans répétition	Avec répétition
Avec ordre	A_n^p	n^p
Sans ordre	$\binom{n}{p}$	$\binom{n+p-1}{p}$

1.4.5. Triangle de Pascal - Binôme de Newton

On appelle triangle de Pascal le tableau de nombres suivants :

C_0^0									
C_1^0	C_1^1								
C_2^0	C_2^1	C_2^2							
C_3^0	C_3^1	C_3^2	C_3^3						
C_4^0	C_4^1	C_4^2	C_4^3	C_4^4					
...				
...			
C_{n-1}^0	C_{n-1}^{p-1}	C_{n-1}^p	C_{n-1}^{n-1}		
C_n^0	C_n^p	C_n^n	
...

On peut également écrire ce tableau de nombres en remplaçant chaque C_n^p par sa valeur numérique.

On a $C_n^0 = 1$ et $C_n^n = 1$ pour tout entier n , donc la diagonale et la première colonne ne contiennent que des « 1 ». Pour obtenir les autres valeurs du triangle, il suffit d'utiliser la propriété :

$C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$. On obtient par exemple le triangle suivant :

1									
1	1								
1	2	1							
1	3	3	1						
1	4	6	4	1					
1	5	10	10	5	1				
1	6	15	20	15	6	1			
1	7	21	35	35	21	7	1		
1	8	28	56	70	56	28	8	1	

Ce triangle peut aider à développer aisément $(a+b)^n$, en admettant la formule ci-dessous :

Formule du binôme de Newton

$$(a+b)^n = C_n^0 a^0 b^n + C_n^1 a^1 b^{n-1} + C_n^2 a^2 b^{n-2} + \dots \quad \dots \quad \sum_{k=0}^n C_n^k a^k b^{n-k} = \sum_{k=0}^n C_n^k b^k a^{n-k}$$

Exemple : Développer $(a+b)^2$ puis $(a+b)^8$.

D'après la formule ci-dessus : $(a+b)^2 = C_2^0 a^0 b^2 + C_2^1 a^1 b^1 + C_2^2 a^2 b^0$.

La troisième ligne du triangle de Pascal donne les valeurs de C_2^0 , C_2^1 et C_2^2 . On a donc :

$$(a+b)^2 = b^2 + 2ab + a^2$$

De façon analogue, on obtient directement :

$$(a+b)^8 = b^8 + 8ab^7 + 28a^2b^6 + 56a^3b^5 + 70a^4b^4 + 56a^5b^3 + 28a^6b^2 + 8a^7b + a^8$$

1.4.6. Exemples de probabilités combinatoires

Exemple 1 : Une urne contient 10 boules parmi lesquelles 3 sont rouges et 7 sont blanches. On tire au hasard une boule de l'urne (« au hasard » veut dire que toutes les boules dans l'urne ont la même probabilité d'être tirées). Quelle est la probabilité p de tirer une boule rouge ?

Réponse :

Il y a C_{10}^1 possibilités de tirer une boule dans l'urne (cas possibles) et il y a C_3^1 possibilités de tirer une boule rouge de l'urne (cas favorables), donc :

$$p = \frac{C_3^1}{C_{10}^1} = \frac{3}{10}$$

Exemple 2 : Une urne contient 10 boules parmi lesquelles 3 sont rouges et 7 sont blanches. On tire simultanément 2 boules de l'urne. Quelle est la probabilité p d'obtenir une boule rouge et une boule blanche ?

Réponse :

Il y a C_{10}^2 possibilités de tirer 2 boules de l'urne (cas possibles) et il y a C_3^1 possibilités de tirer une boule rouge et C_7^1 possibilités de tirer une boule blanche, donc il y a $C_3^1 \times C_7^1$ cas favorables, d'où :

$$p = \frac{C_3^1 \times C_7^1}{C_{10}^2} = \frac{7}{15}$$

Exemple 3 : On tire au hasard 3 cartes d'un jeu contenant 52 cartes. Calculer la probabilité p d'obtenir un « as » parmi les cartes tirées.

Réponse :

Il y a C_{52}^3 façons de tirer 3 cartes d'un jeu de 52 cartes. D'autre part, il y a C_4^1 façons de tirer un « as » parmi les 4 « as » du jeu et C_{48}^2 façons de tirer les 2 autres cartes parmi celles qui restent, d'où :

$$p = \frac{C_4^1 \times C_{48}^2}{C_{52}^3}$$

1.5.- Probabilités conditionnelles

Exemple introductif² : Les décisions dans la vie courante sont souvent influencées ou conditionnées par la survenue d'évènements extérieurs. Admettons, par exemple, que pour vous rendre d'une ville à une autre vous utilisiez la voiture, car c'est le moyen de transport le plus rapide. Devant effectuer un trajet, vous apprenez qu'une manifestation est prévue sur la route, vous décidez de prendre le taxi moto car vous pensez que c'est cette fois-ci le moyen de transport le plus rapide.

- La probabilité du choix de tel ou tel moyen de transport est conditionnée par l'évènement M : Manifestation (\bar{M} : non manifestation)
- Si T désigne l'évènement prendre le taxi moto, V désigne l'évènement prendre la voiture ; alors la probabilité de l'évènement T , sachant (ou conditionné par) M , est notée $P(T/M)$.

On a donc :

$$P(T/M) = 1 \quad P(T/\bar{M}) = 0 \quad P(V/M) = 0 \quad P(V/\bar{M}) = 1$$

On considère dans ce paragraphe un univers Ω quelconque sur lequel on a défini une probabilité P .

Définition : A et B désignent deux évènements (avec $P(B)$ non nul), on appelle **probabilité conditionnelle de A sachant B**, le nombre (compris entre 0 et 1) noté $P(A/B)$ défini par :

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

Exemple : Le gérant d'une station libre-service a constaté que, sur une période de six mois, 8% des conducteurs vérifient le niveau de lave-glace, 3% vérifient le liquide de refroidissement et 2%, les deux liquides. Si un conducteur choisi au hasard a vérifié le niveau de lave-glace, quelle est la probabilité que le liquide de refroidissement fasse également l'objet d'une vérification ?

Réponse

Évènement A : *conducteur vérifiant le niveau de lave-glace*

Évènement B : *conducteur vérifiant le liquide de refroidissement*

$$P(A) = 0.08, P(B) = 0.03 \text{ et } P(A \cap B) = 0.02$$

On doit calculer $P(B/A)$.

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{0.02}{0.08} = 0.25$$

Interprétation : Parmi les conducteurs qui vérifient le niveau de lave-glace, 25% vérifient également le liquide de refroidissement.

Remarque : La probabilité de l'évènement B , soit $P(B) = 0.03$, et la probabilité conditionnelle, soit $P(B/A) = 0.25$, sont différentes. Donc, la valeur de la probabilité $P(B)$ est révisée. Dans ce contexte, $P(B)$ est qualifiée de probabilité *a priori* et $P(B/A)$, de probabilité *a posteriori*.

Remarque : La probabilité de réaliser A sachant que B est réalisé est parfois notée $P_B(A)$ au lieu de $P(A/B)$ et on démontre que P_B définit une probabilité sur Ω . Il en résulte les propriétés suivantes, données en parallèles avec les deux notations utilisées :

²Maurice LETHIELLEUX, *Probabilités – Estimation Statistique*, 3^e édition ; Dunod.

Avec la notation $P_B(A)$	Avec la notation $P(A/B)$
Pour tout évènement $A : 0 \leq P_B(A) \leq 1$	Pour tout évènement $A : 0 \leq P(A/B) \leq 1$
$P_B(\Omega) = 1$	$P(\Omega/B) = 1$
Pour tous évènements A_1 et A_2 incompatibles : $P_B(A_1 \cup A_2) = P_B(A_1) + P_B(A_2)$	Pour tous évènements A_1 et A_2 incompatibles : $P((A_1 \cup A_2)/B) = P(A_1/B) + P(A_2/B)$
$P_B(\emptyset) = 0$	$P(\emptyset/B) = 0$
$P_B(\bar{A}) = 1 - P_B(A)$	$P(\bar{A}/B) = 1 - P(A/B)$
$A_1 \subset A_2 \Rightarrow P_B(A_1) \leq P_B(A_2)$	$A_1 \subset A_2 \Rightarrow P(A_1/B) \leq P(A_2/B)$
Pour tous évènements A_1 et A_2 : $P_B(A_1 \cup A_2) + P_B(A_1 \cap A_2) = P_B(A_1) + P_B(A_2)$	Pour tous évènements A_1 et A_2 : $P((A_1 \cup A_2)/B) + P((A_1 \cap A_2)/B)$ $= P(A_1/B) + P(A_2/B)$

1.6.- Théorème de la multiplication

Il découle directement de la définition des probabilités conditionnelles que :

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

Puisque $A \cap B = B \cap A$, on a aussi : $P(B \cap A) = P(B) \times P(A/B)$.

Généralisation : $E_1, E_2, E_3, \dots, E_n$ étant n évènements donnés, on a :

$$P(E_1 \cap E_2 \cap \dots \cap E_n) = P(E_1) \times P(E_2/E_1) \times P(E_3/(E_1 \cap E_2)) \times \dots \times P(E_n/(E_1 \cap E_2 \cap \dots \cap E_{n-1}))$$

Exemple : Pour ouvrir un coffre, on dispose d'un trousseau de 5 clefs presque identiques dont une seule peut ouvrir le coffre. On essaye donc les clefs une à une, en mettant de côté les clefs essayées qui n'ont pas pu ouvrir le coffre. Déterminer la probabilité que l'on parvienne à ouvrir le coffre au deuxième essai.

Réponse

Réussir au deuxième essai suppose que la première clef n'était pas la bonne (évènement que nous noterons A) et que la deuxième est la bonne (évènement B). On cherche donc : $P(A \cap B)$.

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

Or :

- $P(A) = \frac{4}{5}$ (puisque 4 des 5 clefs ne sont pas bonnes)
- $P(B/A) =$ probabilité que la deuxième clef soit bonne sachant que la première ne l'était pas $= \frac{1}{4}$ (puisque'il ne reste plus que 4 clefs dont une seule est bonne)

D'où :

$$P(A \cap B) = \frac{4}{5} \times \frac{1}{4} = 0.2$$

Exemple : Un lot contient 12 articles dont 4 sont défectueux. On tire au hasard trois articles du lot, l'un après l'autre. Calculer la probabilité p pour que les trois articles ne soient pas défectueux.

Réponse

La probabilité pour que le premier article ne soit pas défectueux est $\frac{8}{12}$ puisque 8 articles sur 12 ne sont pas défectueux. Si le premier article n'est pas défectueux, la probabilité pour que l'article suivant ne soit pas défectueux est $\frac{7}{11}$ puisque 7 articles seulement des onze restants ne sont pas

défectueux. Si les deux premiers articles ne sont pas défectueux, alors la probabilité pour que le dernier article ne soit pas défectueux est $\frac{6}{10}$ puisque 6 articles seulement des 10 qui restent ne sont pas défectueux. Par suite, d'après le théorème de la multiplication, on a ;

$$p = \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} = \frac{14}{55}$$

1.7.- Diagramme en arbre

Une suite (finie) d'expériences dont chacune a un nombre fini de résultats possibles avec des probabilités données est appelée un *processus stochastique (fini)*. Un moyen commode de décrire un tel processus et de calculer la probabilité d'un évènement quelconque est d'utiliser comme sur la figure ci-dessous, un diagramme en arbre. Le théorème de la multiplication du paragraphe précédent permet de calculer la probabilité pour que le résultat représenté par un chemin donné de l'arbre soit réalisé.

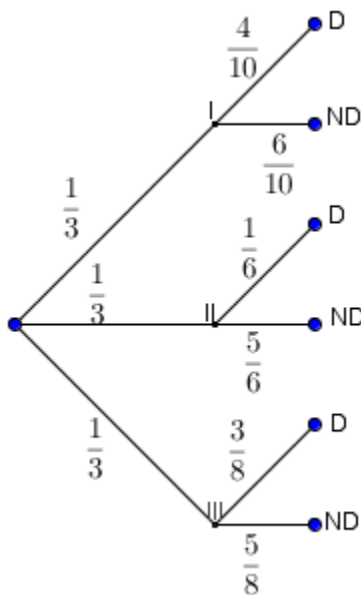
Exemple 1 : On donne trois boites telles que :

- La boite I contient 10 ampoules électriques dont 4 sont défectueuses ;
- La boite II contient 6 ampoules électriques dont 1 est défectueuse ;
- La boite III contient 8 ampoules électriques dont 3 sont défectueuses.

On choisit une boite au hasard et l'on en extrait une ampoule au hasard. Quelle est la probabilité pour que l'ampoule soit défectueuse ?

On représente ici le phénomène par deux expériences :

- On tire l'une des trois boites ;
- On tire une ampoule qui est soit défectueuse (D), soit non défectueuse (ND).



Le diagramme suivant décrit le processus et donne la probabilité de chaque branche de l'arbre.

La probabilité pour qu'un chemin particulier de l'arbre se réalise est, d'après le théorème de la multiplication, le produit des probabilités de chaque branche du chemin. Ainsi, la probabilité de tirer la boite I et ensuite une ampoule défectueuse est : $\frac{1}{3} \times \frac{4}{10} = \frac{4}{30} = \frac{2}{15}$.

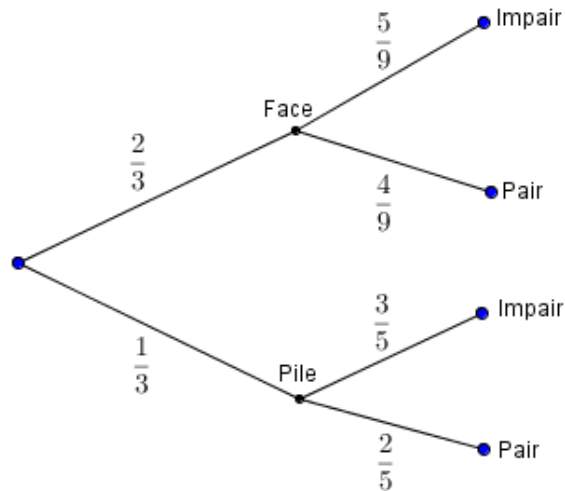
Or, comme il y a trois chemins s'excluant mutuellement qui conduisent à une ampoule défectueuse, la probabilité cherchée est égale à la somme des probabilités de ces chemins :

$$p = \frac{1}{3} \times \frac{4}{10} + \frac{1}{3} \times \frac{1}{6} + \frac{1}{3} \times \frac{3}{8} = \frac{113}{360}$$

Exemple 2 : On jette une pièce de monnaie truquée de telle sorte que $P(\text{Face}) = \frac{2}{3}$ et $P(\text{Pile}) = \frac{1}{3}$. Si c'est face qui apparaît, on choisit au hasard l'un des nombres allant de 1 à 9 ; si c'est pile que l'on obtient, on choisit au hasard l'un des nombres allant de 1 à 5. Calculer la probabilité p pour que ce soit un nombre pair qui ait été choisi.

apparaît, on choisit au hasard l'un des nombres allant de 1 à 9 ; si c'est pile que l'on obtient, on choisit au hasard l'un des nombres allant de 1 à 5. Calculer la probabilité p pour que ce soit un nombre pair qui ait été choisi.

Le diagramme en arbre avec les probabilités correspondantes est représenté ci-dessous.



On remarque que :

- 1) la probabilité de choisir un nombre pair compris entre 1 et 9 est $\frac{4}{9}$ puisqu'il y a 4 nombres pairs parmi les 9 nombres,
- 2) et que la probabilité de choisir un nombre pair entre 1 et 5 est $\frac{2}{5}$ puisqu'il y a 2 nombres pairs parmi ces 5 nombres.

Il y a deux chemins conduisant à un nombre pair : **Face - Pair** et **Pile - Pair**. D'où ;

$$p = \frac{2}{3} \times \frac{4}{9} + \frac{1}{3} \times \frac{2}{5} = \frac{58}{135}$$

1.8.- Probabilités composées et Théorème de Bayes

Supposons que les événements $A_1, A_2, A_3, \dots, A_n$ forment une partition d'un ensemble fondamental Ω ; les événements A_i s'excluent donc mutuellement et leur réunion est Ω . Soit B un autre événement quelconque. Alors ;

$$B = (A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n) \cap B$$

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_n \cap B)$$

où les $A_i \cap B$ s'excluent aussi mutuellement. Par conséquent,

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \dots + P(A_n \cap B)$$

D'après le théorème de la multiplication, on obtient :

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \dots + P(A_n)P(B|A_n) \quad (1)$$

C'est la formule des probabilités totales.

Pour un i donné, la probabilité conditionnelle de A_i , B étant réalisé, est définie par :

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)}$$

Dans cette équation, on remplace $P(B)$ par (1) et $P(A_i \cap B)$ par $P(A_i \cap B) = P(A_i) P(B|A_i)$, ce qui donne :

Théorème (de Bayes)

Soit $A_1, A_2, A_3, \dots, A_n$ une partition de Ω , et B un événement quelconque. Pour tout i , on a alors ;

$$P(A_i|B) = \frac{P(A_i) P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \dots + P(A_n)P(B|A_n)}$$

ou encore

$$P(A_i|B) = \frac{P(A_i) P(B|A_i)}{\sum_i P(B|A_i) \cdot P(A_i)}$$

En utilisant les probabilités *a priori* $P(A_1), P(A_2), \dots, P(A_n)$ et les probabilités conditionnelles appropriées $P(B|A_1), P(B|A_2), \dots, P(B|A_n)$, le théorème de Bayes permet de calculer les probabilités *a posteriori* des événements $A_1, A_2, A_3, \dots, A_n$.

Exemple 1 : Une entreprise manufacturière possède deux fournisseurs différents. Actuellement, 65% des pièces achetées par l'entreprise proviennent du fournisseur 1 et les 35 % restant proviennent du fournisseur 2. La qualité des pièces achetées varie en fonction du fournisseur. Les données historiques révèlent les niveaux de qualité présentés dans le tableau ci-dessous.

	Pourcentage de pièces de bonne qualité	Pourcentage de pièces défectueuses
Fournisseur 1	98	2
Fournisseur 2	95	5

Supposons maintenant que les pièces des deux fournisseurs soient utilisées dans le système de production de l'entreprise et que l'une des machines tombe en panne à cause d'une pièce défectueuse. Sachant que la pièce est défectueuse, quelle est la probabilité qu'elle provienne du fournisseur 1 ? Du fournisseur 2 ?

Réponse

Désignons par :

A_1 : l'évènement « la pièce est fournie par le fournisseur 1 »

A_2 : l'évènement « la pièce est fournie par le fournisseur 2 »

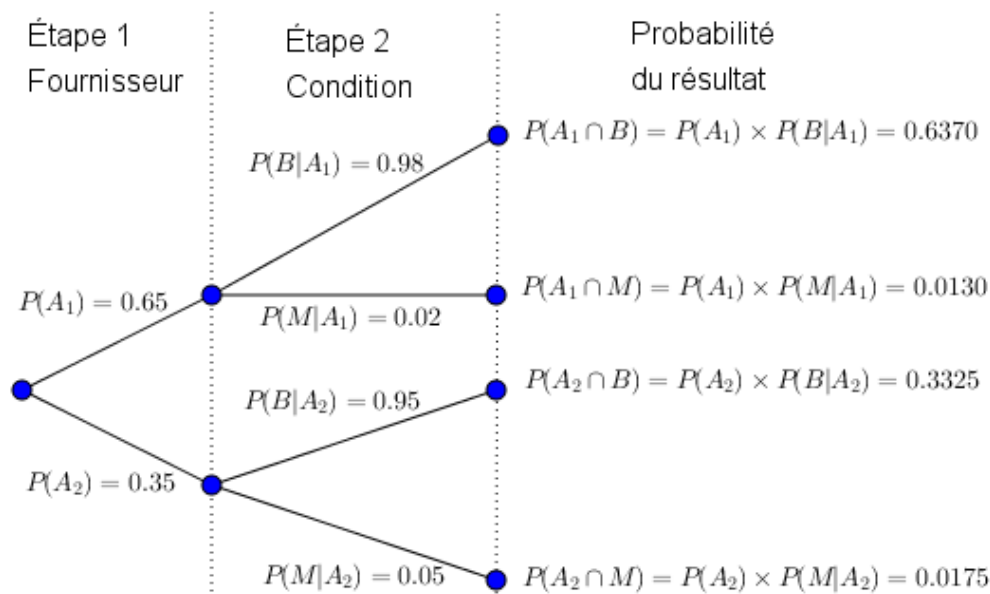
B : l'évènement « la pièce est de bonne qualité »

M : l'évènement « la pièce est défectueuse ».

Avec les données de l'exercice, on obtient les probabilités suivantes :

$P(A_1) = 0.65$; $P(A_2) = 0.35$; $P(B|A_1) = 0.98$; $P(B|A_2) = 0.95$; $P(M|A_1) = 0.02$ et $P(M|A_2) = 0.05$.

Le diagramme arborescent suivant décrit le processus de réception d'une pièce de l'un des deux fournisseurs et de contrôle de sa qualité, comme une expérience en deux étapes. Quatre résultats sont possibles ; deux correspondent à une pièce de bonne qualité et deux correspondent à une pièce de mauvaise qualité.



Remarque : L'étape 1 indique la provenance de la pièce et l'étape 2 indique la qualité de la pièce.

On cherche à déterminer les probabilités a posteriori $P(A_1|M)$ et $P(A_2|M)$ où M correspond à l'évènement « la pièce est défectueuse ».

Par la loi des probabilités conditionnelles, on sait que :

$$P(A_1|M) = \frac{P(A_1 \cap M)}{P(M)} \quad (1)$$

En se référant à l'arbre des probabilités, on note que : $P(A_1 \cap M) = P(A_1) \times P(M|A_1)$ (2).
Puisque l'évènement M ne se produit que dans les deux cas $A_1 \cap M$ et $A_2 \cap M$, on a donc :

$$P(M) = P(A_1 \cap M) + P(A_2 \cap M) = P(A_1) \times P(M|A_1) + P(A_2) \times P(M|A_2) \quad (3)$$

En remplaçant les équations (2) et (3) dans l'équation (1), on obtient le théorème de Bayes :

$$P(A_1|M) = \frac{P(A_1) \times P(M|A_1)}{P(A_1) \times P(M|A_1) + P(A_2) \times P(M|A_2)} = \frac{0.65 \times 0.02}{(0.65 \times 0.02) + (0.35 \times 0.05)}$$

$$P(A_1|M) = \mathbf{0.4262}$$

Si une pièce est choisie au hasard dans la population, la probabilité qu'elle soit venue du fournisseur 1 est 0.65. Si la pièce est défectueuse, la probabilité qu'elle provienne du fournisseur 1 passe de 0.65 à 0.4262.

