

CHAPITRE PREMIER. CALCUL DES PROBABILITES

I.1 Introduction

Dans des domaines très différents comme le domaine scientifique, sociologique, médical, les sciences humaines..., on s'intéresse à de nombreux phénomènes dans lesquels apparaît souvent l'effet du hasard. Ces phénomènes sont caractérisés par le fait que les résultats des observations varient d'une expérience à l'autre.

Une expérience est appelée aléatoire s'il est impossible de prévoir son résultat et si, répétée dans des conditions identiques, elle peut donner, ou aurait pu donner, si l'expérience est unique, des résultats différents. En général, les résultats obtenus varient dans un certain domaine, certains résultats apparaissant plus fréquemment que d'autres.

Le mot probabilité est passé rapidement dans le langage courant bien que la théorie des probabilités soit une branche relativement récente des théories mathématiques.

Le concept des probabilités semblait être connu des Grecs et des Égyptiens. Cependant, ce n'est que vers le milieu du XVIIe siècle que l'on peut situer le début de cette théorie. D'abord limitée à l'étude des jeux de hasard (jeux de pile ou face, roulettes, jeux de cartes...), elle s'est rapidement étendue à tous les domaines de la Science, en Physique (théorie du potentiel, physique statistique, physique corpusculaire...), en Informatique, en Économie et Gestion, en Génétique, en Psychologie... L'influence des jeux de hasard se retrouve encore dans certaines expressions, comme l'espérance mathématique qui était l'espérance du

gain, pouvant être parfois une perte. Le mot probabilité ou l'adjectif probable est bien souvent synonyme du mot chance.

La probabilité d'événements répétitifs est définie à partir de la fréquence d'apparitions de ces événements. On distingue différents concepts :

I.1.1 L'approche fréquentielle

C'est la théorie de Laplace; elle est fondée sur la notion d'épreuves répétées et indépendantes, la probabilité étant définie comme la limite de la fréquence relative des observations.

Cette fréquence, exprimée comme le rapport $\frac{n_a}{n}$ (n_a étant le nombre d'essais où l'événement A a été réalisé au cours de n essais indépendants, répétés dans des conditions identiques), a des fluctuations autour d'une valeur limite qui est la probabilité de l'événement A (loi des grands nombres). Mais, on suppose implicitement que la fréquence relative tend vers cette limite avec une grande probabilité ! C'est-à-dire, que l'on définit la probabilité à partir de la probabilité.

I.1.2 L'approche classique

Dans la notion de probabilité tirée des jeux de hasard, la probabilité est le quotient du nombre de cas favorables par le nombre de cas possibles, mais chaque cas étant supposé également possible, donc équiprobable, on définit encore la probabilité à partir de la probabilité.

$$\Pr(A) = \frac{n}{N} = \frac{\text{Nombre des cas favorables}}{\text{Nombre des cas possibles}}$$

I.2 Définitions de quelques concepts probabilistes

À l'origine le concept probabiliste consistait à répartir, sur chacun des éléments d'un ensemble, un ensemble de valeurs ou probabilités dont la somme était égale à 1.

Si cet ensemble, ou espace des épreuves, est de dimension finie, il n'y a pas de difficultés majeures. En revanche, si cet espace a la puissance du continu, le problème d'associer à chacun de ses éléments, une probabilité, est pratiquement sans solution.

Pour formaliser ces notions, trois étapes sont nécessaires :

- définir le cadre dans lequel on observe les manifestations du hasard, c'est à- dire définir une expérience aléatoire et l'ensemble fondamental Ω ,
- définir un événement aléatoire et la classe C des événements aléatoires,
- définir une probabilité sur l'espace (Ω, C) , c'est-à-dire affecter un poids à chaque événement traduisant la chance de réalisation de cet événement.

I.2.1 Expérience aléatoire ou stochastique

Une expérience est dite aléatoire s'il est impossible d'en prévoir le résultat, c'est-à-dire, si répétée dans les mêmes conditions, elle peut donner des résultats différents, dans un ensemble d'issues considérées comme possibles :

- succession d'appels à un standard téléphonique non surchargé,
- observation de la durée de vie d'un individu anonyme dans une population humaine,
- observation de la durée de fonctionnement sans panne d'un appareil,
- jeu de pile ou face de durée infinie...

Les résultats d'une expérience aléatoire appartiennent à un espace fondamental ou espace des épreuves Ω ; un point quelconque ω de Ω est un résultat élémentaire. D'où la définition :

Une expérience aléatoire est un choix au hasard d'un point ω dans un ensemble Ω .

L'ensemble Ω dépend des connaissances que l'on a, a priori, sur les résultats possibles de l'expérience aléatoire.

Exemples I.1

On lance une pièce de monnaie. Pour l'ensemble Ω , on peut choisir :

- soit l'ensemble $\Omega_1 = \{\text{pile, face}\}$,
- soit l'ensemble $\Omega_2 = \{\text{pile, face, tranche}\}$.

On considère la succession des appels à un standard téléphonique non surchargé et on étudie la répartition des instants où le standard reçoit un appel, à partir d'un instant choisi comme origine (on admet que deux appels ne peuvent se produire rigoureusement au même instant, et que le phénomène n'est pas limité dans le temps).

I.2.2 Événement aléatoire

Un événement aléatoire est lié à une expérience aléatoire; une fois l'expérience réalisée, on peut alors dire si l'événement a été réalisé ou non. Un événement aléatoire A peut être identifié à la partie de Ω dont les éléments réalisent l'événement A .

Exemple I.2

On jette deux dés et soit A l'événement :

« le total des points est supérieur ou égal à 11 ».

L'ensemble des résultats possibles est l'ensemble $\Omega = \{(1, 1), (1, 2), \dots, (6, 6)\}$.

Un total supérieur ou égal à 11 est représenté par l'ensemble des trois couples :

$\omega = \{x, y\}$ tels que $x + y \geq 11$, c'est-à-dire les couples $\{5, 6\}, \{6, 5\}, \{6, 6\}$.

Cet ensemble de trois couples réalise l'événement A.

On pourrait choisir pour l'ensemble des événements, l'ensemble $P(\Omega)$ des Parties de Ω , mais comme cet ensemble est en général trop vaste pour être « probabilisé », on se limite à un ensemble strictement contenu dans $P(\Omega)$, vérifiant les propriétés logiques suivantes, qui servent de base axiomatique à la définition mathématique de la notion d'événement aléatoire.

I.2.3 Opération sur les événements

– À tout événement A est associé son contraire, non A ou $\bar{A}$ qui est réalisé si et seulement si A ne l'est pas.

Dans l'espace Ω des événements, A et $\bar{A}$ sont représentés par des ensembles complémentaires au sens ensembliste.

– Pour tout couple d'événements A et B, l'événement « A et B » est réalisé si A et B sont réalisés.

Dans l'espace Ω des événements, l'événement « A et B » est représenté par l'intersection des ensembles réalisant A et B, on le note « A et B » ou « $A \cap B$ ».