On a de même :

$$P(A_2|M) = \frac{P(A_2) \times P(M|A_2)}{P(A_1) \times P(M|A_1) + P(A_2) \times P(M|A_2)} = \frac{0.35 \times 0.05}{(0.65 \times 0.02) + (0.35 \times 0.05)}$$

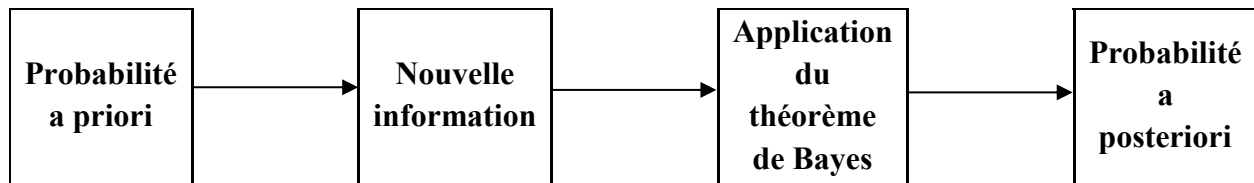
$$P(A_2|M) = \mathbf{0.5738}$$

On a commencé avec une probabilité de 0.65 qu'une pièce aléatoirement sélectionnée provienne du fournisseur 1. Cependant, sachant que la pièce est défectueuse, la probabilité qu'elle provienne du fournisseur 1 chute à 0.4262. En fait, si la pièce est défectueuse, il y a plus d'une chance sur deux qu'elle provienne du fournisseur 2, soit $P(A_2|M) = \mathbf{0.5738}$.

Note : La valeur de la probabilité initiale 0.65 s'appelle **probabilité a priori**. Elle est fondée sur des données historiques. La valeur révisée 0.4262 s'appelle **probabilité a posteriori**, car elle est basée sur des informations complémentaires.

Probabilité a posteriori : Probabilité révisée à partir d'informations complémentaires.

La figure suivante illustre les étapes du processus de révision des probabilités en utilisant le théorème de Bayes



Exemples d'application des différentes formules

Une usine d'embouteillage fonctionne avec 5 machines :

- La machine A_1 fournit 15% de la production globale ;
- La machine A_2 fournit 25 % de la production globale ;
- La machine A_3 fournit 20% de la production globale ;
- La machine A_4 fournit 20% de la production globale
- La machine A_5 fournit 20% de la production globale

Au sortir de ces machines, la production des bouteille à rejeter est de :

- 1% pour A_1
- 2% pour A_2
- 1.5% pour A_3
- 1 % pour A_4
- 0.5% pour A_5

Si nous notons :

- B : l'évènement des bouteilles acceptables (bonnes)
- D : l'ensemble des bouteilles défectueuses

Imaginons qu'on prélève au hasard une bouteille de la production. Nous savons que :

- La probabilité qu'elle provienne de $A_1 = P(A_1) = 0.15$;
- La probabilité qu'elle provienne de $A_2 = P(A_2) = 0.25$;
- La probabilité qu'elle provienne de $A_3 = P(A_3) = 0.20$;
- La probabilité qu'elle provienne de $A_4 = P(A_4) = 0.20$;
- La probabilité qu'elle provienne de $A_5 = P(A_5) = 0.20$.

Nous savons aussi que :

- $P(D|A_1) = 0.01$;
- $P(D|A_2) = 0.02$;
- $P(D|A_3) = 0.015$;
- $P(D|A_4) = 0.01$;
- $P(D|A_5) = 0.005$.

Application de la formule de la multiplication : Quelle est la probabilité que cette bouteille soit défectueuse et qu'elle provienne de A_1 ?

$$P(D \cap A_1) = P(A_1 \cap D) = P(A_1) \cdot P(D|A_1) = 0.15 \times 0.01 = 0.0015$$

Application du théorème des probabilités totales : Quelle est la probabilité que cette bouteille soit défectueuse ?

$$P(D) = \sum_{i=1}^5 P(A_i)P(D|A_i)$$

$$P(D) = (0.01 \times 0.15) + (0.02 \times 0.25) + (0.0015 \times 0.20) + (0.01 \times 0.20) + (0.005 \times 0.20) \\ = 0.0125$$

Application du théorème de Bayes : Si cette bouteille est défectueuse, quelle est la probabilité qu'elle provienne de la machine A_1 ?

$$P(A_1|D) = \frac{P(A_1) \cdot P(D|A_1)}{\sum_{i=1}^5 P(A_i)P(D|A_i)} = \frac{0.15 \times 0.01}{0.0125} = 0.12$$

Exercice 1 : On considère deux urnes, l'une contient une bille noire et une blanche, et l'autre deux noires et une blanche. On désigne une urne au hasard, de laquelle on tire une bille.

- Quelle est la probabilité qu'elle soit noire ?
- Si l'on sait que la bille est blanche, quelle est la probabilité que ce soit la première urne qui ait été désignée ?

Exercice 2 : On dispose de 3 pièces de monnaie : la première est parfaitement équilibrée, la seconde a 2 côtés « pile », la troisième est truquée de façon que la probabilité d'obtenir « pile » soit de $\frac{1}{4}$. On tire une pièce au hasard, on la lance et elle donne « pile ». Quelle est la probabilité d'avoir lancé la deuxième pièce ?

1.9.- Évènements indépendants pour une probabilité P

A. Cas de deux évènements

Intuitivement, deux évènements A et B sont indépendants si la réalisation de l'un n'a pas d'influence sur la réalisation de l'autre. Dans ce cas, la probabilité de réaliser A sachant que B est réalisé doit valoir $P(A)$, et puisque $P(A/B) = \frac{P(A \cap B)}{P(B)}$ (en supposant $P(B) \neq 0$), on doit avoir :

$$P(A) = \frac{P(A \cap B)}{P(B)}, \text{ ce qui conduit à la définition suivante :}$$

A et B sont deux évènements indépendants si et seulement si : $P(A \cap B) = P(A) \times P(B)$.

Remarque :

- Veiller à ne pas confondre « évènements indépendants » et « évènements incompatibles » ;
- Si les probabilités $P(A)$ et $P(B)$ sont non nulles, on peut donner la définition sous la forme :
 - A et B sont deux évènements indépendants si et seulement si $P(A/B) = P(A)$ ou encore ;
 - A et B sont deux évènements indépendants si et seulement si $P(B/A) = P(B)$.

Exemple : On considère une population de jeunes de 25 ans dont 80% sont des étudiants, 70% sont vaccinés contre l'hépatite B, et 10% possèdent une moto. De plus 56% sont des étudiants

vaccinés contre l'hépatite B, et 8% ont une moto et sont vaccinés contre l'hépatite B. on prend au hasard un élément de cette population et on note E, V et M les évènements suivants :

E : « obtenir un étudiant » ; V : « obtenir un jeune qui est vacciné contre l'hépatite B » ;

M : « obtenir un jeune qui possède une moto ».

Les évènements E et V d'une part, et M et V d'autre part sont-ils indépendants ?

Réponse

Compte des données, on a : $P(E) = 0.8$; $P(V) = 0.7$; $P(M) = 0.1$; $P(E \cap V) = 0.56$ et $P(M \cap V) = 0.08$. Avec ces données, on peut écrire que :

- $P(E \cap V) = 0.56 = P(E) \times P(V)$ et en déduire que dans cette population, les évènements « être étudiant » et « être vacciné contre l'hépatite B » sont indépendants en probabilité.
- Par contre $P(M \cap V) = 0.08 \neq P(M) \times P(V) = 0.07$, les évènements « avoir une moto » et « être vacciné contre l'hépatite B » ne sont donc pas indépendants en probabilité.

Propriété :

1) Si A et B sont indépendants pour P alors :

- a. $P(A/B) = P(A)$ si $P(B) \neq 0$
- b. $P(B/A) = P(B)$ si $P(A) \neq 0$
- c. $P(A \cup B) = P(A) + P(B) - P(A) \times P(B)$

2) Si A et B sont indépendants, alors :

- A et \bar{B} sont indépendants ;
- \bar{A} et B sont indépendants ;
- \bar{A} et \bar{B} sont indépendants.

B. Cas de plusieurs évènements

Considérons une famille $A_1, A_2, A_3, \dots, A_n$ d'évènements contenus dans un même univers.

Les évènements $A_1, A_2, A_3, \dots, A_n$ sont **indépendants deux à deux** si l'on a pour tout A_i et tout A_j (avec $i \neq j$) : $P(A_i \cap A_j) = P(A_i) \times P(A_j)$.

Les évènements $A_1, A_2, A_3, \dots, A_n$ sont **mutuellement indépendants** si l'on a, pour toute sous famille $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ (avec $k \geq 2$) :

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \times P(A_{i_2}) \times \dots \times P(A_{i_k})$$

Exemple : On lance une pièce de monnaie jusqu'à ce que l'on obtienne le côté pile. Déterminer la probabilité d'obtenir pile pour la première fois au 5^e essai. En admettant qu'à chaque lancer, on a autant de chances d'obtenir face que pile, et que les résultats des lancers sont mutuellement indépendants.

Réponse

Notons par :

F_i : l'évènement « obtenir le côté face au i ème lancer », et

P_i : l'évènement « obtenir le côté pile au i ème lancer »

On cherche $P(F_1 \cap F_2 \cap F_3 \cap F_4 \cap P_5)$.

On a donc :

$$P(F_1 \cap F_2 \cap F_3 \cap F_4 \cap P_5) = P(F_1) \times P(F_2) \times P(F_3) \times P(F_4) \times P(P_5) = (0.5)^5 = 0.03125$$

Remarque : Si les évènements $A_1, A_2, A_3, \dots, A_n$ sont mutuellement indépendants, ils sont indépendants deux à deux. Mais la réciproque n'est pas toujours vraie.

Exemple : On jette deux pièces de monnaie bien équilibrées. $\Omega = \{PP, PF, FP, FF\}$ est un ensemble équiprobable. Considérons les évènements :

$$A = \{\text{face apparaît sur la première pièce}\} = \{FP, FF\}$$

$$B = \{\text{face apparaît sur la deuxième pièce}\} = \{PF, FF\}$$

$$C = \{\text{face apparaît sur une seule des deux pièces}\} = \{PF, FP\}$$

On a alors : $P(A) = P(B) = P(C) = \frac{2}{4} = \frac{1}{2}$ et $P(A \cap B) = P(\{FF\}) = \frac{1}{4}$, $P(A \cap C) = P(\{FP\}) = \frac{1}{4}$,

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

Les évènements A, B et C sont donc indépendants deux à deux.

Cependant $A \cap B \cap C = \emptyset$, de sorte que $P(A \cap B \cap C) = P(\emptyset) = 0 \neq P(A) \times P(B) \times P(C)$. Alors les trois évènements ne sont pas mutuellement indépendants.

Chapitre 2.- Généralités sur les variables aléatoires

Introduction³ : Dans de nombreuses expériences aléatoires, l'attention se porte non sur le résultat lui-même, mais plutôt sur une valeur numérique associée au résultat. Par exemple, le responsable du contrôle de la qualité a pour principal objectif de maintenir dans des limites acceptables la proportion des articles fabriqués qui ne satisfont pas aux normes de qualité (appelés ici articles défectueux). L'entreprise s'intéresse à ces articles défectueux parce qu'ils nuisent à sa réputation et à la satisfaction du client, et parce que souvent, ils sont retournés afin d'être réparés durant la période de garantie, ce qui augmente les coûts. Ainsi, lorsqu'un inspecteur du contrôle de la qualité évalue un lot fabriqué, il se concentre principalement sur le nombre d'articles défectueux dans le lot. Il s'agit d'une variable aléatoire qui peut prendre les valeurs 0, 1, 2, 3 et ainsi de suite.

2.1. Définition

Soit (Ω, \mathcal{F}, P) un espace probabilisé. On appelle **variable aléatoire** toute application permettant d'associer un nombre réel à toute éventualité.

On note $X(\Omega)$ ou $S(X)$ l'ensemble de toutes les valeurs que peut prendre X , on l'appelle « **support de X** ».

Par exemple, on lance un dé cubique à 6 faces numérotées de 1 à 6. Soit X le résultat obtenu. Ici X est une variable aléatoire et les valeurs possibles de X sont : 1, 2, 3, 4, 5, 6. Pour chacune de ces valeurs, X a une certaine probabilité de lui être égal.

2.2. Les types de variables aléatoires

Nous étudierons deux types de variables aléatoires : *les variables aléatoires discrètes et les variables aléatoires continues.*

Définition : Soit X une variable aléatoire, alors X est dite :

- **Discrète** lorsque l'ensemble des valeurs qu'elle peut prendre est fini ou dénombrable,
- **Continue** lorsqu'elle peut prendre toutes les valeurs d'un intervalle de \mathbb{R} .

Exemples :

- a) Voici quelques exemples de variables aléatoires discrètes :
- Le nombre d'enfants dans une famille sélectionnée au hasard ;
 - Le nombre de voitures vendues par un vendeur un jour donné ;

³ Méthodes statistiques pour les sciences de la gestion, Chenelière McGraw-Hill.

- Le nombre d'articles défectueux produits dans une journée.
- b) Voici quelques exemples de variables aléatoires continues :
- La taille en cm d'une personne sélectionnée au hasard ;
 - Le délai d'attente d'un client pour recevoir une autorisation de carte de crédit dans un magasin ;
- c) On lance deux dés à 6 faces chacun numérotées de 1 à 6. On note S l'application qui, à chaque lancer, associe la somme des résultats obtenus.
 S est une variable aléatoire qui peut prendre toutes les valeurs de l'ensemble :
- $$S(\Omega) = \{2, 3, 4, \dots, 11, 12\}$$
- S est discrète.
- d) On lance plusieurs fois de suite deux dés jusqu'à ce que l'on obtienne un double 6. On note X l'application qui, à chaque série de lancers, associe le nombre d'essais qu'il a fallu faire.
 X est une variable aléatoire discrète qui peut prendre toutes les valeurs entières non nulles.
 $X(\Omega) = \mathbb{N}^*$
- e) Exemples de variables aléatoires discrètes ou continues

Expérience	Variable aléatoire (X)	Type de variable aléatoire	Valeurs que peut prendre la variable aléatoire
Contacter cinq clients	Nombre de clients qui passent commande	Discrète	0, 1, 2, 3, 4, 5
Inspecter une cargaison de 50 radios	Nombre de radios défectueuses	Discrète	0, 1, 2, ..., 49, 50
Gérer un restaurant pendant une journée	Nombre de clients	Discrète	0, 1, 2, 3, ...
Vendre une automobile	Sexe des clients	Discrète	0 si le client est un homme ; 1 si client est une femme.
Gérer une banque	Temps écoulés entre les arrivées des clients en minutes	Continue	$x \geq 0$
Remplir une canette de soda ($max = 30 cl$)	Nombre de centilitres	Continue	$0 \leq x \leq 30$

2.3. Variables aléatoires discrètes

Définition : Une variable aléatoire est discrète si elle est numérique ($E = \mathbb{R}$) et si l'ensemble de ses valeurs est dénombrable $X(\Omega) = \{x_1, \dots, x_N\}$ ou $\{x_n | n \in \mathbb{N}\}$.

Une variable aléatoire discrète est définie par :

- Ses valeurs $\{x_1, \dots, x_N\}$ ou $\{x_n | n \in \mathbb{N}\}$ appelées support de X ;
- Ses probabilités $p_i = P(X = x_i)$.

2.3.1. Loi de probabilité ou fonction de distribution

1. Définition

L'application, qui à chaque valeur possible x d'une variable aléatoire X , associe la probabilité $p_X(x) = P(X = x)$ est appelée **loi de probabilité** ou **fonction de distribution** de la variable aléatoire X .

Remarque : La loi de probabilité de X est l'ensemble des couples (x_k, p_k) ; elle peut être présentée dans un tableau.

x	x_1	x_2	...	x_i	...	x_k	Total
$\mathbb{P}(X = x)$	p_1	p_2	...	p_i	...	p_k	1

Pour tout x_k , $\mathbb{P}(X = x_k) = p_k \geq 0$ et $\sum_k p_k = 1$.

Une loi de probabilité d'une variable aléatoire discrète est définie par les couples (x_k, p_k) avec $\mathbb{P}(X = x_k) = p_k \geq 0$ et $\sum_k p_k = 1$.

Remarque : On peut représenter graphiquement les couples $(x, p_X(x))$, ce qui permet de visualiser la distribution de probabilité d'une variable aléatoire discrète.

Remarque :

- Veiller à ne pas confondre X et x ;
 - X désigne la variable aléatoire considérée, et x une valeur possible de cette variable ;
- Veiller également à distinguer la signification de $P(X = x)$ et celle de $X = x$;
 - $(X = x)$ désigne l'évènement « la variable aléatoire X prend la valeur x », tandis que $P(X = x)$ est la probabilité de réaliser cet évènement.

2. Propriété

Conditions requises pour une fonction de probabilité discrète :

- $P(X = x) \geq 0$, pour tout $x \in S_X = X(\Omega)$;
- La somme des $P(X = x)$ vaut 1. On note : $\sum_{x \in X(\Omega)} P(X = x) = 1$.

3. Exemples

- 1) Supposons qu'on s'intéresse au nombre de faces obtenues après trois lancers d'une pièce de monnaie honnête effectués dans les mêmes conditions. Il s'agit d'une variable aléatoire discrète. Les valeurs possibles de la variable sont 0, 1, 2 et 3. Quelle est la distribution discrète du nombre de faces ?

Réponse

Donner la distribution discrète du nombre de faces revient à donner pour chaque valeur de la variable aléatoire la probabilité correspondante.

- ➔ L'ensemble des résultats possibles à l'expérience aléatoire est :

$$\Omega = \{FFF, FFP, FPF, FPP, PFF, PFP, PPF, PPP\}$$

- ➔ Les résultats possibles de la variable « nombre de faces obtenues » est :

$$X(\Omega) = \{0, 1, 2, 3\}$$

- Puisque la pièce de monnaie est honnête, chacun des huit résultats possibles est également probable et la probabilité de chacun est de $\frac{1}{8}$. On a donc :

$$P(X = 0) = P(\{PPP\}) = \frac{1}{8} ; P(X = 1) = P(\{FPP, PFP, PPF\}) = \frac{3}{8}$$

$$P(X = 2) = P(\{FFP, FPF, PFF\}) = \frac{3}{8} ; P(X = 3) = P(\{FFF\}) = \frac{1}{8}$$

- Le tableau suivant présente la distribution cherchée :

Nombre de faces : x	0	1	2	3	Total
Probabilité : $P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

2) Soit X une variable aléatoire discrète prenant les valeurs $2, \frac{9}{2}, 6$ et 8 , c'est-à-dire $X(\Omega) = \left\{2, \frac{9}{2}, 6, 8\right\}$. Déterminons la loi de probabilité de X sachant que :

$$\mathbb{P}(X < 6) = \frac{1}{3}, \mathbb{P}(X > 6) = \frac{1}{2}, \mathbb{P}(X = 2) = \mathbb{P}\left(X = \frac{9}{2}\right)$$

Pour déterminer la loi de probabilité de X , on va procéder de la manière suivante :

x	2	$\frac{9}{2}$	6	8
$p_X(x)$	a	b	c	d

On sait que :

- $\mathbb{P}(X < 6) = \frac{1}{3}$, donc on en déduit que $a + b = \frac{1}{3}$.
- $\mathbb{P}(X > 6) = \frac{1}{2}$, donc on en déduit que $d = \frac{1}{2}$.
- $\mathbb{P}(X = 2) = \mathbb{P}\left(X = \frac{9}{2}\right)$, donc on en déduit que $a = b$.

De plus, on a la relation suivante : $a + b + c + d = 1$ car $\sum_{x \in X(\Omega)} p_X(x) = 1$. On en déduit que :

$$a = b = c = \frac{1}{6} \text{ et } d = \frac{1}{2}.$$

D'où, la loi de probabilité de X est l'ensemble des couples (x_k, p_k) représentée dans le tableau suivant :

x	2	$\frac{9}{2}$	6	8
$p_X(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$

2.3.2. Fonction de répartition

1. Définition

L'application F , qui a tout réel x , associe la probabilité que la variable aléatoire X prenne une valeur inférieure ou égale à x , est, par définition, la fonction de répartition de X .

Pour tout x réel ; on note définit $F(x) = P(X \leq x)$.

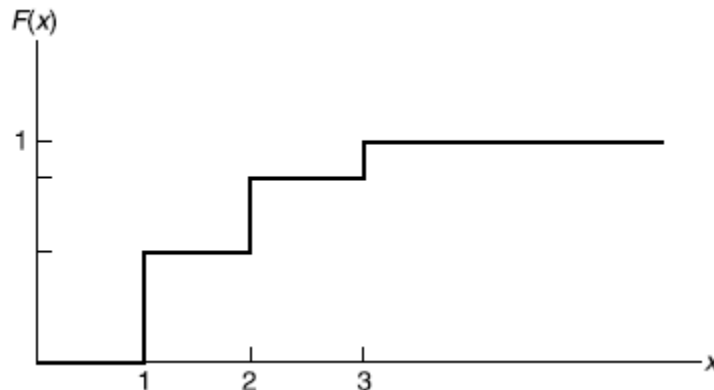
$$F(x) = P(X \leq x) = P(X \in]-\infty, x]) = \sum_{x_i \leq x} p_i$$

Elle est aussi définie par :

$$F(x) = P(X < x) = P(X \in]-\infty, x]) = \sum_{x_i < x} p_i$$

Puisque cette dernière définition modifie la valeur de $F(x)$ aux points $X = x_i$ où la probabilité n'est pas nulle, nous gardons la première définition qui est devenue la plus usuelle.

F est une fonction croissante en escalier, présentant des sauts de p_i en chaque x_i .



Remarque :

$F(x)$ permet de retrouver les probabilités :

$$p_i = F(x_i) - F(x_{i-1}) = (p_1 + \dots + p_{i-1} + p_i) - (p_1 + \dots + p_{i-1})$$

2. Propriétés

- F est une fonction croissante ;
- Pour tout x , on a $0 \leq F(x) \leq 1$;
- $\lim_{x \rightarrow +\infty} F(x) = 1$: on pourrait écrire cette limite sous la forme « $P(X \leq +\infty)$ or $(X \leq +\infty)$ est l'évènement certain, donc $P(X \leq +\infty) = 1$ » ;
- $\lim_{x \rightarrow -\infty} F(x) = 0$: de même $(X \leq -\infty)$ est l'évènement impossible, donc $P(X \leq -\infty) = 0$;
- $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$;
- Pour tous a et b réels : $P(a < X \leq b) = F(b) - F(a)$.

Quelques égalités utiles

- 1) Pour tout réel a , $\mathbb{P}(-\infty < X \leq a) = F_X(a)$;
- 2) Pour tout réel a , $\mathbb{P}(-\infty < X < a) = F_X(a) - \mathbb{P}(X = a)$;
- 3) Pour tout réel a , $\mathbb{P}(a < X < +\infty) = 1 - F_X(a)$;
- 4) Pour tout réel a , $\mathbb{P}(a \leq X < +\infty) = 1 - F_X(a) + \mathbb{P}(X = a)$;
- 5) Pour tous réels a et b , $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$;
- 6) Pour tous réels a et b , $\mathbb{P}(a \leq X \leq b) = F_X(b) - F_X(a) + \mathbb{P}(X = a)$;
- 7) Pour tous réels a et b , $\mathbb{P}(a < X < b) = F_X(b) - F_X(a) - \mathbb{P}(X = b)$;
- 8) Pour tous réels a et b , $\mathbb{P}(a \leq X < b) = F_X(b) - F_X(a) - \mathbb{P}(X = b) + \mathbb{P}(X = a)$.

Remarques :

- Dans le cas où X est discrète, il est important dans la dernière propriété de veiller aux inégalités (l'une stricte, l'autre large) ;
- Si l'on connaît la loi ou distribution de X , on peut déterminer sa fonction de répartition F et réciproquement ;
- Le graphe de la fonction de répartition prendra la forme d'un escalier.

Exemple : Dans une population d'étudiants, 10% d'entre eux sont abonnés à 3 hebdomadaires, 20% à 2 hebdomadaires, 40% à un seul, les autres n'ayant aucun abonnement. On note X la variable aléatoire qui, à chaque étudiant, associe le nombre d'abonnements qu'il a souscrits.

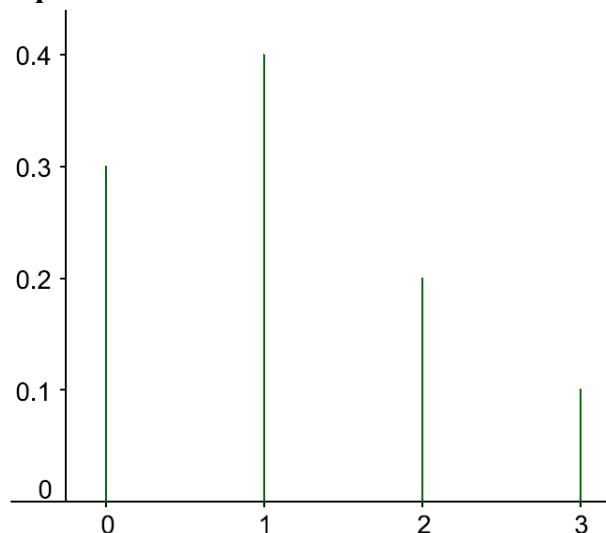
- Déterminer la loi de probabilité de la variable aléatoire X puis la représenter graphiquement.
- F désignant la fonction de répartition de X , calculer $F(0)$, $F(1)$, $F(2)$, $F(3)$ puis donner l'expression de F puis la représenter sur \mathbb{R} .
- Utiliser F pour calculer : $P(X > 2)$, $P(0 < X < 3)$, $P(1 \leq X < 3)$ et $P(0 < X \leq 2)$.

Réponse

- La loi probabilité de X est :

x	0	1	2	3
$P(X = x)$	0.3	0.4	0.2	0.1

Sa représentation graphique est :



- On a donc :

- $F(0) = P(X \leq 0) = P(X = 0) = 0.3$
- $F(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = 0.3 + 0.4 = 0.7$
- $F(2) = P(X \leq 2) = 0.9$ et $F(3) = P(X \leq 3) = 1$
- Expression de F sur \mathbb{R}**

$$\begin{cases} \text{Si } x < 0 & F(x) = 0 \\ \text{Si } x \in [0,1[& F(x) = 0.3 \\ \text{Si } x \in [1,2[& F(x) = 0.7 \\ \text{Si } x \in [2,3[& F(x) = 0.9 \\ \text{Si } x \geq 3 & F(x) = 1 \end{cases}$$

- Lorsque X est discrète, F est une fonction en escalier
- Représentation graphique de F

c) Calculer les probabilités en utilisant F

- $P(X > 2) = 1 - P(X \leq 2) = 1 - F(2) = 1 - 0.9 = 0.1$
- $P(0 < X < 3) = P(1 \leq X < 3) = P(0 < X \leq 2) = F(2) - F(0) = 0.9 - 0.3 = 0.6$

2.3.3. Variables aléatoires discrètes indépendantes

Définition : Deux variables aléatoires discrètes X et Y sont indépendantes si et seulement si pour tous couples (x, y) de valeurs possibles pour X et Y , les événements $(X = x)$ et $(Y = y)$ sont indépendants, ce qui s'écrit :

$$P(X = x \text{ et } Y = y) = P(X = x) \times P(Y = y)$$

Note : X et Y sont indépendantes si et seulement si, pour tous réels x et y , on a :

$$P(X \leq x \text{ et } Y \leq y) = P(X \leq x) \times P(Y \leq y)$$

2.3.4. Espérance - variance - écart-type d'une variable aléatoire discrète - covariance de deux variables aléatoires discrètes

Définitions : Soit X une variable aléatoire discrète définie sur Ω .

1. Espérance mathématique

L'espérance mathématique de X , notée $E(X)$ ou \bar{X} , est définie par :

$$\bar{X} = E(X) = \sum_{x \in X(\Omega)} x P(X = x)$$

$E(X)$ est un paramètre de position ou de tendance centrale de X .

2. Moments non centrés d'ordre h d'une variable aléatoire discrète : $E(X^h)$

Définition :

Si la fonction g est continue : $E(g(X)) = \sum_i p_i g(x_i)$

Si $g(X) = X^h$, on obtient $m_h = E(X^h)$ moment non centré d'ordre h .

$$m_h = E(X^h) = p_1 x_1^h + p_2 x_2^h + \dots + p_i x_i^h + \dots + p_n x_n^h = \sum_i p_i x_i^h$$

h est un nombre entier

- Si $h = 1$, le moment d'ordre 1 est : $m_1 = m = E(X)$
- Si $h = 2$, le moment d'ordre 2 est : $m_2 = E(X^2)$, il sert au calcul de la variance.

3. Moments non centrés d'une variable aléatoire discrète prenant une infinité de valeurs

Si X prend une infinité de valeurs $x_1, \dots, x_i, \dots, x_n, \dots$ sont d'ordre infini, le moment d'ordre h est défini par :

$$m_h = E(X^h) = \lim_{n \rightarrow \infty} \sum_{i=1}^n p_i x_i^h$$

Si cette expression n'a pas de limite, le moment d'ordre h n'existe pas.

De nombreux résultats en statistique mathématique et en probabilité sont liés à l'existence des moments d'ordre 1 et 2 c'est-à-dire $E(X)$ et $E(X^2)$.

4. Variance et écart-type

La variance de X , notée $V(X)$, est l'espérance de la variable aléatoire $(X - \bar{X})^2$:

$$V(X) = E((X - \bar{X})^2)$$

On démontre que :

$$V(X) = \sum_{x \in X(\Omega)} (x - \bar{X})^2 P(X = x)$$

Pour les calculs numériques, on utilise généralement l'expression suivante de $V(X)$:

$$V(X) = E(X^2) - \bar{X}^2 = \sum_{x \in X(\Omega)} x^2 P(X = x) - \bar{X}^2 \quad (\text{Formule de Koenig})$$

On remarque que $V(X)$ est un nombre positif puisque c'est une somme de termes positifs ou nuls.

On peut donc en prendre la racine carrée qui est par définition l'écart type de X , notée $\sigma(X)$:

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

Remarques :

- a) $E(X)$ et $\sigma(X)$ ont même unité ;
- b) $V(X)$ et $\sigma(X)$ sont des caractéristiques qui mesurent la dispersion des valeurs prises par X autour de son espérance \bar{X} .

5. Moment centré d'ordre h

Par définition le moment centré d'ordre h par rapport à $E(X) = m$ est μ_h tel que :

$$\mu_h = E[(X - m)^h] = \sum_i p_i (x_i - m)^h$$

- Si $h = 1$, $\mu_1 = E(X - m) = E(X) - m = 0$
- Si $h = 2$, $\mu_2 = E[(X - m)^2]$ s'appelle la variance (*moment centré d'ordre 2*)

6. Covariance de deux variables aléatoires définies sur un même univers

La covariance de X et de Y , notée $cov(X, Y)$, est l'espérance de la variable aléatoire $(X - \bar{X})(Y - \bar{Y})$:

$$cov(X, Y) = E((X - \bar{X})(Y - \bar{Y}))$$

On démontre que :

$$cov(X, Y) = \sum \sum (x - \bar{X})(y - \bar{Y}) P(X = x \text{ et } Y = y)$$

On utilise généralement l'expression suivante de $cov(X, Y)$, dans la pratique :

$$\text{cov}(X, Y) = E(XY) - \bar{X}\bar{Y} = \sum \sum xy P(X = x \text{ et } Y = y) - \bar{X}\bar{Y}$$

Interprétation de la covariance : La covariance permet d'évaluer l'intensité de la dépendance statistique entre deux variables aléatoires. La covariance peut être positive (les deux variables varient dans le même sens), négative (les deux variables varient en sens contraires) ou nulle et sa valeur se situe entre $-\infty$ et $+\infty$, ce qui la rend difficile d'interprétation.

7. Coefficient de corrélation linéaire

On définit le coefficient de corrélation ρ entre deux variables X et Y par le rapport :

$$\rho = E \left[\left(\frac{X - E(X)}{\sigma(X)} \right) \left(\frac{Y - E(Y)}{\sigma(Y)} \right) \right] = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}, -1 \leq \rho \leq 1$$

Interprétation du coefficient de corrélation linéaire (ρ) : Le coefficient de corrélation est une mesure de l'intensité de la liaison linéaire entre deux variables aléatoires. ρ peut varier entre -1 (corrélation parfaite négative) et $+1$ (corrélation parfaite positive) : $-1 \leq \rho \leq 1$. Si $\rho = 0$, on dit que les variables X et Y sont non corrélées (absence de liaison linéaire).

8. Propriétés

8.1. Propriétés de l'espérance

➤ X étant une variable aléatoire, a et b deux constantes, on a :

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

Et en particulier, l'espérance d'une variable aléatoire constante est égale à cette constante.

➤ L'espérance d'une somme de variables aléatoires est égale à la somme de leurs espérances :

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

Conséquence : $E(X - Y) = E(X) - E(Y)$

8.2. Propriétés de la covariance

➤ $\text{cov}(X, k) = 0$ et $\text{cov}(k, Y) = 0$, k étant une constante donnée.

➤ $\text{cov}(X, Y) = \text{cov}(Y, X)$

➤ $\text{cov}(aX + b, a'Y + b') = aa' \text{cov}(X, Y)$; a, a', b et b' étant des constantes données

8.3. Propriétés de la variance

➤ X étant une variable aléatoire et a et b deux constantes, on a :

$$V(aX + b) = a^2V(X) \quad \text{et} \quad \sigma(aX + b) = |a|\sigma(X)$$

➤ La variance d'une somme de deux variables aléatoires est :

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

Remarques :

- 1) On retrouve la notion d'espérance mathématique en finance lorsqu'on veut déterminer le rendement espéré d'un titre ou encore le rendement espéré d'un portefeuille.
- 2) On retrouve la notion de variance en finance lorsqu'on veut évaluer le risque d'un titre.

- 3) On retrouve la notion de covariance en finance lorsqu'on veut évaluer le risque d'un portefeuille.
- 4) Si les variables X et Y sont indépendantes, alors $cov(X, Y) = 0$ et $\rho = 0$, mais réciproquement si $\rho = 0$ ou $cov(X, Y) = 0$, les variables X et Y ne sont pas obligatoirement indépendantes.

8.4. Cas particuliers des variables aléatoires indépendantes

Si X et Y sont indépendantes, on a $E(XY) = \bar{X}\bar{Y}$ donc $cov(X, Y) = 0$ et par suite :

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

Cette propriété se généralise à n variables aléatoires indépendantes sous la forme :

$$V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$$

NB : La variance d'une différence de deux variables aléatoires indépendantes est égale à la somme de leurs variances : $V(X - Y) = V(X) + V(Y)$.

2.3.5. Variable centrée réduite associée à X

Soit X une variable aléatoire d'espérance m et d'écart-type σ .

- La variable aléatoire X_c définie par : $X_c = X - m$ est appelée **variable centrée** associée à X .

Son espérance est nulle : $E(X_c) = E(X - m) = E(X) - m = m - m = 0$.

- La variable aléatoire X_R définie par : $X_R = \frac{X}{\sigma}$ est appelée **variable réduite** associée à X .

Son écart-type vaut 1 : $V(X_R) = V\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma^2}V(X) = \frac{1}{\sigma^2} \times \sigma^2 = 1$.

- La variable aléatoire Z définie par : $Z = \frac{X-m}{\sigma}$ est appelée **variable centrée réduite** associée à X .

Son espérance est nulle et son écart-type vaut 1.

Exemple :

Une petite agence loue des voitures à la journée. Elle dispose d'un parc de 6 véhicules, et la loi du nombre X de voitures louées par jour est donnée dans le tableau suivant :

x	0	1	2	3	4	5	6
$P(X = x)$	0.05	0.10	0.37	0.27	0.17	0.03	0.01

- Déterminer l'espérance, la variance et l'écart-type de X .
- Le bénéfice, en dollars, rapporté par la location des voitures est $B = 450X - 375$. Déterminer l'espérance, la variance et l'écart-type de B .
- Déterminer la variable centrée réduite Z associée à B puis donner une interprétation des probabilités $P(Z > 2)$ et $P(Z > 5)$.
- Exprimer en fonction de Z la probabilité que B ne s'écarte pas de son espérance de plus d'un demi écart-type.

Réponse :

- a) Déterminer $E(X)$, $V(X)$ et $\sigma(X)$

$$E(X) = (0 \times 0.05) + (1 \times 0.10) + (2 \times 0.37) + (3 \times 0.27) + (4 \times 0.17) + (5 \times 0.03) + (6 \times 0.01) = 2.54$$

L'espérance mathématique de X est de 2.54 voitures louées par jour.

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

Or, $E(X^2) = (0^2 \times 0.05) + (1^2 \times 0.10) + \dots + (6^2 \times 0.01) = 7.84$

D'où : $V(X) = 7.84 - (2.54)^2 = 1.3884$ et $\sigma(X) = \sqrt{1.3884} = 1.18$.

b) Déterminer $E(B)$, $V(B)$ et $\sigma(B)$

$$E(B) = E(450X - 375) = 450E(X) - 375 = 768$$

L'espérance mathématique de B est de 768 dollars par jour.

$$V(B) = V(450X - 375) = 450^2 V(X) = 281\,151$$

$$\sigma(B) = \sqrt{V(B)} = \sqrt{281\,151} = 530.24$$

c) Par définition $Z = \frac{B - E(B)}{\sigma(B)}$. On a donc

$$P(Z > 2) = P\left(\frac{B - E(B)}{\sigma(B)} > 2\right) = P(B > E(B) + 2\sigma(B))$$

$P(Z > 2)$ est donc la probabilité que le bénéfice B dépasse son espérance de plus de 2 fois son écart-type.

De même $P(Z > 5)$ est la probabilité que le bénéfice dépasse son espérance de plus de 5 fois son écart-type.

d) La probabilité que B ne s'écarte pas de son espérance de plus d'un demi écart-type s'écrit :

$$\begin{aligned} P\left(E(B) - \frac{1}{2}\sigma(B) < B < E(B) + \frac{1}{2}\sigma(B)\right) &= P\left(-\frac{1}{2}\sigma(B) < B - E(B) < \frac{1}{2}\sigma(B)\right) \\ &= P\left(-\frac{1}{2} < \frac{B - E(B)}{\sigma(B)} < \frac{1}{2}\right) = P\left(-\frac{1}{2} < Z < \frac{1}{2}\right) \end{aligned}$$

La variable aléatoire Z évalue la « distance » entre B et son espérance, mesurée en nombre d'écarts types.

2.4. Variables aléatoires continues

Si X est une variable aléatoire continue, on a $P(X = x) = 0$ pour tout x .

On ne peut donc pas définir la loi de probabilité d'une telle variable comme dans le cas discret. En revanche la fonction de répartition se définit de la même manière.

2.4.1. Fonction de répartition

1. Définition : L'application F qui, à tout réel x , associe la probabilité que la variable aléatoire X prenne une valeur inférieure ou égale à x , est par définition la fonction de répartition de X .
Pour tout x réel : $F(x) = P(X \leq x)$.

2. Propriétés

- F est une fonction croissante ;
- Pour tout x , on a : $0 \leq F(x) \leq 1$;
- $\lim_{x \rightarrow +\infty} F(x) = 1$ et $\lim_{x \rightarrow -\infty} F(x) = 0$
- Pour tous réels a et b : $P(a < X \leq b) = F(b) - F(a)$ et puisque $P(X = a) = P(X = b) = 0$, on a également :

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

Exemple : Dans une banque, entre 9 et 10 heures le matin, le temps d'attente de l'employé entre deux clients (en minutes) est une variable aléatoire X dont la fonction de répartition est la fonction F définie par :

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-0.1x} & \text{si } x \geq 0 \end{cases}$$

Un client vient juste de partir. Déterminer la probabilité que le client suivant se présente :

- Dans 5 minutes exactement ;
- Dans moins de 3 minutes ;
- Dans plus de 3 minutes mais moins de 10.

Réponse

a) $P(X = 5) = 0$ puisque X est une variable aléatoire continue.

b) $P(X < 3) = \underbrace{P(X \leq 3)}_{\text{puisque } P(X=3)=0} = F(3) = 1 - e^{-0.1 \times 3} \approx 0.259$.

La probabilité que le prochain client se présente dans moins de 3 minutes est donc égale à 0.259.

c) $P(3 < X < 10) = F(10) - F(3) = (1 - e^{-0.1 \times 10}) - (1 - e^{-0.1 \times 3}) \approx 0.373$

La probabilité que le prochain client se présente dans plus de 3 minutes mais de 10 est donc égale à 0.373.

2.4.2. Loi ou densité de probabilité d'une variable aléatoire continue

1. Définition : Si X est une variable aléatoire continue de fonction de répartition F dérivable, la loi de probabilité de X est définie par la dérivée de F , que nous noterons f .

f est appelée **densité de probabilité ou fonction de densité** de X . On note :

$$f(x) = F'(x)$$

2. Propriétés

2.1. La densité est une fonction positive

La fonction de répartition F est croissante, sa dérivée f est donc positive ou nulle. Graphiquement, cela signifie que la courbe représentative de f est entièrement située au-dessus de l'axe des abscisses.

2.2. Calcul de $P(a < X \leq b)$ à partir de la densité

On sait que $P(a < X \leq b) = F(b) - F(a)$, or F est une primitive de f , donc

$$F(b) - F(a) = \int_a^b f(t) dt \quad \text{et} \quad P(a < X \leq b) = \int_a^b f(t) dt$$

$$2.3. \quad \int_{-\infty}^{+\infty} f(t) dt = 1$$

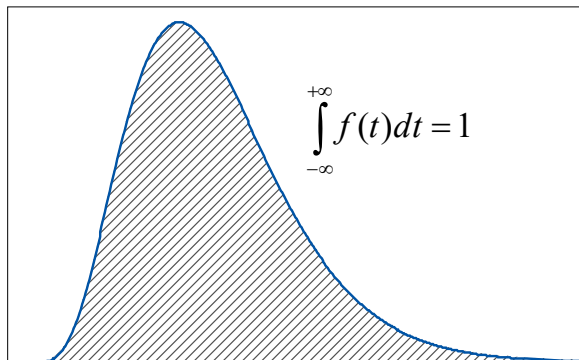
Graphiquement cela signifie que l'aire de la portion de plan située entre l'axe des abscisses et la courbe représentative de f vaut 1 (en unité d'aire).