– Pour tout couple d'événements A et B, l'événement « A ou B » est réalisé si l'un des deux ou si les deux sont réalisés.

Dans l'espace Ω des événements, il est représenté par la réunion des ensembles réalisant A et B, on le note, ou n'étant pas exclusif, « A ou B » ou

« $A \cup B$ ».

– Deux événements A et B sont incompatibles si la réalisation de l'un implique la non réalisation de l'autre,

Dans l'espace Ω des événements, deux événements incompatibles sont représentés par deux parties disjointes.

– Les événements $A_1, A_2, \dots, A_n$ forment un système complet d'événements ou système exhaustif si les ensembles qui leur sont associés forment une partition de l'espace Ω .

Tableau I.1 –Terminologies probabiliste et ensembliste.

Terminologie probabiliste	Terminologie ensembliste	Notation
Événement certain	Espace entier	Ω
Événement impossible	Partie vide	$\emptyset$
Événement contraire	Complémentaire	$\bar{A}$ ou A^c
A et B	Intersection	$A \cap B$
A ou B (ou non exclusif)	Réunion	$A \cup B$
Événements incompatibles	Parties disjointes	$A \cap B = \emptyset$
Système complet d'événements	Partition de Ω	$A_i \cap B_j = \emptyset$ $\cup A_i = \Omega$
Implication $A \Rightarrow B$	Inclusion	$A \subset B$

Implication $A \subset B$ ou $A \Rightarrow B$: l'événement A ne peut être réalisé sans que B le soit.

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 C de parties de Ω (tribu ou σ -algèbre).

En résumé :

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

I.3 Mesure de probabilité et espace probabilisé

I.3.1 Définition de la probabilité

Intuitivement, si A et B sont deux événements incompatibles, la chance de voir se réaliser A ou B doit être égale à la somme des poids traduisant les chances de réalisation de A et B . De même, si (A_n) , n appartenant à N , désigne un ensemble d'événements tel que chacun d'eux est impliqué par le suivant et tel que leur réalisation simultanée est impossible, alors le poids de A_n a une limite nulle quand n tend vers l'infini.

Une probabilité Pr définie sur l'ensemble (Ω, C) , est une application de C dans $[0, 1]$ telle que :

- $Pr(\Omega) = 1$
- $Pr(\cup A_i) = \sum Pr(A_i)$ pour toute réunion finie ou dénombrable d'événements incompatibles.

Le triplet (Ω, C, Pr) est un espace probabilisé, la mesure Pr ainsi définie est une mesure positive de masse totale égale à $+ 1$.

I.3.2 Propriétés élémentaires

Elles se déduisent des axiomes de définition :

- $Pr(\emptyset) = 0$ mais $Pr(A) = 0$ n'implique pas $A = \emptyset$

L'événement A tel que $\Pr(A) = 0$ est un événement presque impossible.

- $\Pr(\bar{A}) = 1 - \Pr(A)$
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \rightarrow$ **théorème d'addition**
- $\Pr(\cup A_i) \leq \sum \Pr(A_i)$ (aucune hypothèse particulière sur les événements A_i)

– si la suite des événements A_i tend vers 0 en décroissant, la limite de $\Pr(A_i)$ est nulle.

– si B_i est un système complet d'événements, alors

$$\forall A, \Pr(A) = \sum \Pr(A \cap B_i)$$

C'est la première forme du théorème des probabilités totales.

Remarque

$\Pr(A) = 1$ n'implique pas $A = \Omega$. L'événement A tel que $\Pr(A) = 1$ est un événement presque certain.

I.4 Échantillons et sous-populations

De nombreux problèmes faisant intervenir le calcul des probabilités se ramènent aux problèmes de tirer des échantillons de taille r dans un ensemble de taille n , appelé population, quelle que soit la nature de ses éléments. Suivant la règle du tirage, cet échantillon est :

- ordonné ou non,
- avec ou sans répétitions (on dit aussi avec ou sans remise).

Deux autres espaces interviennent souvent dans des problèmes élémentaires, l'espace des sous-populations de taille r avec répétitions et l'espace des permutations de n objets.

Remarque

Choisir un élément au hasard, signifie que les divers choix possibles sont équiprobables donc que l'ensemble Ω est muni de la loi de probabilité

uniforme. Dans ce cas, tous les calculs sont simples et se ramènent souvent à des calculs d'analyse combinatoire.

I.5 PROBABILITÉ CONDITIONNELLE ET INDÉPENDANCE

I.5.1 Définition

Soit (Ω, C, Pr) un espace probabilisé. L'intersection de deux événements A et B est l'événement, noté $A \cap B$, réalisé, si et seulement si, les deux événements A et B sont réalisés. Cependant, on peut s'intéresser à la réalisation de l'événement A sachant l'événement B réalisé, si cet événement est de probabilité non nulle, c'est-à-dire on s'intéresse à la probabilité conditionnelle sachant B .

La probabilité conditionnelle sachant B est l'application de C dans $[0, 1]$ définie par :

$$\forall A \in C \quad Pr(A/B) = \frac{Pr(A \cap B)}{Pr(B)} \rightarrow \text{Théorème de}$$

multiplication

Cette application définit une probabilité sur le même espace probabilisé (Ω, C, Pr) , la probabilité conditionnelle $Pr(. / B)$ est définie comme la probabilité

Pr sur la tribu C , le terme $Pr(B)$ est un facteur de normalisation.

Selon les événements A et B , différents cas sont possibles.

I.5.1.1 Les événements A et B sont incompatibles

L'événement A ne se réalisera pas si l'événement B est réalisé :

$$Pr(A/B) = 0$$

Exemple I.5.1

On lance deux dés et on considère les deux événements :

– A : obtenir un chiffre impair sur les deux dés,

– B : la somme des points obtenus sur les deux dés est un nombre impair.
Ces deux événements sont incompatibles.

I.5.1.2 Les événements A et B ne sont pas incompatibles

Deux événements peuvent être totalement dépendants ou dépendants.

– Événements totalement dépendants

Deux événements A et B sont totalement dépendants si $A \subset B$, ou si l'événement B étant réalisé, la probabilité de réalisation de l'événement A est égale à 1 :

$$\Pr(A/B) = 1$$

On dit que A dépend totalement de B.

Exemple I.5.2

Les événements suivants sont totalement dépendants :

- A : le nombre est égal à 4, 6, 8,
- B : le nombre est un nombre pair compris entre 2 et 20.
- Événements dépendants

Deux événements A et B sont dépendants si la probabilité de réalisation de l'événement A change selon que B est réalisé ou non.

Exemple I.5.3

On lance un dé parfaitement équilibré et on considère les événements suivants :

- A : obtenir la face 6,
- B : obtenir un nombre pair,
- C : obtenir un nombre supérieur ou égal à 3.

$$\Pr(A) = 1/6 \quad \Pr(B) = 1/2 \quad \Pr(C) = 4/6 = 2/3$$

Si l'événement B réalisé, la probabilité de réalisation de A est égale à 1/3.

Si l'événement C réalisé, la probabilité de réalisation de A est égale à 1/4.

Les probabilités conditionnelles de A ne sont donc pas égales à la probabilité de A ni égales entre elles :

$$\Pr(A) = 1/6 \quad \Pr(A/B) = 1/3 \quad \Pr(A/C) = 1/4$$

Les événements A et B d'une part, A et C d'autre part sont dépendants.

I.5.2 Principe des probabilités composées

Le principe des probabilités composées découle des axiomes et des définitions.

Il s'écrit :

$$\Pr(A \cap B) = \Pr(A/B) \cdot \Pr(B) = \Pr(B/A) \cdot \Pr(A)$$

Cette formule est valable même si les probabilités $\Pr(A)$ et $\Pr(B)$ sont nulles toutes les deux ; mais dans ces conditions, on ne peut pas définir $\Pr(A/B)$ ni $\Pr(B/A)$.

I.5.3 Événements indépendants

I.5.3.1 Définition

L'événement A est indépendant de l'événement B si la probabilité de réalisation de l'événement A n'est pas modifiée par une information concernant la réalisation de l'événement B, c'est-à-dire si :

$$\Pr(A/B) = \Pr(A)$$

Le principe des probabilités composées entraîne :

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B) = \Pr(B/A) \Pr(A)$$

$$\Pr(B/A) = \Pr(B)$$

L'événement B est donc également indépendant de l'événement A. Les événements A et B sont indépendants et vérifient la propriété :

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$$

I.5.3.2 Événements incompatibles et événements indépendants

– La propriété « les événements A et B sont incompatibles » implique :

$$\Pr(A \cup B) = \Pr(A) + \Pr(B)$$

– La propriété « les événements A et B sont indépendants » implique :

$$\Pr(A \cap B) = \Pr(A) * \Pr(B)$$

Les opérations union $\cup$ et somme semblent jouer le même rôle que les opérations intersection $\cap$ et produit. Cependant, les deux concepts, incompatibles et indépendants, sont totalement différents :

– Le premier « événements incompatibles » est une notion ensembliste.

– Le second « événements indépendants » est une notion probabiliste : deux événements peuvent être indépendants pour une loi de probabilité et non pour une autre loi.

I.5.4 Théorème de Bayes

I.5.4.1 Deuxième forme du théorème des probabilités totales

On considère un événement A de probabilité non nulle et l'ensemble $(C_i) i \in (1, \dots, n)$ de toutes les causes possibles de réalisation de cet événement ; cet ensemble forme un ensemble complet d'événements et l'événement A se produit en même temps qu'un et un seul des C_i , c'est-à-dire :

$$A = (A \cap C_1) \cup (A \cap C_2) \cup \dots \cup (A \cap C_n)$$

On en déduit la deuxième forme du théorème des probabilités totales :

$$\Pr(A) = \sum_{i=1}^n \Pr(A \cap C_i) = \sum_{i=1}^n \Pr(A/C_i) \Pr(C_i)$$

I.5.4.2 Théorème de Bayes

Considérons une des causes susceptibles de réaliser l'événement A , la cause C_k par exemple. Le théorème des probabilités composées donne :

$$\Pr(A \cap C_k) = \Pr(A/C_k) \Pr(C_k) = \Pr(C_k/A) \Pr(A)$$

De la deuxième forme du théorème des probabilités totales, on déduit $\Pr(A)$, puis le théorème de Bayes :

$$\Pr(C_k/A) = \frac{\Pr(A/C_k) \Pr(C_k)}{\sum_{i=1}^n \Pr(A/C_i) \Pr(C_i)}$$

Sous cette forme, le théorème de Bayes (publié après sa mort en 1763) apparaît comme une conséquence logique des axiomes et des définitions. Il présente un grand intérêt, car il permet de modifier notre connaissance des probabilités en fonction d'informations nouvelles, il joue un rôle très important dans la statistique bayésienne.

Exemple I.5.4

Trois machines automatiques produisent des pièces de voitures. La machine M_1 produit 40 % du total des pièces, la machine M_2 25 % et la machine M_3 produit 35 %. En moyenne, les pourcentages des pièces non conformes aux critères imposés sont de 10% pour la machine M_1 , de 5 % pour la machine M_2 et de 1 % pour la machine M_3 .

Une pièce est choisie au hasard dans la production totale des trois machines. On constate qu'elle n'est pas conforme aux critères imposés.

Quelle est la probabilité qu'elle ait été produite par la machine M_1 ?

On peut appliquer directement le théorème de Bayes.

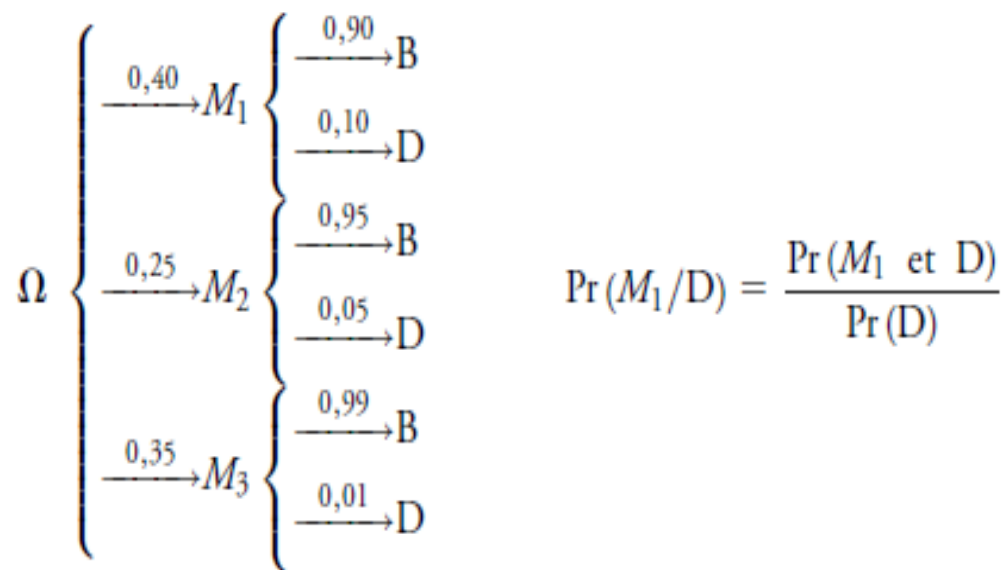
Soit B l'événement « la pièce est bonne » et D l'événement « la pièce est défectueuse ».

Les trois causes possibles de réalisation de l'événement D sont les trois machines.