2.4. Expression de F en fonction de f

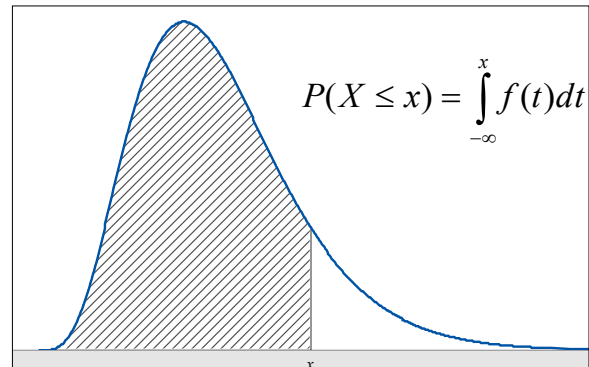
Par définition de la fonction de répartition F , on a : $F(x) = P(X \leq x)$, donc en prenant $a = -\infty$ et $b = x$ dans la propriété donnée au 2.2, on obtient :

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

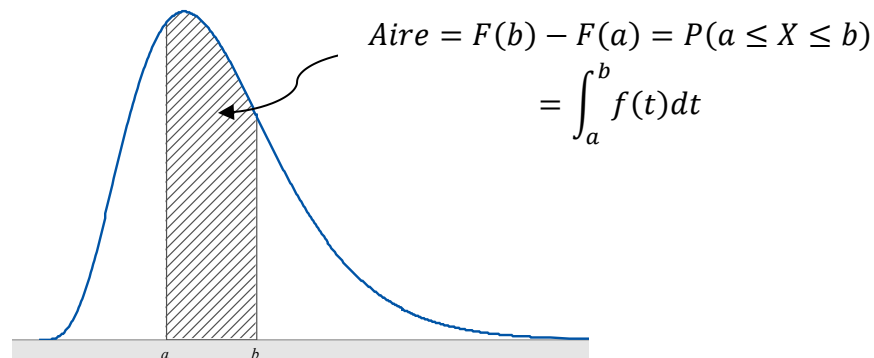
2.5. Représentation des propriétés



L'aire de la surface hachurée vaut 1



L'aire de la surface hachurée vaut $P(X \leq x)$



2.4.3. Variables aléatoires continues indépendantes

Définition : Deux variables aléatoires continues X et Y sont indépendantes si et seulement si pour tous réels x et y , les événements $(X \leq x)$ et $(Y \leq y)$ sont indépendants, ce qui signifie que l'on a :

$$P(X \leq x \text{ et } Y \leq y) = P(X \leq x) \times P(Y \leq y)$$

NB : On peut établir facilement que si X et Y sont indépendantes, on a également :

$$P(X > x \text{ et } Y > y) = P(X > x) \times P(Y > y)$$

2.4.4. Espérance - variance - écart-type d'une variable aléatoire continue

1. Définitions

Soit X une variable aléatoire continue de densité f , prenant dans l'intervalle I de bornes a et b (ces bornes ou l'une des deux bornes étant éventuellement infinies). Par définition :

➔ L'espérance de X , également notée \bar{X} , est : $E(X) = \int_a^b xf(x)dx$.

➔ La variance de X est : $V(X) = \int_a^b (x - \bar{X})^2 f(x)dx$.

Ou encore : $V(X) = \int_a^b x^2 f(x)dx - \bar{X}^2$ (formule de Koëinig)

➔ L'écart-type de X est : $\sigma(X) = \sqrt{V(X)}$

2. Propriétés

On retrouve les mêmes propriétés que dans le cas discret.

Les notions de variable centrée, variable réduite et variable centrée réduite se définissent comme dans le cas discret.

2.5. Valeurs caractéristiques

	Cas discret	Cas continu
Valeurs prises par X	$x_1, x_2, \dots, x_i, \dots$	Tous les réels compris entre a et b (avec éventuellement $a = -\infty, b = +\infty$)
Loi de X	Définie par la donnée de tous les $p_i = P(X = x_i)$ avec $p_i \geq 0$ et $\sum p_i = 1$.	Définie par la densité de probabilité f avec $\int_a^b f(x)dx = 1$ et $f(x) \geq 0, \forall x$.
Fonction de répartition	Pour tout x tel que $x_i < x \leq x_{i+1}$; $F(x) = P(X \leq x) = p_1 + \dots + p_i$	$F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt$
Espérance	$E(X) = \sum x_i \cdot p_i$	$E(X) = \int_{-\infty}^{+\infty} xf(x)dx$
Variance	$V(X) = \sum (x_i - E(X))^2 \cdot p_i$ $V(X) = \sum x_i^2 \cdot p_i - (E(X))^2$ $V(X) = E(X^2) - [E(X)]^2$	$V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 \cdot f(x)dx$ $V(X) = \int_{-\infty}^{+\infty} x^2 \cdot f(x)dx - (E(X))^2$ $V(X) = E(X^2) - (E(X))^2$
Écart-type	$\sqrt{V(X)}$	$\sqrt{V(X)}$
Covariance de X et Y	$cov(X, Y) = E(XY) - E(X) \times E(Y)$	

Chapitre 3 : Lois discrètes et Lois continues usuelles

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

- à la nature du phénomène étudié afin de choisir entre loi discrète et loi continue,
- à la forme de la distribution (histogramme),
- à la connaissance et à l'interprétation des principales caractéristiques de l'ensemble de données : espérance, médiane, variance, écart-type, coefficients d'asymétrie et de dissymétrie, etc.,
- au nombre de paramètres des lois, une loi dépendant de plusieurs paramètres peut s'adapter plus facilement à une distribution.

Domaine d'utilisation

- 1) Les lois discrètes sont utilisées pour modéliser les résultats des jeux de hasard, les sondages d'opinion, les phénomènes biologiques, les processus aléatoires (files d'attente, évolution de l'état de matériels) ... Les plus utilisées sont la loi uniforme, la loi binomiale et les lois dérivées, la loi hypergéométrique, la loi de Poisson.
- 2) Une variable aléatoire continue prend ses valeurs sur un ensemble infini non dénombrable de points, elle décrit par exemple la durée de vie d'une batterie de voiture, l'heure d'arrivée des voitures à un péage donné d'autoroute... La plupart des problèmes rencontrés en statistique peuvent se résoudre à l'aide de quelques lois fondamentales continues. Les principales sont : *loi uniforme, loi exponentielle, les lois gamma, les lois bêta, la loi normale et la loi log-normale* auxquelles il faut ajouter *les lois du chi-deux, de Fischer et de Student* utilisées dans la théorie de l'estimation.

Lois discrètes classiques

3.1.- Loi uniforme

Loi d'une variable aléatoire X prenant ses valeurs dans $\{1, \dots, n\}$ avec la même probabilité :

$$P(X = x) = \frac{1}{n} ; \forall x \in \{1, 2, \dots, n\}$$

Il est possible de vérifier que $\sum_{x=1}^n p_x = 1$.

Moments :

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

$$\text{Rappels : } \sum_{k=1}^n k = \frac{n(n+1)}{2} \text{ et } \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

Exemple : \mathcal{E} = « lancer d'un dé régulier ». X = Numéro apparaissant sur le dé. X suit une loi uniforme de probabilité $\frac{1}{6}$.

3.2.- Loi de Bernoulli $\mathcal{B}(1, p)$

On réalise une expérience aléatoire qui a deux résultats possibles : *soit le succès* qui a une probabilité p de se réaliser, *soit l'échec* qui a une probabilité $q = 1 - p$.

La variable aléatoire X = *nombre de succès obtenus* suit la *loi de Bernoulli*, notée $\mathcal{B}(1, p)$ et définie par :

$$P: \{0,1\} \rightarrow [0,1]$$

$$P(X = 0) = 1 - p \text{ et } P(X = 1) = p$$

Écriture sous forme synthétique :

$$P(X = x) = p^x(1 - p)^{1-x}, \text{ avec } x = 0 \text{ ou } x = 1$$

Loi de probabilité d'une variable de Bernoulli

x	$P(X = x)$
0	q
1	p
Total	1

Propriétés :

Si $X \sim \mathcal{B}(1, p)$, alors $E(X) = p$ et $V(X) = pq$

Exemples :

1. Dans une population un individu choisi au hasard possède un caractère A avec la probabilité p . La variable X définie sur $\Omega = \{A, \bar{A}\}$ par $X(A) = 1$ et $X(\bar{A}) = 0$ est une variable de Bernoulli de paramètre p .
2. On lance une pièce de monnaie une seule fois. Soit X la variable aléatoire qui caractérise le nombre de piles obtenues. X est une variable de Bernoulli, elle prend les valeurs entières 0 et 1 avec la probabilité constante 0,5.

Domaine d'utilisation : elle est utilisée pour modéliser des matériels qui seront soit survivants (valeur 1), soit défectueux (valeur 0) à un instant donné. Elle s'applique aux jeux de hasard de type binaire comme pile ou face...

3.3.- La loi Binomiale $\mathcal{B}(n, p)$ - Loi des tirages avec remise

On réalise n fois successivement et d'une manière indépendante une expérience aléatoire qui a deux résultats possibles, le *succès* (associé au résultat pour lequel nous voulons déterminer la

probabilité) qui a une probabilité p de se réaliser et l'échec qui a une probabilité $q = 1 - p$ de se réaliser.

La variable aléatoire $X = \text{nombre de succès obtenus au cours des } n \text{ épreuves}$ suit la loi binomiale définie par :

$$P: \{0, 1, 2, \dots, n\} \rightarrow [0, 1]$$

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

C'est la probabilité d'obtenir k succès et donc $(n - k)$ échecs au cours de n expériences aléatoires indépendantes.

Cette expression étant un terme du développement du binôme $[p + (1 - p)]^n$, la variable X est appelée *variable binomiale*. On vérifie facilement que : $\sum_{k=0}^n P(X = k) = 1$.

Remarque : La loi binomiale dépend de deux paramètres :

- n : nombre d'expériences aléatoires indépendantes ;
- p : probabilité de succès au cours de chacune des n expériences aléatoires, elle doit rester constante.

Une variable aléatoire X qui suit une loi binomiale de paramètres n et p , est désignée par :

$$X \sim \mathcal{B}(n, p)$$

Comment reconnaître une loi Binomiale ?

On a recours à une loi binomiale si :

- a) L'expérience se compose d'un nombre fixe n d'essais de Bernoulli ;
- b) Les deux résultats de chaque essai sont généralement complémentaires et appelés *succès (S)* et *échecs (E)* ;
- c) Le résultat de n'importe quel essai est indépendant du résultat de n'importe quel autre essai ;
- d) La probabilité p d'un succès reste la même d'un essai à l'autre

Quelques exemples d'application de la loi Binomiale

Exemple 1 :

- Lancement d'une pièce de monnaie (pile ou face) ;
- Qualité d'un produit (bon ou défectueux) ;
- Sondage électoral (pour ou contre) ; ...

Exemple 2 : Un vivier contient 100 truites de même aspect et de même âge. Cinq d'entre elles pèsent moins de 200 grammes, les autres plus de 200 grammes. Trois fois de suite, on va sortir une truite du vivier, la peser et la replonger rapidement dans le vivier. Soit X le nombre de truites de plus de 200 grammes obtenus à l'issue de ces trois expériences. Déterminer la loi de X .

Réponse :

Peser une truite constitue une **épreuve** ayant deux issues possibles :

- La truite pèse plus de 200 grammes (avec une probabilité $p = 0.95$) ;

➤ La truite pèse moins de 200 grammes (avec une probabilité $q = 1 - p = 0.05$)

Cette épreuve est **répétée** 3 fois, et les épreuves sont **indépendantes** puisqu'après chaque pesée, on doit remettre la truite dans le vivier.

X est nombre de truites de plus de 200 grammes obtenus au cours de ces 3 épreuves, donc **X suit une loi binomiale de paramètres $n = 3$ et $p = 0.95$.**

On note $X \sim \mathcal{B}(3, 0.95)$.

Remarque : Le tirage avec remise a deux conséquences :

- La proportion de truites de plus de 200 grammes contenus dans le vivier reste la même à chaque tirage (ce qui justifie l'indépendance des tirages) ;
- Une même truite peut être pesée plusieurs fois.

Exemple 3 : Grâce aux tarifs réduits de toutes sortes, 90% des voyageurs de la compagnie de transport Beurève bénéficient de tarifs réduits. Chaque soir, un contrôleur prend les billets de 5 passagers choisis au hasard dans 5 voitures différentes, et note le nombre X de billets à tarif réduit qu'il trouve parmi eux. Déterminer la loi de X , puis la probabilité qu'il trouve un seul billet à tarif réduit, son espérance et sa variance.

Réponse :

Regarder un billet constitue une épreuve ayant deux issues possibles :

- Tarif réduit (avec une probabilité $p = 0.9$) ;
- Tarif non réduit (avec une probabilité $q = 1 - p = 0.1$)

Cette épreuve est répétée 5 fois, et l'on peut supposer les épreuves indépendantes.

X , nombre de billet à tarif réduit obtenu au cours de ces 5 épreuves, **suit la loi binomiale de paramètres $n = 5$ et $p = 0.9$.**

- On a donc : $P(X = 1) = C_5^1 (0.9)^1 (1 - 0.9)^4 = 0.00045$.
- La formule $P(X = k) = C_5^k (0.9)^k (1 - 0.9)^{5-k}$ appliquée à tous les entiers entre 0 et 5 nous permet de donner le tableau de distribution suivant :

k	0	1	2	3	4	5
$P(X = k)$	0.00001	0.00045	0.00810	0.07290	0.32805	0.59049

- Espérance de X : $E(X) = \sum_{k=0}^5 kP(X = k) = 0 \times 0.00001 + \dots + 5 \times 0.59049 = 4.5$.

- Variance de X :

$$V(X) = \sum_{k=0}^5 k^2 P(X = k) - \bar{X}^2 = 0^2 \times 0.00001 + \dots + 5^2 \times 0.59049 - 4.5^2 = 0.45$$

L'espérance de X est de 4.5 billets, et sa variance est égale à 0.45.

Exemple 4 : Une employée de Télémarketing Inc appelle six ménages un jour donné pour les vendre un produit. D'après son expérience passée, elle réussit à vendre le produit à 15% des ménages qu'elle appelle. Calculer la distribution du nombre de ventes qu'elle fait ce jour-là.

Réponse :

Vérification des conditions d'application d'une expérience binomiale :

- Le nombre n d'essais est six, puisque l'employée fait six appels ce jour-là ;
- Pour chaque essai (appel), il n'y a que deux résultats possibles : *ou bien elle vend le produit (Succès) ou bien elle ne le vend pas (Échec)* ;
- La vente du produit à un ménage n'influe pas sur la vente du produit à un autre ménage, par conséquent, les essais (résultats des différents appels) sont indépendants ;
- La probabilité de 0.15 de vendre le produit est la même pour chaque appel.

Ainsi, l'expérience répond à toutes les conditions, on est donc en présence d'une *expérience aléatoire binomiale*.

Désignons le nombre de vente par la variable aléatoire discrète X . L'ensemble des valeurs possible de X est : $X = \{0, 1, 2, 3, 4, 5, 6\}$. La probabilité de réussite de chaque essai est de 0.15.

Par définition, on sait que :

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

où : n est le nombre d'essais ; k : est le nombre de succès ; p : est la probabilité de succès pour chaque essai.

On a donc : $P(X = k) = C_6^k p^k (1 - p)^{6-k}$. Alors :

- La probabilité qu'aucune vente ne soit effectuée ce jour-là est $P(X = 0)$:

$$P(X = 0) = C_6^0 (0.15)^0 (1 - 0.15)^{6-0} = 0.3771$$
- La probabilité que l'employée fasse exactement une vente est $P(X = 1)$:

$$P(X = 1) = C_6^1 (0.15)^1 (1 - 0.15)^{6-1} = 0.3993$$
- On continue ainsi jusqu'à $P(X = 6)$

On obtient la distribution suivante :

Nombre de ventes (k)	0	1	2	3	4	5	6
$P(X = k)$	0.3771	0.3993	0.1762	0.0415	0.0055	0.0004	0.000

Exercice : Une compagnie de téléphone de la région affirme que, dans 70% des cas, elle peut résoudre les problèmes de clients le jour même où ils sont signalés. Supposez que 15 problèmes signalés cette année par des clients soient sélectionnés au hasard.

- Combien de ces 15 problèmes attendriez-vous à voir résolus le jour où ils ont été signalés ?
- Quelle est la probabilité que 10 des problèmes aient été résolus le jour où ils ont été signalés ?

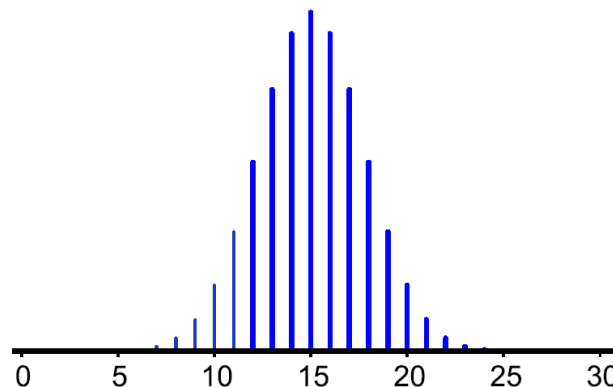
- c) Quelle est la probabilité que 10 ou 11 des problèmes aient été résolus le jour où ils ont été signalés ?
- d) Quelle est la probabilité que plus de 10 des problèmes aient été résolus le jour où ils ont été signalés ?

Propriétés de la loi binomiale

1) Représentation graphique

La loi binomiale étant une loi discrète, son graphe est un diagramme en bâtons. La hauteur des bâtons, proportionnelle à la quantité $P(X = k)$.

Exemples : Représentation graphique de la binomiale de paramètres 30 et 0.5



2) Table de la loi binomiale

Certaines tables donnent les probabilités individuelles, $P(X = k)$, et d'autres, les probabilités cumulées $\sum_{i=0}^k P(X = i)$, pour toutes les valeurs de i et k et pour des valeurs du nombre n d'épreuves et de la probabilité p .

3) Espérance - Variance - Écart-type

Si $X \sim \mathcal{B}(n, p)$, alors :

- Son espérance mathématique est : $E(X) = np$;
- Sa variance est : $V(X) = npq$
- Son écart-type est : $\sigma(X) = \sqrt{npq}$

4) Domaine d'utilisation

- La loi binomiale décrit des phénomènes ne pouvant prendre que deux états s'excluant mutuellement, succès ou échec dans un jeu, bonne pièce ou pièce défectueuse dans une fabrication, lot acceptable ou lot refusé, défaillance ou fonctionnement d'un matériel...
- Elle est utilisée dans le domaine technique pour déterminer la probabilité de défaillance à la sollicitation de matériels, en contrôle qualité, mais elle ne peut s'appliquer

rigoureusement que si les expériences sont non exhaustives, c'est la loi du *tirage avec remise*.

- Les événements considérés doivent être indépendants et la probabilité de réalisation d'un événement doit être constante.

5) Somme de variables aléatoires, binomiales, indépendantes et de même paramètre p

Si $X_1 \sim \mathcal{B}(n_1, p)$ et $X_2 \sim \mathcal{B}(n_2, p)$, et si ces deux variables aléatoires sont indépendantes, alors, $Y = X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$.

De façon générale, la somme de n variables aléatoires binomiales indépendantes et de même paramètre p est une variable aléatoire binomiale.

Remarque : Une variable binomiale est la somme de n variables de Bernoulli indépendantes, Si $X = X_1 + \dots + X_n$ telle que $X_i \sim \mathcal{B}(1, p)$ alors $X \sim \mathcal{B}(n, p)$.

3.2.1. - La distribution d'une proportion

Dans les enquêtes statistiques, en marketing, en contrôle de la qualité, ... l'intérêt porte plus sur la proportion de succès que sur le nombre de succès lui-même. On peut alors modifier la variable aléatoire X (qui représente le nombre de succès observés) pour obtenir un nombre relatif, une proportion de succès en n épreuves. On n'a qu'à diviser la variable aléatoire X par n et la nouvelle variable $\frac{X}{n}$ suivra aussi la loi binomiale car la variable est toujours X et de support $S_{\frac{X}{n}} = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\right\}$.

L'espérance mathématique et la variance de cette nouvelle variable.

$$\mu_{\frac{X}{n}} = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n} \times np = p \quad \text{et} \quad \sigma_{\frac{X}{n}}^2 = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{1}{n^2} \times npq = \frac{pq}{n}$$

Les principaux paramètres de la distribution binomiale lorsque les variables sont le nombre de succès X et la proportion de succès $\frac{X}{n}$ sont résumés dans le tableau suivant :

	Moyenne	Variance	Ecart-type
Nombre de succès X	np	$np(1-p)$	$\sqrt{np(1-p)}$
Proportion de succès $\frac{X}{n}$	p	$\frac{p(1-p)}{n}$	$\sqrt{\frac{p(1-p)}{n}}$

Remarque : La division des valeurs d'une variable aléatoire par une constante ne change pas la distribution de cette variable ; toutefois, elle en modifie les paramètres.

3.4. - La loi hypergéométrique $\mathcal{H}(N, n, p)$ ou Loi du tirage exhaustif

La loi hypergéométrique permet de modéliser de situations analogues à celles modélisées à l'aide de la loi binomiale. La différence réside le fait que les tirages sont sans remise (tirages exhaustifs).

Définition : On considère un tirage équiprobable sans remise ou tirage exhaustif dans une population d'effectif N , cette population étant composée de deux parties disjointes :

- Une partie A à N_1 éléments (éléments possédant certain caractère) ;
- Une partie B à $N_2 = N - N_1$ éléments (éléments n'ayant pas ce caractère).

C'est le cas, par exemple, d'un lot de N pièces comprenant N_1 pièces défectueuses et donc $N - N_1$ pièces fonctionnant bien.

Quelle est la probabilité qu'un sous-ensemble de n éléments contienne x éléments de l'ensemble A ?

Les éléments peuvent être tirés, soit un par un, soit d'un seul coup, mais sans remise.

Soit X la variable aléatoire représentant le nombre d'individus ayant le caractère considéré dans l'échantillon. On veut calculer $P(X = x)$.

Le nombre d'échantillons de taille n est égal à : C_N^n .

Un échantillon de taille n comprend :

- x individus, pris parmi les N_1 individus ayant le caractère considéré, donc $C_{N_1}^x$ choix possibles,
- $n - x$ individus pris parmi les $N - N_1$ individus n'ayant pas ce caractère, donc $C_{N-N_1}^{n-x}$ choix possibles.

Le nombre d'échantillons de taille n , comprenant x individus pris parmi les N_1 individus, est donc égal à : $C_{N_1}^x \times C_{N-N_1}^{n-x}$.

La probabilité cherchée est égale à : $P(X = x) = \frac{C_{N_1}^x \times C_{N-N_1}^{n-x}}{C_N^n}$.

Les valeurs extrêmes de x sont : $\min x = \max\{0, n - N(1 - p)\}$ et $\max x = \min\{n, N_1\}$. Le quotient $\frac{n}{N}$ est appelé *taux de sondage*.

La variable aléatoire ainsi définie suit *une loi hypergéométrique $\mathcal{H}(N, n, p)$* . Cette loi dépend de trois paramètres N , n et p .

Caractéristiques d'une variable hypergéométrique

Les distributions hypergéométriques possèdent des propriétés semblables à celles des distributions binomiales.

- La proportion des individus de la population qui possèdent le caractère étudié est : $p = \frac{N_1}{N}$
- La proportion des individus de la population ne possédant le caractère étudié est : $q = \frac{N_2}{N}$
- Espérance mathématique de X est : $E(X) = np$.

- Variance et écart-type de X : $V(X) = \frac{N-n}{N-1}npq$ et $\sigma(X) = \sqrt{\frac{N-n}{N-1}}\sqrt{npq}$. $\frac{N-n}{N-1}$ est le *facteur d'exhaustivité*.

Remarque : Le facteur d'exhaustivité tend vers 1 si N est grand devant n et la variance de la variable X tend alors vers la variance d'une variable suivant la loi binomiale $\mathcal{B}(n, p)$.

Domaine d'utilisation

La loi hypergéométrique est utilisée, en particulier dans les contrôles de qualité où on retire, de la population étudiée, les éléments défectueux.

Exemple : Dans une assemblée de 30 personnes, il y a 20 hommes et 10 femmes. On tire un échantillon de 15 personnes (tirage sans remise).

Soit X la variable aléatoire « *nombre d'hommes* » dans cet échantillon.

La probabilité d'avoir 10 hommes dans un échantillon de taille 15 est égale à :

$$P(X = 10) = \frac{C_{20}^{10} \times C_{10}^5}{C_{30}^{15}} = 0.30$$

Le modèle de cette loi correspond très souvent à des sondages d'opinion. Par exemple, X est le nombre de personnes qui expriment une intention de vote favorable à un candidat dans un échantillon de 1000 personnes. Si la même personne ne peut être interrogée plusieurs fois, c'est un tirage sans remise et la loi exacte est la loi hypergéométrique, la loi binomiale constitue simplement une excellente approximation.

3.5.- La loi de Poisson $\mathcal{P}(\lambda)$

C'est une classe importante de variables aléatoires discrètes, nommée ainsi en mémoire du mathématicien Siméon Denis Poisson.

Contrairement aux lois binomiales, la loi de Poisson comporte un nombre infini de valeurs. Elle convient à la description d'évènements dont les chances de réalisation sont faibles, c'est-à-dire d'évènements *rare* (*succès*), qui se produisent durant un intervalle précis. L'intervalle peut être *le temps, une distance, un endroit précis ou un volume*.

Les journalistes, et un peu tout le monde, la connaissent bien sans le savoir. La constatation :
Il y a en moyenne 5 accidents par jour à ce rond-point
 signifie que la variable aléatoire comptant le nombre d'accidents pendant une journée au fameux rond-point, suit la loi de Poisson de paramètre 5.

Exemple :

Nombre d'occurrences d'un évènement dans un certain laps de temps ou dans une région donnée, comme :

- Nombre de clients arrivant à une banque ;
- Nombre d'imperfections d'un véhicule automobile fraîchement peint ;
- Nombre d'accidents par semaine sur une autoroute ;
- Nombre d'appels téléphoniques dans un intervalle de temps ;
- Nombre de naissances par année dans une petite municipalité.

Pour que la distribution de Poisson en constitue une bonne approximation, ces situations doivent respecter les critères suivants, appelés caractéristiques de l'expérience de Poisson :

- 1) La valeur de la variable aléatoire est le nombre de succès enregistrés durant un intervalle ;
- 2) Le nombre de succès enregistrés durant un intervalle donné ne dépend pas du nombre de succès dans les autres intervalles ;
- 3) À mesure que l'intervalle se réduit, la probabilité que deux ou plusieurs succès soient enregistrés dans l'intervalle s'approche de zéro. On connaît aussi cette caractéristique sous le nom de propriété d'un évènement rare ;
- 4) La probabilité qu'un succès ait lieu dans un intervalle est la même pour tous les intervalles de longueur égale, et le nombre moyen de succès durant un intervalle est proportionnel à la longueur de l'intervalle.

Définition : La probabilité d'observer exactement k occurrences d'un certain évènement dans une unité de temps ou de région si $X \sim \mathcal{P}(\lambda)$, est donnée par :

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} ; k = 0, 1, 2, 3, \dots$$

où

$P(X = k)$ est la probabilité que la variable X prenne la valeur k ;

λ est le nombre moyen d'occurrences (succès) dans un intervalle donnée.

Il est possible de vérifier que $\sum_{k=0}^{+\infty} p_k = 1$.

Outils : On sait que $\sum_{x=0}^{+\infty} \frac{\lambda^x}{x!} = e^\lambda$

A.- Espérance - Variance - Écart-type

Si X suit la loi de Poisson $\mathcal{P}(\lambda)$, alors :

- Son espérance mathématique est : $E(X) = \lambda$,
- Sa variance est : $V(X) = \lambda$,
- Son écart-type est : $\sigma(X) = \sqrt{\lambda}$

NB : La loi de Poisson est la seule loi discrète possédant la propriété, $E(X) = V(X)$.

B.- Propriété

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

- 1) Si $X_1 \sim \mathcal{P}(\lambda_1)$ et $X_2 \sim \mathcal{P}(\lambda_2)$, X_1, X_2 indépendantes, alors $Y = X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$
- 2) **Généralisation :** Si $Z = X_1 + X_2 + \dots + X_n$ avec $X_i \sim \mathcal{P}(\lambda_i)$ pour tout $i = 1, \dots, n$; alors :

$$Z = X_1 + X_2 + \dots + X_n \sim \mathcal{P}(\lambda_1 + \lambda_2 + \dots + \lambda_n)$$

D.- Principaux cas où la loi de Poisson s'applique

La loi de Poisson est la loi discrète d'une variable aléatoire représentant un nombre d'événements. Elle est utilisée pour décrire :

- la réalisation d'événements peu probables, dans une succession d'épreuves très nombreuses, au moins 50,
- le nombre d'accidents dans un atelier, le nombre de défauts sur un appareil,

Elle a des applications dans le domaine des files d'attente. On retiendra les principaux exemples suivants :

- nombre d'erreurs dans une comptabilité, une page ... ;
- nombre de pannes d'un appareil au cours d'une période donnée ;
- nombre d'appels téléphoniques reçus au cours d'une période donnée ;
- nombre de personnes arrivant à un guichet au cours d'une période donnée (arrivées dites poissonnières)

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

Exemple : Selon les données recueillies depuis plusieurs années, le nombre de pannes hebdomadaires du système informatique d'une entreprise suit une loi de Poisson de paramètre $\lambda = 0.05$.

Soit X la variable aléatoire « nombre de pannes hebdomadaires », on a donc :

$$P(X = k) = \frac{e^{-0.05} (0.05)^k}{k!}$$

La probabilité que le système tombe en panne une fois au cours d'une quelconque ($k = 1$) est égale à : 0.04756.

La probabilité qu'il fonctionne sans panne ($k = 0$) est égale à 0.95122.

On considère une année (50 semaines) de fonctionnement de ce système. Le nombre de pannes Y obéit à une loi de Poisson de paramètre $\mu = 0.05 \times 50 = 2.5$.

$$P(Y = k) = \frac{e^{-2.5} (2.5)^k}{k!}$$

La probabilité d'observer 2 pannes au cours de l'année ($k = 2$) est égale à 0.2565 et la probabilité d'en observer 4 est égale à 0.1336.

3.6.- Loi Géométrique $\mathcal{G}(p)$

Modèle aléatoire de cette loi

L'expérience aléatoire consiste à effectuer des tirages avec remise dans une urne contenant une proportion p de boules blanches et $1 - p$ de boules noires. Les tirages s'arrêtent dès qu'une boule blanche est obtenue.

La variable aléatoire X géométrique de paramètre p associée à cette expérience aléatoire est le nombre de tirages nécessaires pour obtenir la première fois une boule blanche.

Définitions : La variable aléatoire X est un nombre entier de support $S_X = \{1, 2, \dots, k + 1, \dots, n, \dots\}$

$$P(X = k) = p_k = (1 - p)^{k-1} p = q^{k-1} p; \forall x \in \mathbb{N}^*$$

Il est possible de vérifier que $\sum_{k=1}^{+\infty} p_k = 1$.

$$\text{Outils : } \sum_{x=0}^{+\infty} p^x = \frac{1}{1-p}$$

Propriétés :

- 1. Espérance et variance de X :** $X \sim \mathcal{G}(p) \Leftrightarrow E(X) = \frac{1}{p}; V(X) = \frac{q}{p^2}$.
- 2. Domaine d'application :** Cette loi sert généralement lorsque nous intéressons au temps d'attente du premier succès, c'est-à-dire au nombre d'essais nécessaires pour obtenir un succès, lors d'une succession d'expériences indépendantes n'ayant que deux issues possibles : le succès avec une probabilité p et l'échec avec une probabilité $1 - p$. La loi est parfois utilisée pour modéliser des durées de vie. Par exemple, la loi géométrique est le modèle discret de la mort d'une particule radioactive. La loi géométrique est la version discrète d'une loi absolument continue : la loi exponentielle (*voir plus loin*).
- 3. Cette loi est un cas particulier de la loi de Pascal ;**

3.6.- Approximation des lois (Lois limites)

La loi hypergéométrique dépend de trois paramètres, la loi binomiale de deux et la loi de Poisson d'un seul, les calculs pour la première loi sont donc plus longs que pour la deuxième et que pour la troisième. Il est intéressant de chercher dans quelles conditions, on peut utiliser la loi binomiale comme approximation de la loi hypergéométrique puis la loi de Poisson comme approximation de la loi binomiale.

Cependant, comme les domaines de définition et d'utilisation de ces trois lois sont différents, il faut définir avec précision dans quelles conditions ces approximations seront valables.

3.5.1.- Approximation d'une loi hypergéométrique par une loi binomiale

Si l'échantillon est constitué lors d'un sondage d'opinion avec de tirages sans remise, dès que la taille n de l'échantillon est petite vis-à-vis de la taille N de la population, la loi hypergéométrique peut être approximée par une loi binomiale de paramètres n et p .

Si $\frac{n}{N} \leq 0.1$, l'approximation est très correcte.

Autrement dit, quand un prélèvement se fait sans remise, mais qu'il ne dépasse pas 10% de la population, on peut l'assimiler à un tirage avec remise.

3.5.2.- Approximation d'une loi binomiale par une loi de Poisson

Si les trois conditions suivantes sont vérifiées :

- n grand (par exemple n supérieur ou égal à 30),
- p petit (par exemple p inférieur ou égal à 0.1),
- np pas trop grand (par exemple np inférieur ou égal à 10),

On peut approcher la loi binomiale $\mathcal{B}(n, p)$ par la loi de Poisson $\mathcal{P}(\lambda)$, avec $\lambda = np$, ce qui permet décrire pour k entier inférieur ou égal à n :

$$P(X = x) = C_n^k p^k (1 - p)^{n-k} \rightarrow \frac{e^{-np} (np)^x}{x!}$$

Convergence de la loi binomiale vers la loi de Poisson

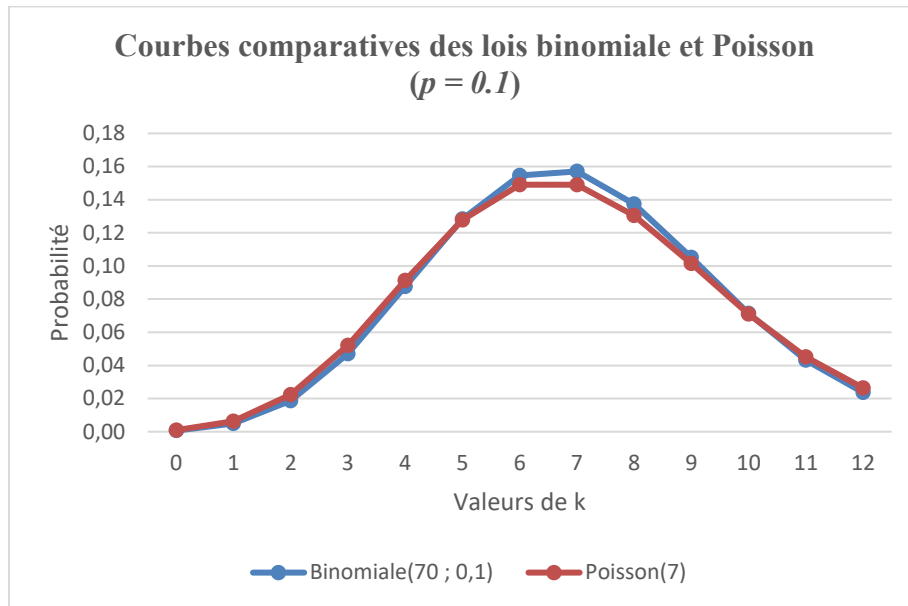
Soit $X \sim \mathcal{B}(n, p)$ alors, si n grand et p petit, on peut approximer la loi binomiale par une loi de Poisson $\mathcal{P}(\lambda)$, avec $\lambda = np$ (il s'agit d'une convergence en loi).

En pratique : $n > 50, p < 0.1$.

Exemple : Voici, pour $0 \leq k \leq 10$, les valeurs de $P(X = k)$ lorsque X suit la loi binomiale $\mathcal{B}(70 ; 0.1)$ et celles de $P(X = k)$ lorsque X suit la loi de Poisson de même espérance égale à 7. L'approximation est correcte.

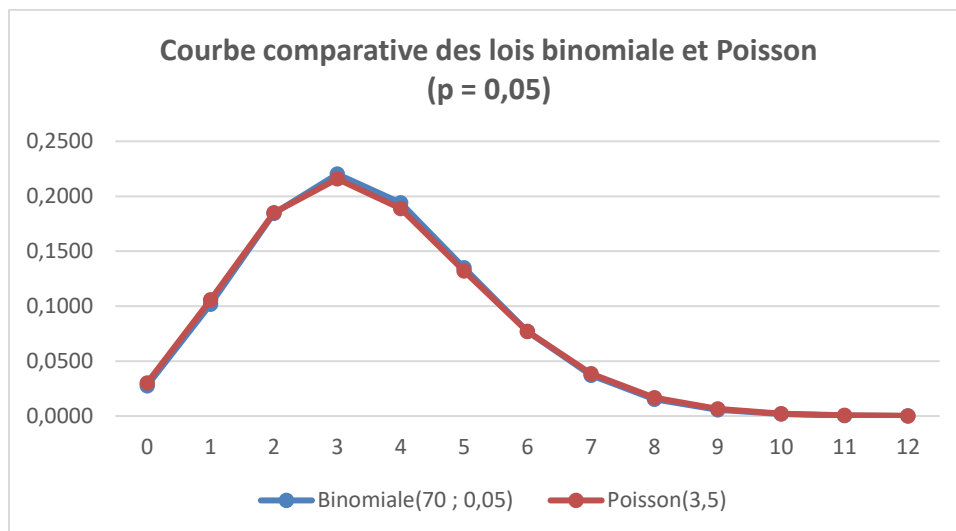
k	0	1	2	3	4	5	6	7	8	9	10	11	12
Binomiale(70 ; 0,1)	0,0006	0,0049	0,0187	0,0470	0,0876	0,1284	0,1546	0,1570	0,1374	0,1052	0,0713	0,0432	0,0236
Poisson(7)	0,0009	0,0064	0,0223	0,0521	0,0912	0,1277	0,1490	0,1490	0,1304	0,1014	0,0710	0,0452	0,0263

Graphiquement, on a :



On constate que l'approximation est beaucoup plus parfaite pour $p = 0,05$.

k	0	1	2	3	4	5	6	7	8	9	10	11	12
Binomiale(70 ; 0,05)	0,0276	0,1016	0,1845	0,2201	0,1941	0,1348	0,0769	0,0370	0,0153	0,0056	0,0018	0,0005	0,0001
Poisson(3,5)	0,0302	0,1057	0,1850	0,2158	0,1888	0,1322	0,0771	0,0385	0,0169	0,0066	0,0023	0,0007	0,0002



En résumé nous avons les tableaux récapitulatifs suivants :

A. Approximation

Loi réelle	Loi approchée	Conditions d'approximation
Hypergéométrique $\mathcal{H}(N, n, p)$	Binomiale $\mathcal{B}(n, p)$	Si $n \leq 0,1N$
Binomiale $\mathcal{B}(n, p)$	De Poisson $\mathcal{P}(np)$	$\begin{cases} n \geq 30 & p \leq 0,1 \\ np \leq 10 \end{cases}$

B. Caractéristiques des principales lois discrètes

Loi	Formule	Application	Espérance	Variance
Hypergéométrique $\mathcal{H}(N, n, p)$	$P(X = x) = \frac{C_{N_1}^x \times C_{N-N_1}^{n-x}}{C_N^n}$	Tirage sans remise	np	$npq \frac{N-n}{N-1}$
Binomiale $\mathcal{B}(n, p)$	$P(X = x) = C_n^x p^x (1-p)^{n-x}$	Tirage avec remise Probabilité constante	np	npq
Poisson $\mathcal{P}(\lambda)$	$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$	Événements instantanés et indépendants	λ	λ

Exemples :**I.- Approximation de la loi binomiale par une loi de Poisson**

D'année en année, en moyenne 4% des étudiants d'un cours de statistique obtiennent plus de 80% à l'examen final. Si on tire un échantillon de taille $n = 100$ parmi les dossiers des anciens, quelle est la probabilité d'y trouver exactement 2 où le résultat est supérieur à 80% ?

Réponse :

L'expérience aléatoire consiste à tirer un échantillon de taille $n = 100$ parmi les dossiers des anciens étudiants et de trouver le nombre d'étudiants ayant obtenu plus de 80% à l'examen final.

Cette expérience suit donc une loi binomiale de paramètre $n = 100$ et $p = 0.04$. La probabilité d'avoir k succès est : $P(X = k) = C_{100}^k p^k (1-p)^{100-k}$

Probabilité de trouver exactement 2 dossiers où le résultat est supérieur à 80% :

$$P(X = 2) = C_{100}^2 (0.04)^2 (1 - 0.04)^{98} = 0.1449$$

Puisque $p < 0.10$ et $np = 4$, alors on peut utiliser la loi de Poisson pour approximer cette probabilité. On a de façon approximative $X \sim \mathcal{P}(\lambda = 100 \times 0.04 = 4)$. Donc ;

$$P(X = 2) = \frac{e^{-4} \times 4^2}{2!} = 0.1465$$

II.- Approximation de la loi hypergéométrique par une loi binomiale

Voir le TD1

Utilisation des logiciels commerciaux

En plus des connaissances théoriques, il faut savoir calculer rapidement $P(X = k)$ ou $P(X \leq k)$ pour une variable aléatoire quelconque. Les logiciels commerciaux contiennent généralement des fonctions intégrées pour les lois classiques.

A.- Loi binomiale $\mathcal{B}(n, p)$ **1. Calculer la valeur de $P(X = k)$**

Logiciels	Formule
Mupad	binomialPF(n,p)(k)
Xcas	binomial(n,k,p)

Excel	loi.binomiale.n(k;n;p;FAUX)
-------	-----------------------------

2. Calculer $P(X \leq k)$.

Le nombre k est nécessairement un entier

Logiciels	Formule
Mupad	binomialCDF(n,p)(k)
Xcas	Binomial cdf(n,p,k)
Excel	loi.binomiale.n(k;n;p;VRAI)

Exemples : Pour $n = 30$ lancers, $p = 0.5$ et $k = 20$, on obtient :

- $P(X = 20) = 0,027981601$
- $P(X \leq 20) = 0,978613027$

3. Calculer le plus petit entier k tel que $P(X \leq k) \geq x$

Les commandes nous donnent l'unique entier k tel que $P(X \leq k - 1) < x$ et $P(X \leq k) \geq x$.

Logiciels	Formule
Mupad	binomialQuantile(n,p)(x)
Xcas	binomial icdf(n,p,x)
Excel	loi.binomiale.inverse(n;p;x)

Exemple : Pour $n = 30$ lancers, $p = 0.5$ et $x = 0.975$, on obtient 20.

On vérifie que $P(X \leq 19) = 0.9506$ est strictement plus petit que 0.975 et $P(X \leq 20) = 0.9786$ est supérieur ou égal à 0.975.

B.- Loi de Poisson $\mathcal{P}(\lambda)$

1. Calculer la valeur de $P(X = k)$

Logiciels	Formule
Mupad	poissonPF(λ)(k)
Xcas	poisson(λ ,k)
Excel	loi.poisson.n(k; λ ;FAUX)

2. Calculer $P(X \leq k)$.

Le nombre k n'est pas nécessairement un entier

Logiciels	Formule
Mupad	poissonCDF(λ)(k)
Xcas	poisson cdf(λ ,k)
Excel	loi.poisson.n(k; λ ;VRAI)

Exemples : Pour $\lambda = 7$ et $k = 5$, on obtient : $P(X \leq 5) = 0.3007$

3. Calculer le plus petit entier k tel que $P(X \leq k) \geq x$

Les commandes nous donnent l'unique entier k tel que $P(X \leq k - 1) < x$ et $P(X \leq k) \geq x$.

Logiciels	Formule
Mupad	poissonQuantile(λ)(x)
Xcas	poisson icdf(λ ,x)

C.- Loi hypergéométrique $\mathcal{H}(N, n, p)$

On a un sac de N boules, dont M boules rouges. On tire sans remise n boules du sac et, X désigne le nombre de boules rouges tirées.

1) Calculer la valeur $P(X = k)$

Mupad	hypergeometricPF(N,M,n)(k)
Excel	Loi.hypergeometrique.N(k ; n ; M ; N ; FAUX)

Explications de la Formule Excel :

Loi.hypergeometrique.N (nombre de succès dans l'échantillon(k) ; taille de l'échantillon(n) ;
 nombre de succès dans la population(M) ; Taille de la population(N) ;
 FAUX(on ne souhaite pas de cumul))

2) Calculer $P(X \leq k)$

Mupad	hypergeometricCDF(N,M,n)(k)
Excel	Loi.hypergeometrique.N(k ; n ; M ; N ; VRAI)

3) Calculer le plus petit entier k tel que $P(X \leq k) \geq x$

Les commandes nous donnent l'unique entier k tel que $P(X \leq k - 1) < x$ et $P(X \leq k) \geq x$.

Mupad	hypergeometricQuantile(N,M,n)(k)
-------	----------------------------------

3.7.- Lois continues usuelles

3.7.1- Loi exponentielle

Cette loi a beaucoup d'application en recherche opérationnelle, entre autres en fiabilité où l'on mesure l'aptitude d'une pièce à fonctionner sans panne, une fois mise en service. On l'utilisera aussi dans les phénomènes d'attente où la durée de certains services est exponentielle.