On connaît les probabilités de ces causes par exemple $\Pr(M_1) = 0,40$ ainsi que les probabilités conditionnelles $\Pr(D/M_1) = 0,10$.

$$\Pr(M_1/D) = \frac{\Pr(D/M_1) \Pr(M_1)}{\sum_{i=1}^3 \Pr(D/M_i) \Pr(M_i)}$$

On peut visualiser ce problème par l'arbre suivant :



$$\Pr(D) = 0,40 \times 0,10 + 0,25 \times 0,05 + 0,35 \times 0,01 = 0,056$$

$$\Pr(M_1/D) = \frac{0,40 \times 0,10}{0,056} = 0,714$$

CHAPITRE DEUXIEME. VARIABLES ALÉATOIRES ET DISTRIBUTION DES PROBABILITES

II.1 Généralités sur les variables aléatoires

II.1.1 Définition d'une variable aléatoire

Les variables aléatoires constituent un espace fondamental d'éléments aléatoires, un tel élément étant défini par référence à une expérience aléatoire.

Si $(\Omega, \mathcal{C}, \Pr)$ désigne un espace probabilisé et (E, ε) un espace probabilisable, un élément aléatoire, défini sur $(\Omega, \mathcal{C}, \Pr)$ et à valeurs dans (E, ε) , est une application mesurable de $(\Omega, \mathcal{C})$ dans (E, ε) . Cet élément est appelé :

– variable aléatoire réelle si l'espace (E, ε) est l'espace $(\mathbb{R}, \mathcal{B})$, où $\mathcal{B}$ est la tribu de Borel de $\mathbb{R}$,

Dans ce chapitre, on ne définira que des variables aléatoires réelles. Les propriétés de ces variables sont donc celles des fonctions réelles mesurables.

Exemple II.1 Variable aléatoire

On jette n fois une pièce de monnaie. L'espace fondamental est $\Omega = (P, F)^n$ où P désigne pile et F face ; la tribu associée est la tribu $\mathcal{P}(\Omega)$ des parties de Ω .

On peut s'intéresser :

– soit aux résultats élémentaires :

$\omega = (\omega_1, \omega_2, \dots, \omega_n)$ où ω_i désigne soit pile, soit face.

On obtient, par exemple, la succession $\omega = (P, F, F, F, P, F)$ pour $n=6$

– soit au nombre de fois où « pile » est sorti au cours des n jets.

On obtient, par exemple, 2 fois pile quand on a lancé 6 fois la pièce.

On définit une fonction X application de Ω dans Ω' ($1, 2, \dots, n$) où $X(\omega)$ est le nombre de fois où « pile » apparaît dans ω .

Si $\omega = (P, F, F, F, P, F)$, $X(\omega) = 2$. Si la pièce est parfaitement équilibrée, il semble logique de munir $(\Omega, P(\Omega'))$ de la loi de probabilité uniforme :

$$\Pr(P) = \Pr(F) = 1/2$$

Sur l'espace $(\Omega', P(\Omega'))$, on définit une probabilité $\Pr_x$ ou $\Pr'$, image de $\Pr$ par l'application :

$$\forall A' \in P(\Omega') \quad \Pr'(A') = \Pr(X^{-1}(A'))$$

Cette application X est une variable aléatoire.

II.1.2 Loi de probabilité d'une variable aléatoire X

La loi de probabilité d'une variable aléatoire réelle X est la loi de probabilité $\Pr_x$ définie sur l'espace (R, B) par :

$$\forall B \in B \quad \Pr_x(B) = \Pr(\omega/X(\omega) \in B) = \Pr(X^{-1}(B))$$

On montre facilement que $\Pr_x$ est une mesure positive sur (R, B) et comme :

$$\Pr_x(R) = \Pr(X^{-1}(R)) = \Pr(\Omega) = 1$$

Cette mesure est une probabilité. $\Pr_x$ est la mesure image de $\Pr$ par X .

$(R, B, \Pr_x)$ est l'espace de probabilité associé à la variable aléatoire réelle X .

Une variable aléatoire réelle traduit donc l'idée de résultat numérique associé à un phénomène aléatoire.

Exemple II.2 Loi de probabilité

On jette deux dés équilibrés et on s'intéresse à la somme S des points figurant sur les deux dés. On définit les espaces Ω et Ω' par :

$$- \Omega = (1, 2, \dots, 6)^2$$

$$- \text{et } \Omega' = (2, 3, \dots, 12)$$

Ω est l'espace fondamental ou ensemble des couples $\omega = (n_1, n_2)$, n_1 et n_2 prenant les valeurs entières entre 1 et 6, bornes comprises, et Ω' est l'ensemble des résultats possibles, c'est-à-dire l'ensemble des valeurs que la somme S peut prendre.

Soit X l'application de Ω dans Ω' telle que :

$$X(\omega) = (n_1, n_2)$$

$\Pr(\omega) = 1/36$ car tous les éléments de Ω ont la même probabilité de réalisation et le cardinal de Ω est égal à 36.

Par définition, $P'(A') = \Pr(X^{-1}(A'))$

Ainsi, $\Pr'(6) = \Pr\{X^{-1}(6)\} = \Pr\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\} = 5/36$.

En résumé, les propriétés des variables aléatoires réelles sont donc celles des fonctions mesurables. Il en résulte, en particulier que la composition, la somme, le produit de deux variables aléatoires réelles, la limite d'une suite dénombrable de variables aléatoires réelles est une variable aléatoire réelle.

II.1.3 Fonction de répartition

Une loi de probabilité est une mesure abstraite sur $\mathbb{R}$, elle est donc en général peu maniable et peu utilisée dans les applications concrètes. Or, la tribu de Borel $\mathcal{B}$ de $\mathbb{R}$ contient les intervalles du type $] -\infty, x[$, on en déduit la notion de fonction de répartition.

II.1.3.1 Définition

La fonction de répartition de la variable aléatoire réelle X est l'application F de $\mathbb{R}$ dans $\mathbb{R}$ définie par :

$$\forall x \in \mathbb{R} \quad F(x) = \Pr_x(-\infty, x) = \Pr\{\omega / X(\omega) < x\}$$

On écrit plus simplement :

$$\forall x \in \mathbb{R} \quad F(x) = \Pr(X < x)$$

II.1.3.2 Principales propriétés

Elles se déduisent de la définition et des propriétés d'une probabilité (mesure positive, finie, définie sur $\mathbb{R}$) :

- une fonction de répartition est une fonction F définie sur $\mathbb{R}$ et à valeurs dans $[0, 1]$,
- une fonction de répartition est une fonction croissante au sens large,
- la limite de $F(x)$ quand x tend vers $-\infty$ est égale à 0,
- la limite de $F(x)$ quand x tend vers $+\infty$ est égale à 1,
- une fonction de répartition est continue à gauche, c'est-à-dire $F(x) = F(x-)$
- si la variable aléatoire réelle est continue, la fonction F est continue à droite et dérivable,
- la fonction de répartition permet de traiter tous les problèmes faisant intervenir une seule variable aléatoire X . La probabilité de tout intervalle de $\mathbb{R}$ est égale à :

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

II.1.4 Densité de probabilité

II.1.4.1 Définition

Si la loi de probabilité $\Pr_x$ d'une variable aléatoire réelle X admet une densité par rapport à la mesure de λ sur $\mathbb{R}$, cette densité est appelée densité de probabilité de la variable X . Plus simplement, on peut définir la densité f si elle existe par :

$$f(x) dx = \Pr(x \leq X < x + dx)$$

II.1.4.2 Relation entre fonction de répartition et densité

Soit X une variable aléatoire réelle et F sa fonction de répartition.