➔ Loi $X \sim \mathcal{E}(\lambda)$, $\lambda > 0 \Leftrightarrow f(x) = \lambda e^{-\lambda x} 1_{x \geq 0} = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \geq 0 \end{cases}$

➔ Il est possible de vérifier que : $\int_{-\infty}^{+\infty} f(x) dx = 1$

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

➔ Fonction de répartition : $F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$

➔ **Domaine d'application** : Loi souvent utilisée en fiabilité : durée de vie d'un composant

➔ **Exemple** :

3.7.2- Loi gamma

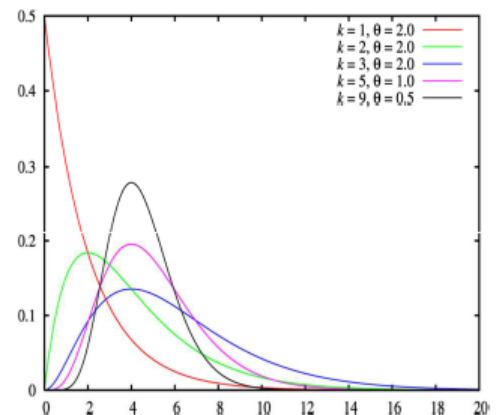
➔ Loi $X \sim \Gamma(k, \theta)$; $k > 0, \theta > 0 \Leftrightarrow f(x) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\Gamma(k) \theta^k} 1_{x \geq 0}$.

➔ Il est possible de vérifier que : $\int_{-\infty}^{+\infty} f(x) dx = 1$

➔ Moments : $E(X) = k\theta$; $V(X) = k\theta^2$.

➔ Propriétés :

- Si $k = 1$, $X \sim \mathcal{E}\left(\frac{1}{\theta}\right)$;
- Si $\theta = 2$, $X \sim \chi^2(2k)$;
- Si k entier, la loi de X s'appelle loi d'Erlang ;
- Si X_i i.i.d $\sim \mathcal{E}\left(\frac{1}{\theta}\right) \Rightarrow Y = \sum_{i=1}^n X_i \sim \Gamma(n, \theta)$.



Outils : $\Gamma(k) = \int_0^{+\infty} t^{k-1} e^{-t} dt$; $\Gamma(k+1) = k\Gamma(k)$

3.7.3.- La loi normale (Laplace-Gauss) : $\mathcal{N}(\mu, \sigma)$

C'est la plus importante des lois de probabilité continues. Des questions tant théoriques que pratiques font appel à cette loi (souvent loi limite). Elle est également très utilisée dans le domaine

de l'Inférence Statistique. Dans de telles applications, la loi normale fournit une description des résultats possibles obtenus grâce à un échantillon.

Historiquement, elle apparaît vers 1773 comme la forme limite de la loi binomiale (Abraham de Moivre). Gauss en 1809 et Laplace en 1812 lui donnèrent sa forme définitive.

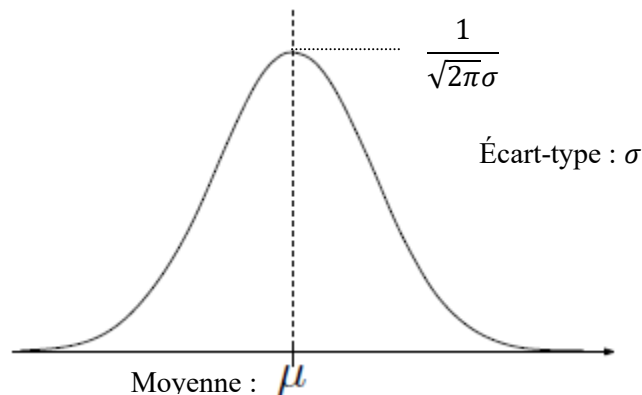
Définition – Fonction de densité : Une variable aléatoire X suit une loi de Laplace-Gauss de paramètres μ et σ si sa fonction densité est :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma \in \mathbb{R}_+^*$$

On note : $X \sim \mathcal{N}(\mu, \sigma)$ où : μ correspond à la moyenne et σ correspond à l'écart type.

A.- La courbe normale

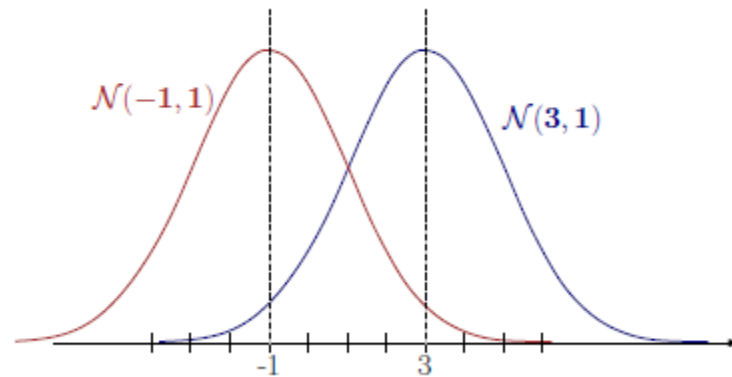
La loi normale est représentée par une courbe en forme de cloche. Voici une courbe en forme de cloche de la loi normale de moyenne μ et d'écart-type σ .



N.B : La courbe normale a deux paramètres, μ et σ . Ils déterminent la position et la forme de la distribution.

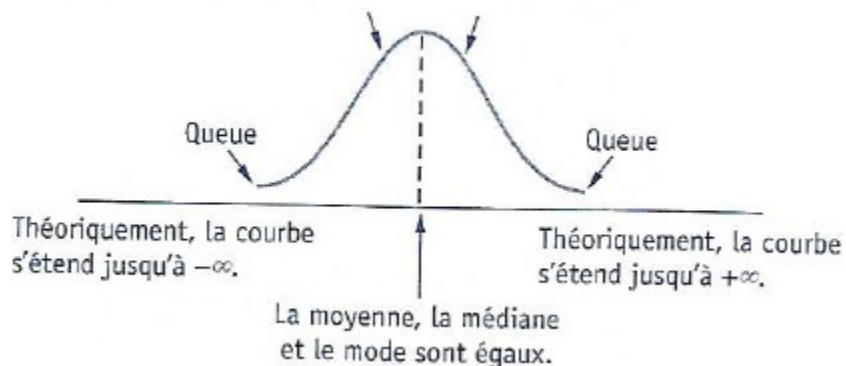
A.1.- Les caractéristiques des lois normales

- Il existe une famille entière de lois normales. Elles se différencient par leur moyenne μ et leur écart-type σ ;
- Le point le plus élevé de la courbe normale correspond à la moyenne qui est également la médiane et le mode de la distribution ;
- La moyenne de la distribution peut être négative, nulle ou positive. Deux courbes normales ayant le même écart-type mais deux moyennes différentes sont représentées ci-dessous :



- la courbe normale est en forme de cloche : elle a un sommet unique au centre et décroît à mesure que l'on s'éloigne de la valeur centrale dans l'une ou dans l'autre des directions. Elle est asymptotique, c'est-à-dire qu'elle s'approche de plus en plus de l'axe horizontal sans jamais réellement le toucher. Autrement dit, les « queues » de la courbe s'étendent indéfiniment dans les deux directions et ne touchent jamais l'axe horizontal ;
- la courbe est symétrique autour de la valeur centrale. Si l'on coupe la courbe verticalement à sa valeur centrale, les deux moitiés sont le miroir l'une de l'autre. La valeur centrale est donc la moyenne arithmétique μ , et elle est également la médiane et le mode de la distribution. L'aire sous la courbe à droite de la moyenne représente la moitié, et l'aire sous la courbe à gauche de la moyenne représente l'autre moitié ;

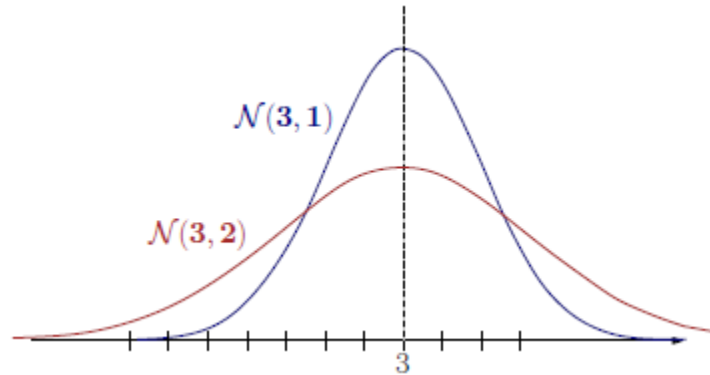
**La courbe normale est symétrique.
Les deux moitiés sont identiques**



- les probabilités d'une variable aléatoire normale sont données par l'aire sous la courbe. L'aire totale située sous la courbe d'une distribution de probabilité est égale à 1. Puisque la distribution est symétrique, l'aire sous la courbe à gauche de la moyenne est égale à 0.5 et l'aire sous la courbe à droite à 0.5 également. On note :

$$P(X < \mu) = P(X > \mu) = 0.5$$

- l'écart-type détermine la largeur de la courbe. Plus l'écart-type est grand, plus la courbe sera large, aplatie, traduisant ainsi une plus grande dispersion des données. Deux distributions normales de même moyenne mais avec des écarts types différents sont représentés ci-dessous :



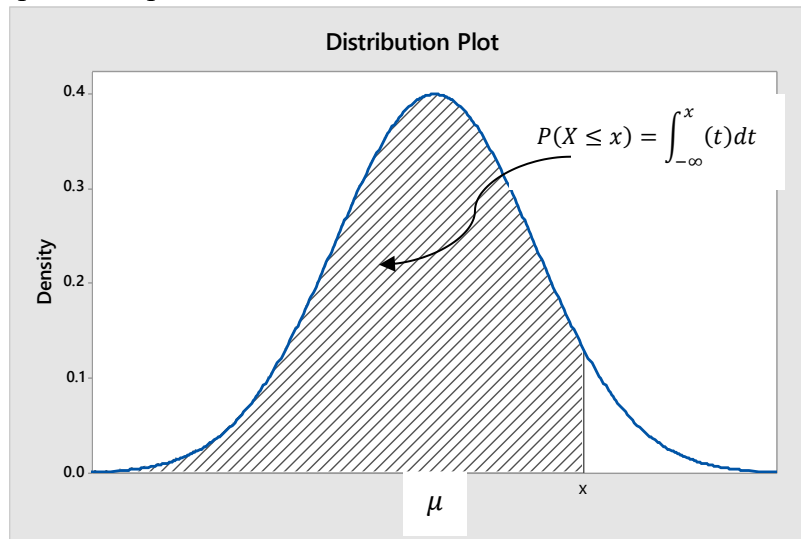
- pour toute variable aléatoire normale X de moyenne μ et d'écart-type σ , la variable centrée réduite $\frac{X-\mu}{\sigma}$ obéit à une normale centrée réduite de moyenne nulle et d'écart-type égal à 1 ;

A.2.- Fonction de répartition :

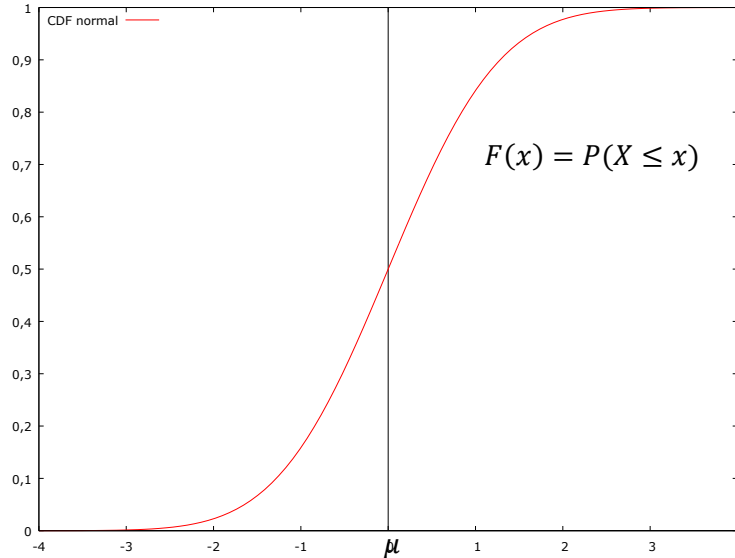
La fonction de répartition de cette loi est :

$$F(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

Cette valeur est représentée par la surface hachurée sous la courbe normale.



La forme de la fonction de répartition est alors :



A.3.- Propriétés : Moyenne et variance de la loi normale

Si $X \sim \mathcal{N}(\mu, \sigma)$ alors $E(X) = \mu$ et $V(X) = \sigma^2$.

Ces deux paramètres μ et σ^2 caractérisent complètement la loi normale.

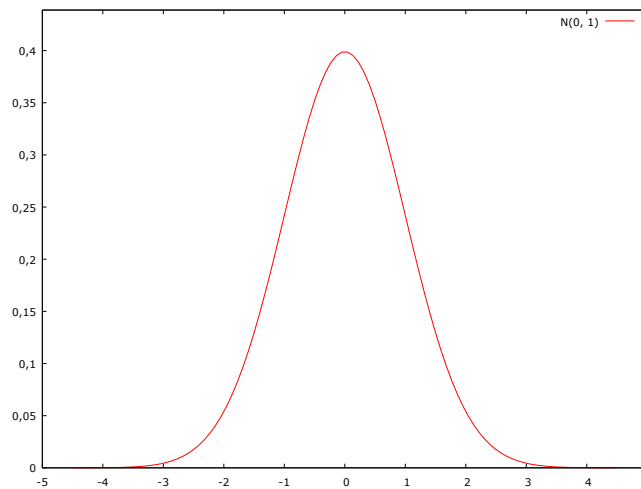
B.- Loi normale centrée réduite ou Loi normale standard

Soit $X \sim \mathcal{N}(\mu, \sigma)$, alors $Z = \frac{X - \mu}{\sigma}$ suit la loi normale centrée réduite et sa densité en tout point réel z est la suivante :

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad \text{et} \quad X = \sigma Z + \mu$$

f est une fonction positive sur \mathbb{R} et on admettra que $\int_{-\infty}^{+\infty} f(z) dz = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 1$.

B.1.- Représentation graphique de la densité



f est une fonction paire, sa courbe représentative est **symétrique** par rapport à l'axe des ordonnées.

Pour tout z positif, on peut lire la valeur de $f(z)$ dans une table (donnée en annexe), et pour tout z négatif, on a $f(z) = f(-z)$.

B.2.- Espérance - Variance et Écart-type

Si Z suit la loi normale centrée réduite, son espérance est $E(Z) = 0$, sa variance $V(Z) = 1$ et son écart-type $\sigma(X) = 1$.

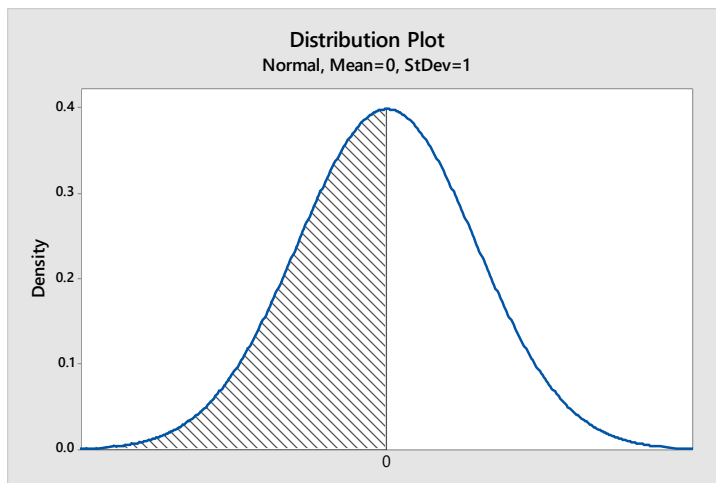
On note $Z \sim \mathcal{N}(0,1)$, le premier paramètre étant l'espérance et le second l'écart-type.

B.3.- Fonction de répartition

Si Z est une variable aléatoire continue suivant la loi normale centrée réduite, sa fonction de répartition, souvent notée Π , est définie par :

$$\Pi(z) = P(Z \leq z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \text{ pour tout } t \text{ réel.}$$

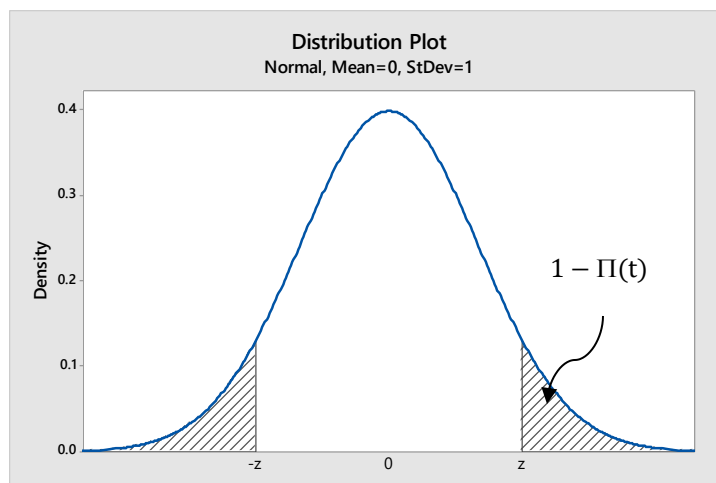
B.4.- Propriétés



a) $\Pi(0) = P(Z \leq 0) = 0.5$

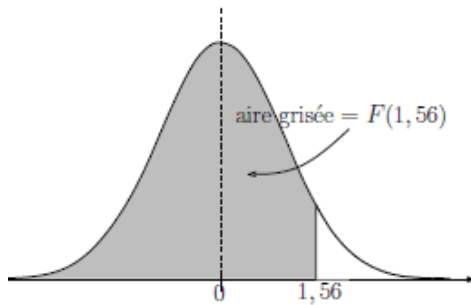
On sait que l'aire de la portion de plan située entre l'axe des abscisses et la courbe de la densité de Z vaut 1. Compte tenu de la symétrie de cette courbe par rapport à l'axe des ordonnées, $\Pi(0) = P(Z \leq 0) = 0.5$.

b) $\Pi(-z) = 1 - \Pi(z)$ pour tout z .



Cette propriété, illustrée ci-contre, résulte également de la symétrie de la courbe de la densité de Z par rapport à l'axe des ordonnées.

c) La fonction Π est strictement croissante.



On cherche **1.56** dans la table :

	...	0,06	...
...			
1,5	...	0.9406	...
...			

Donc $P(X \leq 1.56) = 0.9406$.

La variable X est inférieure à 1.56 pour 94.06% des individus.

b) On cherche $P(X \geq 1.49)$.

On écrit d'abord $P(X \geq 1.49) = 1 - P(X \leq 1.49) = 1 - F(1.49)$

	0,09
...			
1,4	0.9319
...			

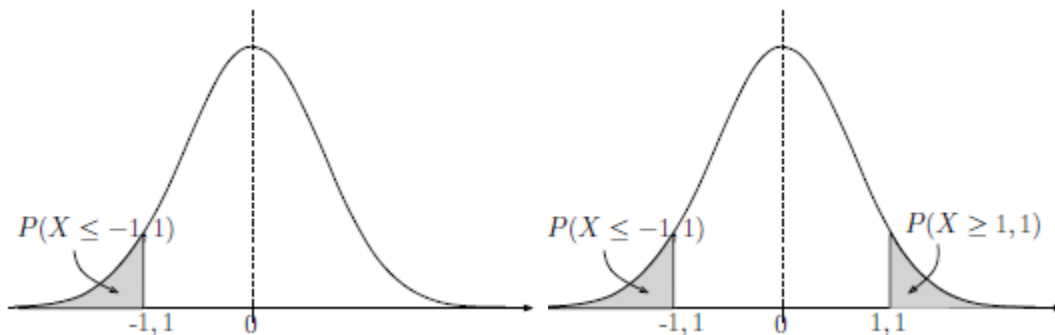
On cherche **1.49** dans la table :

Donc $P(X \leq 1.49) = 0.9319$.

Finalemnt $P(X \geq 1.49) = 1 - 0.9319 = 0.0681$.

La variable X est supérieure à 1.49 pour 6.81% des individus.

c) On cherche $P(X \leq -1.1)$, c'est-à-dire $F(-1.1)$.



Par symétrie, on sait que $P(X \leq -1.1) = P(X \geq 1.1)$. On a donc :

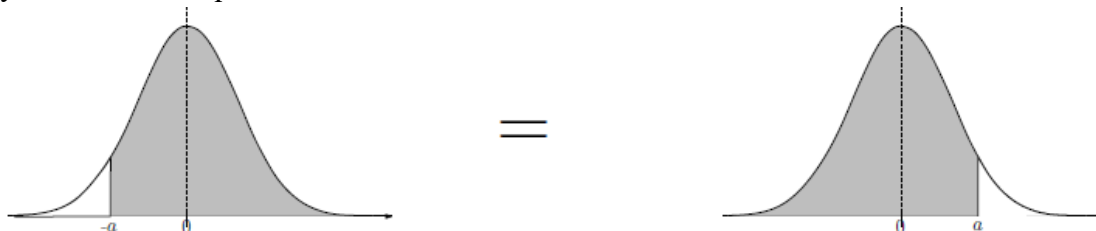
$$P(X \leq -1.1) = P(X \geq 1.1) = 1 - P(X \leq 1.1) = 1 - 0.8643 = 0.1357$$

Finalemnt $P(X \leq -1.1) = 0.1357$.

Donc, pour 13.57% des individus, la variable X est inférieure à -1.1 .

d) On cherche $P(X \geq -2.5)$

Par symétrie, on sait que $P(X \geq -2.5) = P(X \leq 2.5)$



On cherche **2.50** dans la table. Donc $P(X \geq -2.5) = P(X \leq 2.5) = 0.9938$

C.2.- Calcul avec la loi normale quelconque, tous les cas

Pour faire des calculs avec une $\mathcal{N}(\mu, \sigma)$, on se ramène à la loi $\mathcal{N}(0,1)$

Théorème : si $X \sim \mathcal{N}(\mu, \sigma)$ alors $\frac{X-\mu}{\sigma} = Z \sim \mathcal{N}(0,1)$.