Si de plus f est continue en x , F est dérivable et $F'(x) = f(x)$. On peut alors écrire :

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

II.2 Moments d'une variable aléatoire

II.2.1 Espérance mathématique

II.2.1.1 Variable aléatoire discrète

X est une variable aléatoire réelle prenant un ensemble fini ou dénombrable de valeurs sur un espace probabilisé $(\Omega, \mathcal{C}, \Pr)$

$\{\alpha_i\}$ ($i \in I$) est l'ensemble des valeurs prises par X avec les probabilités

$$P_i = \Pr(X = \alpha_i)$$

L'espérance mathématique de X , notée $E(X)$ est définie par (si la série converge) :

$$E(X) = \sum_{i=0}^n p_i \alpha_i$$

II.2.1.2 Variable aléatoire continue

L'espérance mathématique de la variable aléatoire X , définie sur l'espace probabilisé $(\Omega, \mathcal{C}, \Pr)$, est donnée par l'intégrale, si elle converge :

$$E(X) = \int_{\Omega} X d\Pr = \int_{\Omega} x \Pr_x dx$$

que l'on peut écrire, si f est la densité de probabilité de X :

$$E(X) = \int_{\Omega} x f(x) dx$$

II.2.1.3 Propriétés de l'espérance mathématique

X et Y sont deux variables aléatoires admettant chacune une espérance mathématique, a et b sont deux constantes :

$$E(a) = a$$

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

$$E(X + Y) = E(X) + E(Y)$$

Une variable aléatoire est dite centrée si son espérance mathématique est nulle : $E(X) = 0$

Remarques

- L'espérance mathématique peut être considérée comme le centre de gravité de la distribution de masses Pr_X .
- Il existe des variables aléatoires qui n'ont pas d'espérance mathématique.

II.2.2 Variance d'une variable aléatoire

II.2.2.1 Définition

La variance (moment centré d'ordre 2), ou carré de l'écart-type σ , est donnée par l'intégrale si cette intégrale et la densité f existent :

$$E[(X - E(X))^2] = \text{Var}(X) = \sigma^2 = \int_{\Omega} [x - E(X)]^2 f(x) dx$$

L'écart-type σ s'exprime avec la même unité que la variable.

$$\begin{aligned} E[(X - a)^2] &= E[(X - E(X) + E(X) - a)^2] \\ &= E[(X - E(X))^2] - 2[E(X) - a][E[X - E(X)]] + [E(X) - a]^2 \end{aligned}$$

ou, comme $E[X - E(X)] = 0$,

$$E[(X - a)^2] = \text{Var}(X) + [E(X) - a]^2$$

La borne inférieure de $E[(X - a)^2]$ est obtenue pour $a = E(X)$, elle est égale à $\text{Var}(X)$.

La variance est le plus petit moment d'ordre deux d'une variable aléatoire.

II.2.2.2 Quelques propriétés de la variance

- Pour $a = 0$, la formule précédente devient :

Cette formule peut être utilisée pour calculer rapidement $\text{Var}(X)$.

– Si a et b sont des constantes :

$$\text{Var}(b) = 0$$

CHAPITRE TROISIEME. DESCRIPTION DES DISTRIBUTIONS THEORIQUES

III.1 DISTRIBUTION ET LOIS DE PROBABILITE DISCRETES

III.1.1 Définition d'une variable discrète

Une variable aléatoire discrète prend ses valeurs sur un ensemble fini ou dénombrable de points. La loi de probabilité d'une telle variable est appelée loi discrète.

Une loi de probabilité discrète est caractérisée par l'énumération des valeurs x_i , appartenant à $\mathbb{R}$ ou à un intervalle de $\mathbb{R}$, prises par la variable aléatoire X et par les probabilités associées, c'est-à-dire les nombres réels positifs p_i tels que :

$$\Pr(X = x_i) = p_i \qquad 0 \leq p_i \leq 1 \qquad \sum_i p_i = 1$$

La fonction de répartition est une fonction en escalier, constante sur tout intervalle $[x_i, x_{i+1}[$, admettant en chaque point x_i un saut égal à $p_{i+1} = \Pr(X \leq x_{i+1})$.

III.1.1.1 Moments

$$E(X) = \sum_i x_i p_i \qquad \text{Var}(X) = \sum_i x_i^2 p_i - [E(X)]^2$$

Exemple

Soit X la variable aléatoire prenant trois valeurs 0, 1, 2 avec les probabilités :

$$\Pr(X=0) = 1/2 \qquad \Pr(X=1) = 1/3 \qquad \Pr(X=2) = 1/6$$

$$(1/2 + 1/3 + 1/6 = 1)$$

Espérance mathématique :

$$E(X) = 0 \cdot 1/2 + 1 \cdot 1/3 + 2 \cdot 1/6 = 2/3$$

Variance :

$$\text{Var}(X) = (\sigma^2) = E(X^2) - [E(X)]^2 = 1 \cdot 1/3 + 4 \cdot 1/6 - (2/3)^2 = 5/9$$

III.1.1.2 Domaine d'utilisation

Les lois discrètes sont utilisées pour décrire 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. Dans le cadre de ce cours nous allons présenter trois distributions.

III.1.2 Distribution binomiale ou loi des tirages avec remise

III.1.2.1 Définition et propriétés d'une variable de Bernoulli

On considère une expérience dont le résultat ne peut prendre que deux valeurs appelées, par convention, succès ou échec : un candidat est reçu ou non à un examen, une pièce usinée est bonne ou défectueuse, une porte est ouverte ou fermée...

À une expérience de ce type, est associée une variable aléatoire X prenant la valeur 1 pour le succès et la valeur 0 pour l'échec, avec les probabilités respectives p et $(1 - p) = q$. Cette variable est appelée variable de Bernoulli.

La loi de probabilité d'une variable de Bernoulli est définie par :

$$\Pr(X=1) = p$$

$$\Pr(X=0) = 1-p = q$$

Ses moments sont :

$$E(X) = p$$

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

Domaine d'utilisation : elle est utilisée pour décrire 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...

III.1.2.2 Définition d'une variable binomiale

On réalise n épreuves indépendantes de la même expérience telles que :

- chaque épreuve ne peut avoir que deux résultats, s'excluant mutuellement, soit le succès, soit l'échec,
- la probabilité p de succès est constante à chaque épreuve, la probabilité d'échec est également constante et égale à $1 - p = q$.

❖ Probabilité d'obtenir k succès au cours de ces n épreuves

Soit X la variable aléatoire qui « compte » le nombre de succès au cours de n épreuves.

- Si au cours de n épreuves, on obtient k succès, on obtient également $(n - k)$ échecs. La probabilité de réalisation d'un tel événement est égale à $p^k(1 - p)^{n-k}$ (les épreuves sont indépendantes).
- Il y a différentes façons d'obtenir k succès au cours de n épreuves, chaque épreuve pouvant être, indépendamment les unes des autres, un succès ou un échec. Le nombre de réalisations possibles de l'événement « obtenir k succès au cours de n épreuves » est le nombre de combinaisons sans répétitions de n objets pris k à k , soit :

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

D'où

$$\Pr(X=k) = C_n^k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

Cette expression étant un terme du développement du binôme $[p + (1-p)]^n$, la variable X est appelée variable binomiale. On vérifie facilement que :

$$\sum_{k=0}^n \Pr(X = k) = 1$$

La loi de la variable X est appelée loi binomiale de paramètres (n, p) , notée $B(n ; p)$.

Une variable binomiale est égale à la somme de n variables aléatoires indépendantes de Bernoulli, la loi binomiale est donc la loi des épreuves répétées. Elle est également la loi d'un tirage sans remise dans une urne contenant des boules de deux types différents.

❖ Moments

Une variable binomiale, suivant la loi $B(n ; p)$, peut être considérée comme la somme de n variables indépendantes de Bernoulli. D'où les résultats :

$$E(X) = np$$

$$\text{Var}(X) = np(1-p) = npq$$

❖ 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.

Exemple

On veut réaliser une étude clinique sur des malades se présentant à une consultation hospitalière. Pour cette étude, seuls les malades répondant à un ensemble de critères C sont retenus. Des statistiques antérieures ont montré que 20 % des consultants présentent les critères C.

10 malades viennent consulter le premier jour.

Soit X la variable aléatoire « nombre de malades retenus » c'est-à-dire répondant à l'ensemble des critères C. La variable X suit la loi binomiale $B(10 ; 0,20)$.

La probabilité qu'aucun malade ne soit recruté ce jour est égale à :

$$\Pr(X = 0) = C_{10}^0 (0,20)^0 (0,80)^{10} = 0,107$$

La probabilité pour qu'il y ait au moins un malade recruté est égale à :

$$\Pr(X \geq 1) = 1 - \Pr(X = 0) = 1 - 0,107 = 0,893$$

III.1.3 Distribution hypergéométrique ou loi du tirage exhaustif

III.1.3.1 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_p éléments (éléments possédant un certain caractère),

– une partie B à $(N - N_p)$ éléments (éléments n'ayant pas ce caractère).
C'est le cas, par exemple, d'un lot de N pièces comprenant N_p pièces défectueuses et donc $(N - N_p)$ 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 $\Pr(X = x)$.

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

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

La probabilité cherchée est égale à :

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

Le quotient n/N est appelé taux de sondage.

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

III.1.3.2 Moments

Une variable aléatoire X suivant une loi hypergéométrique $H(N ; n ; p)$ peut être considérée comme une somme de n variables aléatoires de Bernoulli,

$X_1 \dots X_n$, non indépendantes, correspondant aux tirages successifs de n individus, tirage sans remise. On en déduit le calcul des moments d'ordre 1 et 2.

❖ **Espérance mathématique :**

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

❖ **Variance**

Des calculs analogues conduisent au résultat suivant :

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

$\frac{N-n}{N-1}$ est le facteur d'exhaustivité

❖ **Remarques**

- Les variables aléatoires X_i sont des variables aléatoires de Bernoulli non indépendantes.
- 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 $B(n ; p)$.

❖ **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.

Valeurs extrêmes de cette variable :

- 5 (la valeur 0 ne peut pas être obtenue car il n'y a pas 15 femmes dans l'assemblée),
- 15 (dans l'échantillon, il n'y aura pas de femmes).

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

$$\Pr(X = 10) = \frac{C_{20}^{10} C_{10}^5}{C_{30}^{15}} = 0,30$$

III.1.4 Distribution ou loi de Poisson

III.1.4.1 Définition

La loi de Poisson de paramètre λ est la loi d'une variable aléatoire discrète réelle X , prenant toutes les valeurs entières non négatives, avec les probabilités :

$$\Pr(X = k) = p_k = \frac{e^{-\lambda} \lambda^k}{k!} \quad K \in [0, +\infty[\quad \sum_{k=0}^{\infty} p_k = 1$$

La loi de Poisson dépend d'un seul paramètre λ . On la note $P(\lambda)$.

❖ Moments

$$E(X) = \text{Var}(X) = \lambda$$

La loi de Poisson est la seule loi discrète possédant la propriété, $E(X) = \text{Var}(X)$

❖ Propriétés et domaine d'utilisation

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 est la loi limite de la loi binomiale quand n tend vers l'infini et p tend vers zéro, le produit np restant fini.

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 » :

$$\Pr(X = k) = \frac{e^{-0,05}(0,05)^k}{k!}$$

La probabilité que le système tombe en panne une fois au cours d'une semaine 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 $\mu = 0,05 \times 50 = 2,5$.

$$\Pr(Y = k) = \frac{e^{-2,5}(2,5)^k}{k!}$$

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.

Caractéristiques des principales lois discrètes

Loi	Formule	Application	Approximation
Hypergéométrique	$\Pr(X = x) = \frac{C_{Np}^x C_{N-Np}^{n-x}}{C_N^n}$	Tirage sans remise	Loi binomiale si $n < 0,1N$
Binomiale	$\Pr(X = x) = C_n^x p^x (1-p)^{n-x}$	Tirage avec remise Probabilité constante	Loi de Poisson si $n > 50$ et $p < 0,10$
Poisson	$\Pr(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$	Événements instantanés et indépendants	

III.2 DISTRIBUTION OU LOIS DE PROBABILITE CONTINUES

III.2.1 Généralités

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...

Il existe une fonction f non négative, définie pour toute valeur x appartenant à $\mathbb{R}$ et vérifiant, pour toute partie A de $\mathbb{R}$, la propriété :

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

Telle que

$$\int_{\mathbb{R}} f(x) \, dx = 1$$

La fonction f est la densité de probabilité de la variable aléatoire X .

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

$$F(a) = \Pr(X < a) = \int_{-\infty}^a f(x) \, dx$$

Pour toutes les valeurs a et b appartenant à $\mathbb{R}$, on a donc la relation :

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

III.2.2 Loi de Laplace-Gauss ou loi normale

III.2.2.1 Définition

Une variable aléatoire réelle X , prenant ses valeurs dans $\mathbb{R}$, suit une loi de Laplace-Gauss ou loi normale, de paramètres m et σ , si sa densité de probabilité est donnée par :

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

La fonction f définit une densité. En effet :

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

Cette loi est notée, en général $N(m ; \sigma)$. On dit indifféremment qu'une variable suivant une telle loi est une variable normale ou gaussienne.

Fonction de répartition :

$$\Pr(X < a) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

Cette intégrale n'ayant pas d'expression mathématique simple, des tables donnent les valeurs de la fonction de répartition.