On dit que l'on centre et réduit X . Dans ce cours, on utilise la lettre Z pour désigner une loi normale centrée réduite.

Exemple : On suppose qu'une certaine variable $X \sim \mathcal{N}(11,2)$. Pour quelle proportion d'individus est-ce que $X \leq 14$?

Réponse

On cherche $P(X \leq 14)$

- On centre et on réduit X : $\frac{X-11}{2} = Z \sim \mathcal{N}(0,1)$.
- On a donc :

$$P(X \leq 14) = P\left(\frac{X-11}{2} \leq \frac{14-11}{2}\right) = P(Z \leq 1.5)$$

- On cherche 1.5 dans la table.

On trouve finalement $P(X \leq 14) = 0.9332$

C.3.- Quantile >50% d'une $\mathcal{N}(0,1)$.

Exemple : On cherche le quantile à 97.5% pour la $\mathcal{N}(0,1)$. Cela revient à trouver a tel que $P(Z \leq a) = 0.975$

On lit la table dans le sens inverse :

	...	0,06	...
⋮			
1,9	...	0.9750	...
⋮			

Donc $P(X \leq 1.96) = 0.9750$

Le quantile recherché est donc 1.96.

Notation : Le quantile d'ordre α pour la loi normale centrée réduite est noté z_α .

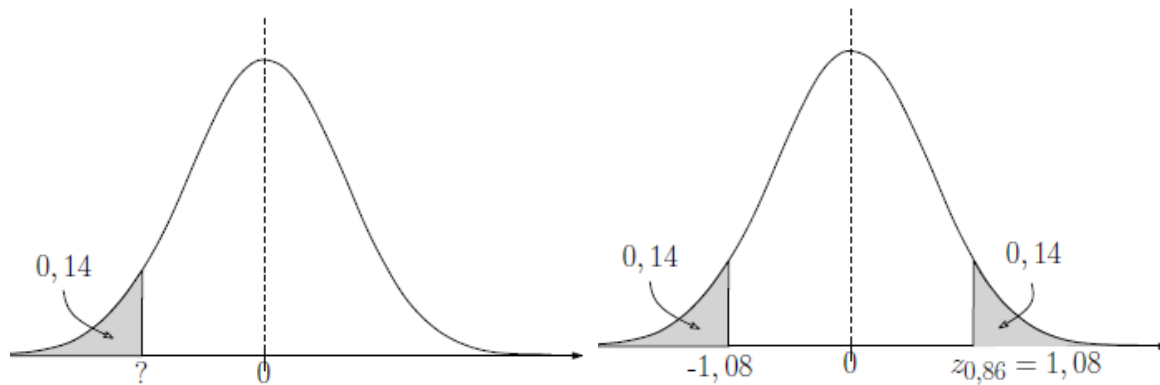
Par exemple, $z_{0.975} = 1.96$.

C.4.- Quantile <50% d'une $\mathcal{N}(0,1)$.

Exemple : On cherche le quantile à 14% pour la $\mathcal{N}(0,1)$.

Réponse

Cela revient à trouver a tel que $P(Z \leq a) = 0.14$. Il n'y a pas de nombre inférieur à 0.5 dans la table.

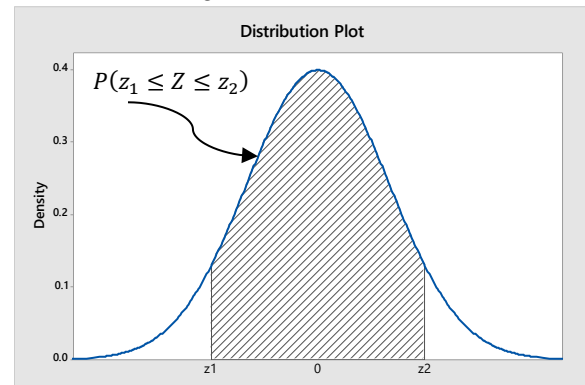
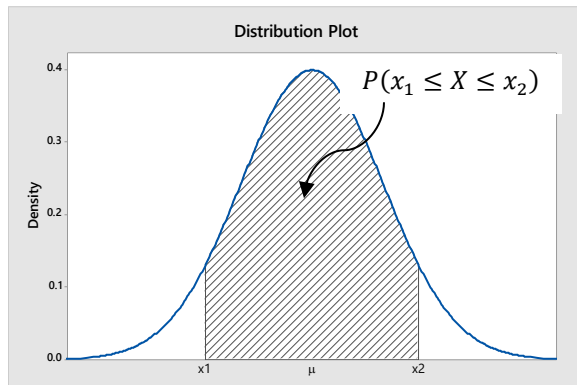


Le quantile est donc $z_{0,14} = -1,08$.

C.5.- Calculs pratiques

a) Si $X \sim \mathcal{N}(\mu, \sigma)$, alors :

$$\begin{aligned} P(x_1 < X < x_2) &= P\left(\frac{x_1 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{x_2 - \mu}{\sigma}\right) = P\left(\frac{x_1 - \mu}{\sigma} < Z < \frac{x_2 - \mu}{\sigma}\right) \\ &= P(z_1 < Z < z_2) \\ &\text{avec } z_1 = \frac{x_1 - \mu}{\sigma} \quad \text{et} \quad z_2 = \frac{x_2 - \mu}{\sigma} \end{aligned}$$



De symétrique par rapport à $X = \mu$, la courbe devient symétrique par rapport à $Z = 0$.

b) $Z \sim \mathcal{N}(0,1)$, si $P(Z < a)$, $a > 0$ est connue, alors :

- $P(X < -a) = P(X > a) = 1 - P(X < a) = 1 - P(Z < a)$;
- $P(-a < Z < a) = P(Z < a) - P(Z < -a) = P(Z < a) - (1 - P(Z < a)) = 2P(Z < a) - 1$
- Application numérique : $a = 1,87$
 - 1) $P(Z < 1,87) = 0,9693$
 - 2) $P(Z < -1,87) = 1 - 0,9693 = 0,0307$.
 - 3) $P(-1,87 < X < 1,87) = 0,9693 - 0,0307 = 0,9386$

D.- Probabilités de quelques intervalles centrés sur l'espérance

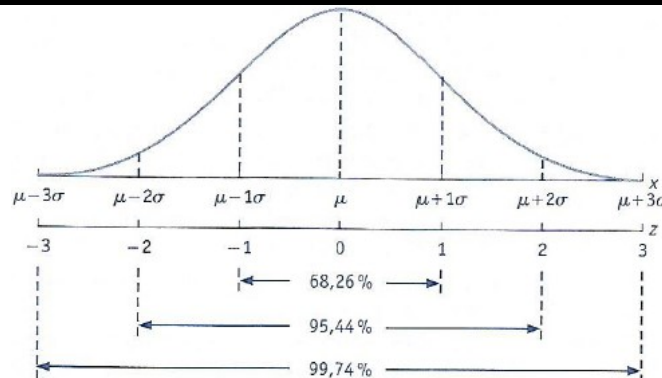
Si $X \sim \mathcal{N}(\mu, \sigma)$, il y a environ :

- 68 chances sur 100 que X ne s'écarte pas de son espérance de plus d'un écart-type ;
- 95 chances sur 100 que X ne s'écarte pas de son espérance de plus de 2 écarts types ;
- Plus de 99 chances sur 100 que X ne s'écarte pas de son espérance de plus de 3 écarts types.

En règle générale :

- 68.26% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu - \sigma ; \mu + \sigma]$;
- 95.44% des valeurs d'une variable aléatoire normale sont dans l'intervalle $[\mu - 2\sigma ; \mu + 2\sigma]$;
- 99.74% des valeurs d'une variable aléatoire normale sont dans l'intervalle $[\mu - 3\sigma ; \mu + 3\sigma]$.

N.B : Ces pourcentages sont à la base de la règle empirique. La figure ci-dessous illustre graphiquement les propriétés énoncées ci-dessus :

**Nous avons aussi :**

- 90% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu - 1.64\sigma ; \mu + 1.64\sigma]$
- 95% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu - 1.96\sigma ; \mu + 1.96\sigma]$
- 99.8% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu - 3.09\sigma ; \mu + 3.09\sigma]$

E. - Propriétés**1) Fonction affine d'une variable normale**

a et b étant deux réels donnés, et X une variable aléatoire qui suit la loi normale de paramètres μ et σ , la variable $Y = aX + b$ suit la loi normale d'espérance $a\mu + b$ et d'écart-type $|a|\sigma$.

2) Somme et différence de variables aléatoires normales indépendantes)

Soient $X_1 \sim \mathcal{N}(\mu_1, \sigma_1)$ et $X_2 \sim \mathcal{N}(\mu_2, \sigma_2)$ indépendantes, alors ;

$$X_1 + X_2 \sim \mathcal{N}\left(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right) \quad \text{et} \quad X_1 - X_2 \sim \mathcal{N}\left(\mu_1 - \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right)$$

Cas particulier

La somme de deux variables aléatoires gaussiennes, centrées, réduites est une variable gaussienne centrée non réduite :

$$\left. \begin{array}{l} \text{Loi de la variable X : } \mathcal{N}(0;1) \\ \text{Loi de la variable Y : } \mathcal{N}(0;1) \\ \text{X et Y variables indépendantes} \end{array} \right\} \text{Loi de la somme } S = X + Y : \mathcal{N}\left(0; \sqrt{2}\right)$$

Généralisation :

a) Si $X_i \sim \mathcal{N}(\mu_i, \sigma_i)$, $i = 1, \dots, n$ indépendantes, alors :

$$\sum_{i=1}^n X_i \sim \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sqrt{\sum_{i=1}^n \sigma_i^2}\right)$$

b) Si $X_i \sim \mathcal{N}(\mu, \sigma)$, $i = 1, \dots, n$ indépendantes, alors :

$$\frac{1}{n}(X_1 + X_2 + \dots + X_n) \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

F. - Théorème de limite centrale

Soient X_1, \dots, X_n , n variables aléatoires indépendantes, de même loi, de même espérance μ et de même écart-type σ .

Posons $S_n = X_1 + \dots + X_n$. Alors : $E(S_n) = n\mu$, $V(S_n) = n\sigma^2$ et

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \rightarrow \mathcal{N}(0,1) \text{ en loi quand } n \rightarrow \infty.$$

$$S_n \sim \mathcal{N}(n\mu, \sigma\sqrt{n}) \text{ quand } n \rightarrow \infty.$$

Exemple : Convergence de la loi binomiale (somme de n lois de Bernoulli) vers la loi normale.

G. - Approximation par une loi normale

Cette section est consacrée à l'étude des situations permettant d'approcher une loi binomiale, puis une loi de Poisson, par une loi normale.

G.1. - Approximation d'une loi binomiale par une loi normale

Conditions d'approximation : Si la variable aléatoire discrète X suit la loi binomiale $\mathcal{B}(n, p)$, et si l'une des conditions suivantes est réalisée :

[n « grand » ($n \geq 30$) et p et q voisins de 0.5] ou [$n \geq 30$, np et $nq > 18$]. On peut approcher la loi X par la loi normale de paramètres $\mu = np$ et $\sigma = \sqrt{npq}$.

Convergence de la loi binomiale vers la loi normale

Soit $X \sim \mathcal{B}(n, p)$, alors $\frac{X - np}{\sqrt{npq}} \rightarrow \mathcal{N}(0,1)$ en loi quand $n \rightarrow \infty$.

ou bien $\mathcal{B}(n, p) \approx \mathcal{N}(np, \sqrt{npq})$ ($n \rightarrow \infty$).

Ceci signifie que lorsque n est assez grand, on peut approximer la loi binomiale par la loi normale ; en pratique $p \in [0.1 ; 0.9]$, $n \geq 30$.

Dans certains ouvrages, on trouve la condition : $np(1 - p) > 9$ ou $np, nq \geq 5$.

G.2. - Approximation d'une loi de Poisson par une loi normale

Conditions d'approximation : Si la variable aléatoire discrète X suit la loi de Poisson $\mathcal{P}(\lambda)$, et si λ est « grand » ($\lambda > 10$), on peut approcher la loi de X par la loi normale de paramètres λ et $\sqrt{\lambda}$.

Convergence de la loi de Poisson vers la loi normale

Soit $X \sim \mathcal{P}(\lambda)$ alors si $\lambda \rightarrow \infty$, $\frac{X - \lambda}{\sqrt{\lambda}} \rightarrow \mathcal{N}(0, 1)$ en loi. L'approximation est très satisfaisante pour $\lambda > 18$.

H. - Correction de continuité

Lorsqu'on effectue une approximation d'une loi discrète (loi binomiale, loi Poisson ; par exemple) par une loi continue (loi normale), il faut donc faire une correction de continuité pour que l'approximation soit de bonne qualité.

Sans une telle correction, on obtiendrait toujours 0 comme valeur approximative de $P(X = a)$, alors que ce n'est généralement pas le cas.

Règles régissant de la correction de continuité dans l'évaluation d'une probabilité (a et b sont des entiers)

Énoncé avant la correction $X \sim \mathcal{B}(n, p)$	Énoncé après la correction $Y \sim \mathcal{N}(np, \sqrt{npq})$
$P(X = a)$	$P(a - 0.5 < Y < a + 0.5)$
$P(a < X < b)$	$P(a + 0.5 < Y < b - 0.5)$
$P(a \leq X \leq b)$	$P(a - 0.5 < Y < b + 0.5)$

Remarque :

Soit X une variable aléatoire suivant une loi binomiale de paramètres n et p . Alors la variable aléatoire $Z = \frac{X - np}{\sqrt{np(1-p)}}$ tend vers la loi $\mathcal{N}(0 ; 1)$ quand n tend vers l'infini. On obtient donc :

$$P(X = k) \cong P\left(\frac{k - 0.5 - np}{\sqrt{np(1-p)}} < Z < \frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

et de la même façon :

$$P(X \leq k) \cong P\left(Z < \frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

Exemple : Si $X \sim \mathcal{B}(50, 0.3)$, calculer $P(17 < X \leq 20)$.

Les conditions d'approximation de la loi binomiale par la loi normale sont remplies ($n = 50 \geq 30$; $np = 50 \times 0.3 = 15 \geq 5$ et $nq = 35 \geq 5$). Par conséquent :

$P(17 < X \leq 20) = P(17 + 0.5 < Y < 20 + 0.5)$ où $Y \sim \mathcal{N}(15, \sqrt{10.5})$. D'où :

$$P(17.5 < Y < 20.5) = P\left(\frac{17.5 - 15}{\sqrt{10.5}} < \frac{Y - 15}{\sqrt{10.5}} < \frac{20.5 - 15}{\sqrt{10.5}}\right) = P(0.77 < Z < 1.70) = 0.1760$$

Ainsi, $P(17 < X \leq 20) \approx 0.1760$

Facteur de correction de continuité

Valeur 0.5 soustraite d'une valeur sélectionnée ou ajoutée à cette dernière, selon la question, lorsqu'une distribution discrète est approximée par une distribution continue.

Illustration des cas d'exceptions pour calculer la probabilité que le nombre de succès soit entre x_1 et x_2 à l'aide d'exemples où $n = 20$.

Évènement	x_1	x_2	Aire sous la courbe normale
Exactement 9 succès	9	9	Entre 8.5 et 9.5
6 succès ou moins	0	6	A gauche de 6.5
Moins de 6 succès, c'est-à-dire $\{0, 1, \dots, 5\}$	0	5	A gauche de 5.5
Au moins 17 succès, c'est-à-dire $\{17, 18, 19, 20\}$	17	20	A droite de 16.5
Plus de 6 succès, mais pas plus de 9	7	9	Entre 6.5 et 9.5
Au moins 8 succès, mais pas plus de 10	8	10	Entre 7.5 et 10.5

Exemple : On lance n fois une pièce équilibrée. En prenant les valeurs successives $n = 40, 80, 100, 200$, calculer la probabilité d'obtenir au moins 60% de faces, à savoir les probabilités de sortir au moins 24, 48, 60, 120 faces suivant les cas. On fera le calcul exact avec la loi binomiale $\mathcal{B}\left(n, \frac{1}{2}\right)$ et avec son approximation normale, avec ou sans correction de continuité.

Réponse

Pour les calculs avec la loi binomiale exacte, on peut utiliser la valeur formelle :

$$P(X \geq k) = \frac{1}{2^n} \sum_{i=k}^n C_n^i$$

Voici les valeurs calculées, directement avec la loi binomiale, avec l'approximation normale sans correction de continuité, avec l'approximation normale avec correction de continuité. Par exemple, pour $n = 100$, on calcule $P(Z > 60)$ sans la correction de continuité et $P(Z > 59.5)$ avec la correction de continuité.

n	Exacte	Sans correction	Avec correction
20	0.25172	0.18554	0.25116
40	0.13409	0.10295	0.13419
80	0.04645	0.03681	0.04676
100	0.02844	0.02275	0.02871
200	0.00284	0.00233	0.00291

Remarque :

- 1) On constate bien que l'approximation avec correction de continuité est meilleure que celle faite sans cette correction ;
- 2) La première ligne pour $n = 20$ est un peu limite pour les conditions d'application de l'approximation, 20 étant un peu faible. L'approximation sans correction de continuité est franchement mauvaise, mais celle avec cette correction est plutôt bonne quand même ;
- 3) Si n est grand, il est clair que la correction de continuité ne va pas avoir une grande influence puisque les probabilités ponctuelles $P(X = a)$ et $P(X = b)$ seront très faibles.

3.7.3.- Lois dérivées de la loi normale

Certaines lois sont définies à partir de la loi normale et elles sont utilisées pour des applications statistiques bien précises. Ce sont :

A.- La loi du Khi-deux à n degrés de liberté : χ_n^2

Définition et Notation : Soient X_1, \dots, X_n , n variables aléatoires indépendantes telles que $X_i \sim \mathcal{N}(0, 1)$, $\forall i \in \{1, 2, \dots, n\}$. Si $X = X_1^2 + X_2^2 + \dots + X_n^2$, alors X est variable aléatoire continue soumise à une fonction de densité de probabilité dite du « **Khi-Carré** » ou « **Khi-deux** » à n degrés de liberté, et l'on note :

$$X \sim \chi_n^2$$

n , le nombre de degrés de liberté de la distribution, est le paramètre spécifique de celle-ci.

Plus généralement, si X_1, \dots, X_n sont indépendantes et suivent la loi normale $\mathcal{N}(\mu, \sigma^2)$ alors la

variable aléatoire : $X = \left(\frac{X_1 - \mu}{\sigma}\right)^2 + \dots + \left(\frac{X_n - \mu}{\sigma}\right)^2$ suit cette loi du χ^2 à n degrés de liberté.

Espérance et Variance

Si $X \sim \chi_n^2$, alors les valeurs respectives de l'espérance et de la variance de cette variable sont :

$$E(X) = n \text{ et } V(X) = 2n$$

En effet ;

- La linéarité de l'espérance nous dit que :

$$E(X) = E(X_1^2) + \dots + E(X_n^2) = V(X_1) + \dots + V(X_n) = \underbrace{1 + \dots + 1}_{n \text{ termes}} = n$$

Puisque les variables X_i centrées vérifient : $V(X_i) = E(X_i - E(X_i))^2 = E(X_i^2) = 1$.

- L'indépendance des variables X_i^2 nous dit que : $V(X) = V(X_1^2) + \dots + V(X_n^2) = nV(X_1^2)$.

Calculer donc la variance de Y^2 , lorsque $Y \sim \mathcal{N}(0, 1)$.

Ce calcul se fait en utilisant une intégration :

$$V(Y^2) = E(Y^4) - E(Y^2)^2 = E(Y^4) - V(Y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^4 e^{-\frac{x^2}{2}} dx - 1$$

On fait une intégration par partie, en posant :

$$u = x^3 \rightarrow u' = 3x^2 \quad \text{et} \quad v' = x e^{-\frac{x^2}{2}} \rightarrow v = -e^{-\frac{x^2}{2}}$$

Donc ;

$$V(Y^2) = \frac{1}{\sqrt{2\pi}} \left(\left[-x^3 e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} 3x^2 e^{-\frac{x^2}{2}} dx \right) - 1 = \frac{3}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx - 1$$

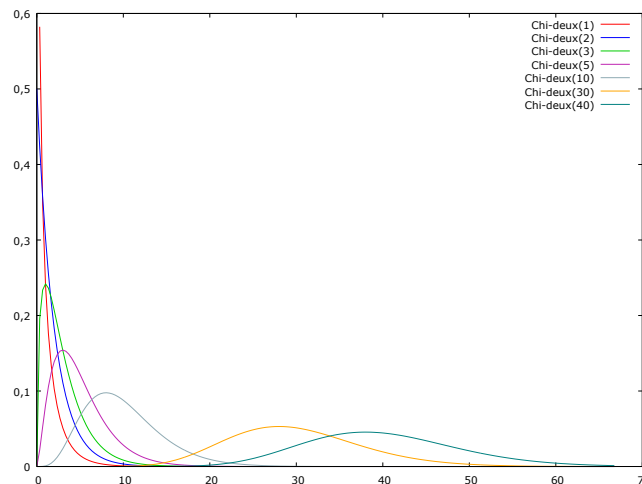
$$V(Y^2) = 3E(Y^2) - 1 = 3V(Y) - 1 = 2$$

Car, par croissance comparée, $\left[-x^3 e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} = 0$.

On a donc : $V(X) = 2n$.

Représentation graphique

Voici, pour $n = 1 ; n = 2 ; n = 3 ; n = 5 ; n = 10 ; n = 30$ et $n = 40$ la représentation graphique de la fonction de densité de probabilité d'une variable X soumise à une loi χ_n^2 .



Densité de la loi du chi-deux pour différentes valeurs du paramètre n .

Remarque :

- 1) Toute variable aléatoire soumise à une loi χ_n^2 ne peut prendre que des valeurs positives de 0 à $+\infty$ (étant définie comme somme de carrés de variables) ;
- 2) La loi du Khi-Deux joue un rôle important dans les intervalles de confiance et les tests d'hypothèses sur la variance ;

- 3) L'aspect de la courbe d'une distribution χ_n^2 variera selon le nombre de degrés de liberté n qui est le seul paramètre de cette distribution ;
- 4) Lorsque n est grand, elle se rapproche de la courbe d'une loi $\mathcal{N}(n, 2n)$.