Sur la courbe représentant la densité de probabilité d'une variable gaussienne, la valeur de $F(a)$ est représentée par la partie non hachurée. Cette courbe a un axe de symétrie vertical pour $x = m$ et du fait de sa forme, elle est souvent appelée « courbe en cloche ».

III.2.2.2 Variable aléatoire centrée réduite

La variable centrée réduite associée à la variable aléatoire X est la variable

$$: Z = \frac{X - \mu}{\sigma}$$

La variable Z suit la loi normale $N(0 ; 1)$ dont les paramètres sont $\bar{X} = 0$ et $\sigma = 1$.

Le changement de variables permettant de passer d'une variable à l'autre.

:

$$Z = \frac{X - \mu}{\sigma} \quad \text{et} \quad X = \sigma Z + \mu$$

❖ Exemple : Utilisation de la table de la loi normale

Soit X une variable suivant la loi normale $N(3 ; 2)$, donc de moyenne 3 et d'écart-type

2. On veut calculer les probabilités suivantes : $\Pr(X < 4)$, $\Pr(X < -1)$, $\Pr(X > 1)$ ou les nombres a_i tels que $\Pr(X < a_1) = 0,75$, $\Pr(X > a_2) = 0,85$.

On utilise la variable centrée réduite Z associée à la variable X :

$$Z = \frac{X - 3}{2} \quad \text{et} \quad X = 2Z + 3$$

$$a) \Pr(X < 4) = \Pr(2Z + 3 < 4) = \Pr(Z < 0,50) = 0,6915$$

$$b) \Pr(X < -1) = \Pr(2Z + 3 < -1) = \Pr(Z < -2) = 0,0228$$

$$c) \Pr(X > 1) = \Pr(2Z + 3 > 1) = \Pr(Z > -1) = 0,8413$$

$$d) \Pr(X < a_1) = \Pr(2Z + 3 < a_1) = \Pr(Z < \frac{a_1 - 3}{2}) = 0,75$$

$$\Pr(Z < 0,675) = 0,75 ; \text{ d'où : } a_1 = 4,35$$

$$e) \Pr(X > a_2) = \Pr(2Z + 3 > a_2) = \Pr(Z > \frac{a_2 - 3}{2}) = 0,85$$

$$\Pr(Z < -1,035) = 0,15 \quad \Pr(Z > -1,035) = 0,85$$

$$\text{D'où : } a_2 = -1,035 * 2 + 3 = 0,93$$

- Résultats remarquables :

$$\Pr (\mu - 1,64 \sigma < X < \mu + 1,64 \sigma) = 0,90$$

$$\Pr (\mu - 1,96 \sigma < X < \mu + 1,96 \sigma) = 0,95$$

$$\Pr (\mu - 3,09 \sigma < X < \mu + 3,09 \sigma) = 0,998$$

III.2.2.3 Domaine d'utilisation

- La loi normale est une des lois de probabilité la plus utilisée. Elle dépend de deux paramètres, la moyenne $\bar{X}$, paramètre de position, et l'écart-type σ , paramètre mesurant la dispersion de la variable aléatoire autour de sa moyenne.
- Elle s'applique à de nombreux phénomènes, en physique, en économie (erreurs de mesure). De plus, elle est la forme limite de nombreuses distributions discrètes...

III.2.3 Distribution ou loi de chi-deux ou chi-carré

Soient les variables aléatoires $X_1, X_2, \dots, X_n$ ayant respectivement pour moyennes et écart-type $\mu_1, \mu_2, \dots, \mu_n ; \sigma_1, \sigma_2, \dots, \sigma_n$. Ainsi, les variables centrées réduites sont

$$Z_1 = \frac{X_1 - \mu_1}{\sigma_1}$$

$$Z_2 = \frac{X_2 - \mu_2}{\sigma_2}$$

$$Z_n = \frac{X_n - \mu_n}{\sigma_n}$$

Alors, la variable aléatoire, égale à la somme des carrés de ces n variables aléatoires indépendantes, centrées, réduites, gaussiennes, suit la loi du chi-deux, χ^2 , à k degrés de liberté :

$$Z = \sum_{i=1}^n Z_i^2 \sim \chi^2(k)$$

❖ Propriétés

- La variable aléatoire **Z** varie de 0 à $+\infty$.
- Le paramètre **k** est le nombre de degrés de liberté de la variable, il représente la dimension de l'espace dans lequel se trouve le point représentatif de l'échantillon X. Si les variables aléatoires X_i vérifient k relations linéaires, le nombre de degrés de liberté diminue de k.
- La loi suivie par la somme de variables aléatoires indépendantes, suivant chacune des lois du chi-deux, est une loi du chi-deux dont le degré de liberté est la somme des degrés de liberté de chaque loi.
- La distribution du χ^2 est asymétrique à droite

III.2.4 Distribution ou loi de Student

La distribution t de Student, joue un rôle important dans l'étude de la statistique X pour une distribution normale dont on ne connaît pas la variance. La loi de Student est la loi de la variable aléatoire t définie par : soit Z_1 une variable aléatoire centrée réduite normale $N(0,1)$, et Z_2 une variable aléatoire qui suit la loi de chi-deux avec k degré de liberté, alors l'expression $t = \frac{Z_1}{\sqrt{Z_2/k}} = \frac{Z_1 \sqrt{k}}{\sqrt{Z_2}}$ suit une loi de student avec k degré de liberté.

La distribution de t de student comme distribution normale est symétrique par rapport à la valeur moyenne mais elle est plus aplatie que la distribution normale. Elle a une moyenne nulle et la variance égale à $\frac{k}{(k-2)}$.

Les différentes valeurs de t pour différents nombres de degré de liberté sont comprises à la table de student.

CHAPITRE QUATRIEME. ÉCHANTILLONNAGE ET DISTRIBUTION D'ÉCHANTILLONNAGE

IV.1 INTRODUCTION A LA STATISTIQUE INFÉRENTIELLE

IV.1.1 INTRODUCTION

L'étude descriptive des données se limite à un seul ensemble soit une population, soit un échantillon et n'établit pas de liaison entre les deux.

D'un autre côté, les rappels fondamentaux des probabilités (variables aléatoires, distributions, paramètres, convergences, etc.) nous ont confronté à l'aléatoire, avec notamment les subtils passages à la limite, les convergences qui conduisent au fondement de la statistique mathématique.

La statistique inférentielle, pont entre la statistique descriptive et la statistique mathématique, établit des relations entre populations et échantillons. On distingue deux types de démarche :

- la démarche d'échantillonnage (de la population vers l'échantillon)
- la démarche d'estimation (de l'échantillon vers la population).

IV.1.2 DÉMARCHE D'ÉCHANTILLONNAGE

La démarche d'échantillonnage est une démarche statistique classique de type déductif c'est à dire qui va du "général au particulier" : on connaît la population, on s'intéresse à l'échantillon.