Convergence de la loi χ_n^2 vers la loi normale (approximation)

Soit $X \sim \chi_n^2$ alors $\frac{X - n}{\sqrt{2n}} \rightarrow \mathcal{N}(0, 1)$ en loi quand $n \rightarrow \infty$, ou bien on écrit : $\chi_n^2 \approx \mathcal{N}(n, \sqrt{2n})$

($n \rightarrow \infty$).

En pratique $n > 30$.

Additivité (variables aléatoires indépendantes)

Soient $X_1 \sim \chi_{n_1}^2, \dots, X_k \sim \chi_{n_k}^2$ indépendantes. Alors $Z = X_1 + \dots + X_k \sim \chi_n^2$ avec $n = n_1 + \dots + n_k$.

B. - La loi de Student à n degrés de libertés : T_n

Définition et Notation : Soient $X \sim \mathcal{N}(0, 1)$ et $Y \sim \chi_n^2$ deux variables aléatoires indépendantes,

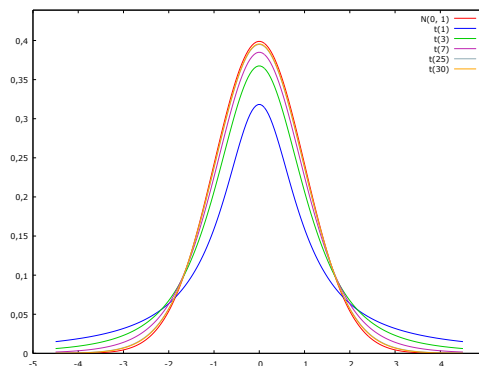
si $T = \frac{X}{\sqrt{\frac{Y}{n}}}$, alors T est une variable aléatoire continue soumise à une fonction de densité de

probabilité dite « T de Student » à n degrés de liberté et l'on note :

$$T \sim T_n$$

Représentation graphique

L'aspect de la représentation graphique d'une T_n est celle d'une $\mathcal{N}(0, 1)$ aplatie, incluant les propriétés de symétries que cela sous-entend.



Densité de la loi de Student pour différentes valeurs du paramètre n et celle de $\mathcal{N}(0, 1)$.

Espérance et variance

Si $T \sim T_n$ alors, les valeurs respectives de l'espérance et de la variance de cette variable sont :

$$E(T) = 0, \quad n > 1 \quad \text{et} \quad V(T) = \frac{n}{n-2}, \quad n > 2.$$

Remarque :

- 1) Elle joue un rôle important dans l'estimation par intervalle de confiance et dans la théorie des tests. Elle est symétrique, de moyenne nulle et dépend d'un paramètre n appelé nombre de degrés de liberté ;

- 2) L'aspect de la courbe variera selon le nombre de degrés de liberté n (de façon générale, elle est plus aplatie que $\mathcal{N}(0, 1)$ et quand n augmente ($n > 30$) les deux courbes se confondent).
- 3) Plus la valeur de n est petite, plus la forme de la cloche est large et plate. A mesure que n augmente, celle-ci devient plus étroite et plus haute pour tendre progressivement vers celle d'une $\mathcal{N}(0, 1)$, lorsque $n \rightarrow \infty$.
- 4) Pour n grand $E(T) = 0$ et $V(T) \rightarrow 1$.

Convergence de la loi Student vers la loi normale (approximation)

Soit $X \sim T_n$ alors $X \rightarrow \mathcal{N}(0, 1)$ en loi quand $n \rightarrow \infty$. En pratique : $n > 30$.

C. - La loi de Fisher-Snedecor : $\mathcal{F}(n_1, n_2)$

Définition : Soient $Y_1 \sim \chi_{n_1}^2$ et $Y_2 \sim \chi_{n_2}^2$, deux variables aléatoires indépendantes. Alors :

$$F = \frac{\frac{Y_1}{n_1}}{\frac{Y_2}{n_2}} \sim \mathcal{F}(n_1, n_2)$$

C'est la loi de **Fisher-Snedecor** à n_1 et n_2 degrés de libertés. Cette loi intervient aussi en théorie des tests.

Espérance et Variance

Si $F \sim \mathcal{F}(n_1, n_2)$ alors $E(F) = \frac{n_1}{n_2 - 2}$, $n_2 > 2$ et $V(F) = \frac{2n_1^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}$, $n_2 > 4$.

Remarque

- 1) Si X suit $\mathcal{F}(n_1, n_2)$ alors $\frac{1}{X}$ suit $\mathcal{F}(n_2, n_1)$;
- 2) La variable $\mathcal{F}(n_1, n_2)$ varie de 0 à $+\infty$;
- 3) La loi de probabilité de la variable F dépend de deux paramètres, les degrés de libertés, n_1 et n_2 .

Chapitre 4 : Inégalités et Notions de convergence de variables aléatoires

Inégalités de Markov

Andreï Andreïevitch Markov, 1856-1922, né et mort en Russie, professeur à l'Université de Saint-Pétersbourg. Ses premiers travaux portèrent sur la théorie des nombres et l'analyse mathématique. Il a démontré le *théorème central limite* sous des conditions assez générales. Son étude, de ce qu'on appelle maintenant *chaines de Markov*, a lancé le domaine de processus stochastiques. Il s'intéressait aussi à la poésie.

Théorème : Soit X une variable aléatoire qui ne prend que des valeurs non négatives ; alors on a :

$$P(X \geq a) \leq \frac{E(X)}{a} ; \forall a > 0$$

Inégalités de Bienaymé-Tchebychev

Irénée-Jules Bienaymé, 1796-1878, né et mort en France. Il étudia à l'École Polytechnique de Paris et enseigna dans une académie militaire pendant deux ans. Ensuite, il devint inspecteur général des Finances. En 1848, il fut nommé professeur de probabilités à la Sorbonne. Ami de Tchebychev, il a traduit ses travaux du russe au français. L'œuvre de Laplace, en particulier la *Théorie analytique des probabilités*, eut une grande influence sur sa pensée scientifique.

Soit X une variable aléatoire sur un univers Ω . Nous supposons qu'elle admet une espérance et une variance. Ceci est toujours réalisé si X prend un nombre fini de valeurs (que Ω soit fini ou non), mais pas toujours le cas si X prend une infinité de valeurs.

Les valeurs « sympathiques » et « espérées » d'une variable aléatoire sont celles qui sont proches de son espérance. Les valeurs loin de $E(X)$ sont donc aberrantes. On aimerait donc savoir majorer le risque qu'une variable aléatoire prenne des valeurs trop excentriques. L'inégalité de Bienaymé-Tchebychev est là pour cela.

Théorème : Soit X une variable aléatoire d'espérance $E(X) = \mu$ et de variance $V(X) = \sigma^2$. Alors, pour tout $\alpha > 0$, on a :

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

Exemple fondamental : une moyenne de valeurs observées est plus précise qu'une mesure faite au hasard

Soient X_1, X_2, \dots, X_n des variables aléatoires indépendantes et de même loi d'espérance μ et de variance σ^2 . Notons $\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ la moyenne de ces variables aléatoires.

- La linéarité de l'espérance nous donne :

$$E(\bar{X}_n) = \frac{E(X_1) + E(X_2) + \dots + E(X_n)}{n} = \frac{n\mu}{n} = \mu$$

- Comme les variables sont indépendantes, la variance de la somme est égale à la somme des variances. Nous avons donc :

$$V(\bar{X}_n) = \frac{1}{n^2} V(X_1 + X_2 + \dots + X_n) = \frac{1}{n^2} (V(X_1) + V(X_2) + \dots + V(X_n)) = \frac{1}{n^2} (n\sigma^2) = \frac{\sigma^2}{n}$$

L'espérance reste la même mais la variance est divisée par n , ce qui signifie que la moyenne est beaucoup plus précise que chaque X_i . L'inégalité de Bienaymé-Tchebychev s'écrit alors :

$$P(|\bar{X}_n - \mu| \geq \alpha) \leq \frac{\sigma^2}{n\alpha^2}$$

Pour un α fixé, quand n devient grand, la probabilité que $|\bar{X}_n - \mu|$ soit plus grand que α devient très faible. La moyenne \bar{X}_n prend des valeurs presque certainement concentrées autour de l'espérance μ , et elle est donc un outil très fiable de mesure.

Exemple : Le confiturier annonce que ses pots de confiture pèsent 250 grammes avec un écart-type de $\sigma = 2$ grammes. De façon précise, le poids d'un pot est une variable aléatoire X (dont on ne connaît pas la loi) d'espérance 250 et de variance 4.

La probabilité qu'un pot pris au hasard ait un poids inférieur ou égal à 247 grammes ou bien supérieur ou égal à 253 grammes est donc majorée par $P(|X - 250| \geq 3) \leq \frac{4}{1 \times 3^2} = \frac{4}{9}$.

Il y a donc moins de 44% de chances que $X \leq 247$ ou $X \geq 253$; ce qui est très élevé.

Par contre, si l'on prend la moyenne des poids de 100 pots de confiture, on va trouver $P(|\bar{X}_n - 250| \geq 3) \leq \frac{4}{100 \times 3^2} = \frac{4}{900}$.

Il y a donc moins de 0.44% de chances que $\bar{X}_n \leq 247$ ou $\bar{X}_n \geq 253$; ce qui est déjà plus acceptable.

Si l'on prend donc 100 pots au hasard, et que l'on constate que la moyenne des poids est inférieure à 246 grammes, alors que la probabilité de cette éventualité est inférieure à 5%, il est raisonnable de se dire que le confiturier nous roule ou bien que ses pots sont trop remplis.

Bienaymé-Tchebychev ~ Bernoulli et Binomiale

Considérons maintenant X_1, \dots, X_n une suite de variables aléatoires indépendantes suivant la même loi de Bernoulli de paramètre p . En pratique, on répète indéfiniment et de façons indépendantes la même expérience se traduisant par un succès avec probabilité p et un échec avec probabilité $1 - p$.

La variable aléatoire X_i prend la valeur 1 si l'expérience numéro i est un succès, et prend la valeur 0 sinon.

L'espérance commune est égale à p , et la variance vaut $p(1 - p)$. L'inégalité de Bienaymé-Tchebychev s'écrit :

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - p\right| \geq \alpha\right) \leq \frac{p(1-p)}{n\alpha^2}$$

Le réel p est compris entre 0 et 1. Le maximum de la fonction $x(1-x)$ sur $[0; 1]$ est atteint lorsque $2x - 1$ (sa dérivée) s'annule, c'est-à-dire lorsque $x = \frac{1}{2}$, et il est égal à $\frac{1}{4}$.

On peut donc majorer $p(1-p)$ par $\frac{1}{4}$. On obtient alors :

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - p\right| \geq \alpha\right) \leq \frac{1}{4n\alpha^2}$$

Remarque : $S_n = X_1 + \dots + X_n$ compte le nombre de succès parmi les n expériences, et suit la loi binomiale de paramètres n et p . Le quotient $\frac{S_n}{n}$ mesure la proportion de succès parmi les n essais, ce que l'on appelle la fréquence du succès.

Quand n devient très grand, la fréquence de succès parmi les expériences est de plus en plus proche de la probabilité réelle p de succès.

Pafnouti Lvovitch Tchebychev, 1821-1894, né et mort en Russie. Noble russe et professeur à l'Université de Saint-Petersbourg, il est considéré comme le fondateur de la grande école de mathématiques russe du XX^e siècle. En plus de ses contributions à la théorie des nombres et à la théorie de l'approximation des fonctions, il s'intéressa beaucoup à la mécanique. En utilisant l'inégalité qui porte son nom, il donna une preuve simple de la *loi des grands nombres*. Il a aussi travaillé intensément sur le *théorème central limite*.

Loi faible des grands nombres – Théorème Central Limite (TCL)

Ces deux théorèmes fondamentaux sont valables aussi bien pour des variables aléatoires discrètes que continues, à partir du moment où l'espérance et la variance existent.

NB : « équidistribués » signifie « suivant toutes la même loi »

Loi faible des grands nombres

Soit X_1, \dots, X_n, \dots une suite infinie de variables aléatoires indépendantes équidistribuées ayant une même espérance μ et une variance finie. Alors, pour tout $\alpha > 0$, on a :

$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| \geq \alpha\right) = 0$$

Il suffit en effet de faire tendre n vers l'infini dans l'inégalité de Bienaymé-Tchebychev

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| \geq \alpha\right) \leq \frac{V}{n\alpha^2}$$

où V est la variance commune.

On dit que la suite (X_n) converge en probabilité vers la variable certaine égale à μ .

Cas particulier des variables de Bernoulli

Considérons l'exemple où les (X_i) suivent la loi de Bernoulli de paramètre p . La variable $\frac{X_1 + \dots + X_n}{n}$ représente la fréquence du nombre de succès parmi les n expériences.

La loi faible des grands nombres affirme que la suite de ces fréquences converge vers p .

Soit p la probabilité réelle mais inconnue de voter A à une élection qui ne compte que deux candidats : c'est donc le quotient entre le nombre d'électeurs qui vont voter A le jour du scrutin et le nombre de votants. Cette proportion est inconnue avant le scrutin, mais elle existe.

Supposons que les électeurs soient bien fixés sur leur choix un certain temps à l'avance.

On interroge n personnes avant le vote, et on leur demande de dire pour qui ils vont voter. La proportion observée de votants pour A va converger vers p , et on aura ainsi une idée de p .

Ceci sous-entend que les réponses sont indépendantes les unes des autres mais c'est le cas dans une grande population. Le plus difficile consiste ensuite, à partir de la réponse observée, d'en déduire une fourchette dans laquelle on est raisonnablement certain que se trouve la proportion p inconnue. Nous donnerons quelques éléments de réponse avec les intervalles de confiance (*voir plus loin*).

Convergence en probabilité

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires. On dit que (X_n) converge en probabilité vers la variable aléatoire X ($X_n \xrightarrow{\text{proba}} X$) ssi :

$$\forall \alpha, \eta, \exists N, (n \geq N) \Rightarrow P(|X_n - X| > \alpha) < \eta$$

ou simplement :

$$\lim_{n \rightarrow \infty} P(|X_n - X| \geq \alpha) = 0$$

Corollaire de la loi des grands nombres

Soient X_1, \dots, X_n , n variables aléatoires indépendants, de même loi. Si $\mu = E(X_i)$, alors $\bar{X} \rightarrow \mu$ en probabilité.

Définition : On dit que (X_n) converge en loi vers la variable aléatoire X ($X_n \rightarrow X$ en loi) ssi :

$$\forall x, F_n(x) \rightarrow F(x)$$

$F_n(x)$ et $F(x)$ étant les fonctions de répartition de X_n et X .

N.B : La convergence en probabilité implique la convergence en loi mais la réciproque est fausse.

Théorème Central Limite (TCL)

Soit X_1, \dots, X_n, \dots une suite infinie de variables aléatoires indépendantes équidistribuées ayant une même espérance μ et une variance σ^2 .

Notons $\bar{X}_n = \frac{X_1 + X_2 + \dots}{n}$. Alors $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$ converge en loi vers la loi normale centrée réduite

$\mathcal{N}(0, 1)$, ce qui signifie que :

$$\lim_{n \rightarrow +\infty} P\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

Annexe : Rappels sur les ensembles

Définition 1 : Un ensemble est une collection d'objets. Les objets sont appelés *éléments*.

Notation : On note les ensembles par des lettres majuscules et les éléments par des lettres minuscules ou des chiffres.

Relation entre un ensemble et un élément : Soit A un ensemble. On écrit $p \in A$ (et on lit « p appartient à A ») si p est un élément de A .

Exemple : $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ est l'ensemble des entiers naturels. On a $1 \in \mathbb{N}$, $4 \in \mathbb{N}$ mais $-2 \notin \mathbb{N}$.

Définition 2 : Soient A et B deux ensembles. Si chaque élément de A est également un élément de B , alors A est un sous-ensemble de B . On écrit $A \subset B$ (et on lit « A est inclus dans B »).

Exemple : Soit $B = \mathbb{N} = \{0, 1, 2, 3, \dots\}$ et $A = \{0, 2, 4, 6, \dots\}$. Il est clair que $A \subset B$.

Remarques :

1. Si $A \subset B$ et $B \subset A$, alors $A = B$.

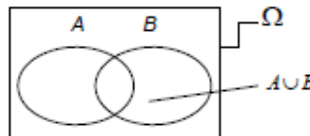
2. La négation de $p \in A, A \subset B, A = B$ est $p \notin A, A \not\subset B, A \neq B$ respectivement.

Définition 3 : Dans le contexte de la théorie des probabilités, tous les ensembles considérés sont des sous-ensembles d'un ensemble universel que l'on désigne par Ω . L'ensemble vide qui est l'ensemble ne contenant aucun élément est désigné par \emptyset .

Opérations sur les ensembles

1. Union

Soient A et B deux ensembles. L'union de A et B est l'ensemble des éléments qui appartiennent à A ou à B . On désigne l'union de A et B par $A \cup B$. On dira que $x \in A \cup B$ si x appartient à au moins l'un des deux ensembles A et B .



Remarque : La figure ci-dessus est appelée diagramme de Venn.

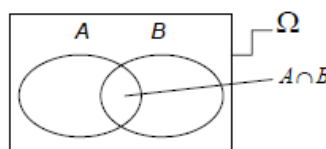
Exemple :

1) Soient $A = \{1, 2, 3, 4\}$ et $B = \{2, 4, 5, 6\}$. Alors $A \cup B = \{1, 2, 3, 4, 5, 6\}$.

2) Soit A l'ensemble des individus aux cheveux blonds ; soit B l'ensemble des individus aux cheveux bruns. Alors $A \cup B$ est l'ensemble des individus qui ont les cheveux blonds ou les cheveux bruns.

2. Intersection

Soient A et B deux ensembles. L'intersection de A et B est l'ensemble des éléments qui appartiennent à la fois à A et à B . On désigne l'intersection de A et B par $A \cap B$.

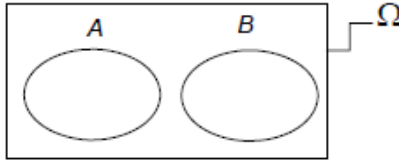


Exemples :

1) Soit $A = \{1, 2, 3, 4\}$ et $B = \{2, 4, 5, 6\}$. Alors $A \cap B = \{2, 4\}$.

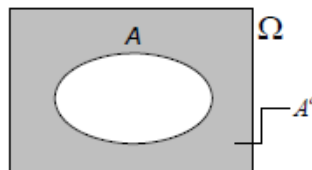
2) Soit $A = \{1, 3, 5, \dots\}$ et $B = \{0, 2, 4, 6, \dots\}$. Alors $A \cap B = \emptyset$.

Remarque : Si $A \cap B = \emptyset$, on dit que A et B sont *mutuellement exclusifs* ou *disjoints* ou *incompatibles*. Dans ce cas, le diagramme de Venn est le suivant :



3. Complémentaire

Le complément d'un ensemble A est l'ensemble des éléments qui n'appartiennent pas à A . Le complément de A est désigné par $A^c = \bar{A}$.



Exemple : Si A est l'ensemble des fumeurs en Haïti, \bar{A} est l'ensemble des non-fumeurs.

Remarque : A et \bar{A} sont toujours disjoints, c'est-à-dire $A \cap \bar{A} = \emptyset$.

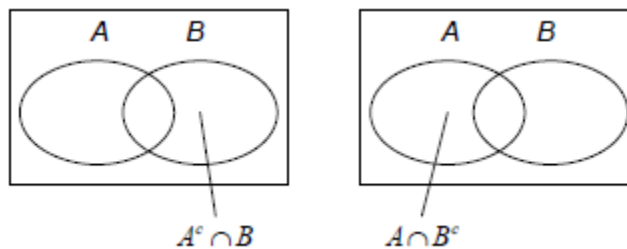
Règle de base

Union	Intersection
$A \cup A = A$	$A \cap A = A$
$A \cup B = B \cup A$	$A \cap B = B \cap A$
$A \cup \emptyset = A$	$A \cap \emptyset = \emptyset$
$A \cup \Omega = \Omega$	$A \cap \Omega = A$
$A \cup \bar{A} = \Omega$	$A \cap \bar{A} = \emptyset$
$(A^c)^c = A$	$\Omega^c = \emptyset, \quad \emptyset^c = \Omega$
Loi de DeMorgan	
$\overline{A \cup B} = \bar{A} \cap \bar{B}$	$\overline{A \cap B} = \bar{A} \cup \bar{B}$

Autres ensembles intéressants

- $A \cap \bar{B} : x \in A \cap \bar{B} \Leftrightarrow x \in A \text{ et } x \notin B$
- $\bar{A} \cap B : x \in \bar{A} \cap B \Leftrightarrow x \notin A \text{ et } x \in B$

Diagramme de Venn



5.- Références Bibliographiques

1. G. BAILLARGEON, *Méthodes Statistiques avec applications en gestion, production, marketing, relations industrielles et sciences comptables*, 3^{ème} édition. Les Éditions SMG ;
2. Anderson, Sweeney et Williams, *Statistiques pour l'économie et la gestion*, traduction de la 5^e édition américaine par Claire Borsenberger, 3^e édition ;
3. F. BERTRAND, M. MAUMY-BERTRAND et E. PÉRINEL, *Probabilités, Statistiques et Économétrie*. DUNOD ;
4. G. DENGLOS, *Statistiques et Probabilités appliquées*, puf ;
5. M.R. SPIEGEL, *Statistique*, EdiScience ;
6. Sanders, Murph, Eng, *Les statistiques une approche nouvelle*, Mc Graw-Hill, Inc ;
7. Luc Amyotte, *Méthodes Quantitatives*, Éditions du Renouveau Pédagogique ;
8. G. BAILLARGEON, *Introduction à la statistique appliquée*, Les Éditions SMG ;
9. Jean M. Martel, R. Nadeau, *Statistiques en Gestion et en Économie*, gaëtan morin, éditeur ;
10. Michèle Avenel, Jean-François Riffault, *Mathématiques appliquées à la gestion*, Sup'FOUCHER
11. E. Bosia, L. Gautier, A. Ruchalski, *Statistiques appliquées à la gestion*, Escem 2012-2013.