Exemples :

- On connaît les professions d'une population cible dans laquelle est prélevé un échantillon. Est-ce que cet échantillon peut être considéré comme représentatif de la population selon la variable profession ?
- On s'intéresse au contrôle de la qualité de fabrication de tablettes de chocolat. Est-ce qu'on peut considérer comme constant le poids moyen

garanti d'une tablette ? Pour cela, on prélève régulièrement un échantillon de n tablettes dont l'étude statistique permettra de répondre à la question.

- Dans la fabrication d'aliment pour poulets conditionné en sacs de 10 kilos, on indique sur les sacs la composition de l'aliment (proportions des composants). Des échantillons sont prélevés sur les lieux de vente pour contrôler le respect de ces indications.

IV.1.3 DÉMARCHE D'ESTIMATION

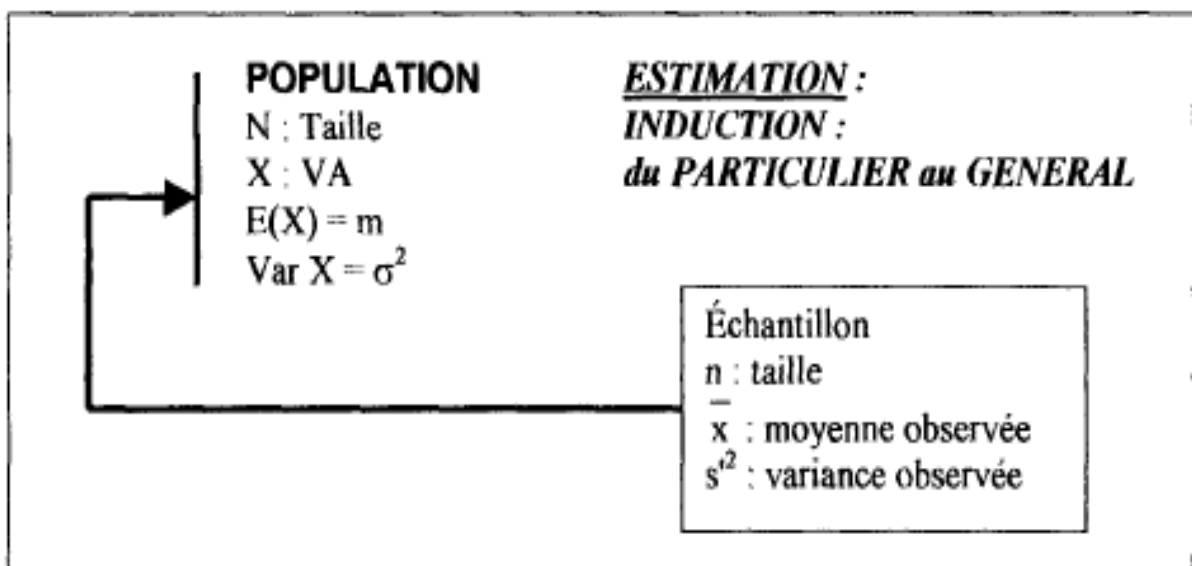
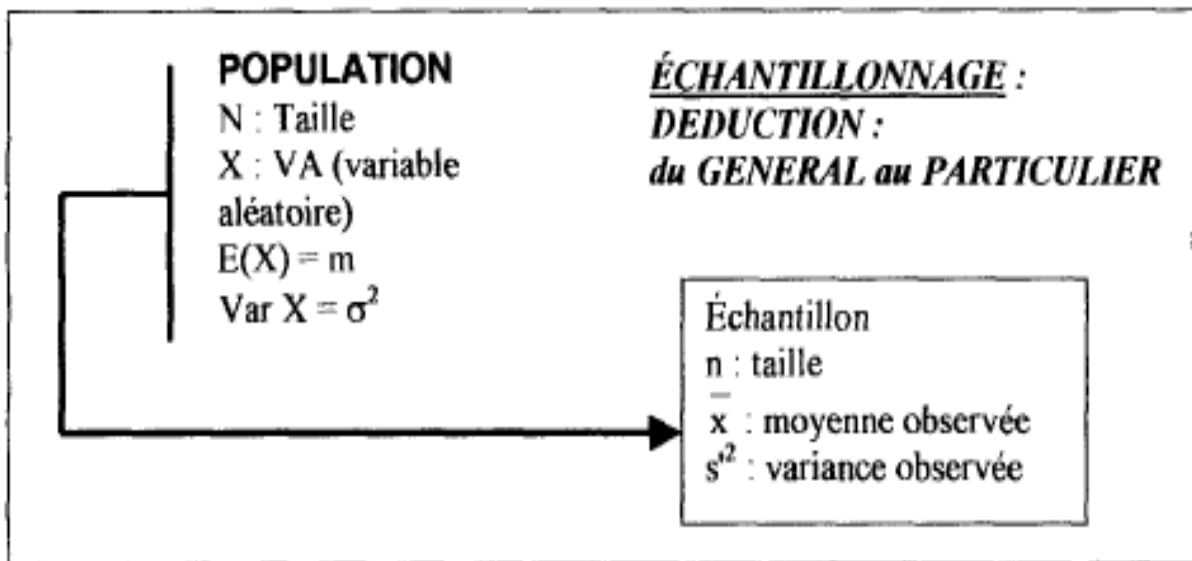
La démarche d'estimation, opposée à la précédente, vise à étudier, à prédire les paramètres d'une population inconnue à partir des résultats obtenus grâce à des échantillons.

C'est une démarche inductive "du particulier au général". Inférence est d'ailleurs synonyme d'induction, d'où le terme de statistique inférentielle même si dans la pratique ce qualificatif de la statistique a été élargi aux deux démarches.

Exemples :

- Avant des élections, des sondages sont effectués pour "estimer" les chances des candidats.
- Pour évaluer l'image d'un "produit" fourni par une société de services, on réalise un sondage auprès d'un échantillon de clients ; son analyse permet d'estimer l'indice de satisfaction moyen pour ce produit.

❖ RESUME

**IV.2 ÉCHANTILLONNAGE****IV.2.1. NOTION DE POPULATION ET D'ÉCHANTILLON**

Après avoir rappelé les notions fondamentales de "population" et d'"échantillon", nous définirons l'échantillon aléatoire et simple et son approche pratique en précisant nos choix de "grand" et "petit" échantillon.

Nous présenterons ensuite les concepts de base des distributions d'échantillonnage des moyennes, des variances et des proportions.

Le nombre d'applications concrètes sera volontairement limité car nous le développerons par la suite dans le cadre plus large des tests de conformité.

IV.2.1.1. Population

La "population" est l'ensemble des éléments auxquels on s'intéresse. Chaque élément est appelé "unité statistique" (u.s.) ou "individu" ou "observation".

La population peut être parfaitement définie (ensemble dénombrable fini) :

- ensemble des clients d'une banque
- ensemble des habitants d'une ville donnée ou d'un quartier donné
- ensemble des chevaux d'une région donnée
- ensemble des arbres d'un verger.

Pour de telles populations, l'étude statistique peut être parfaitement conduite sur l'intégralité de la population (petites populations, recensements, feuilles d'impôts)

La population peut également être non définie car infinie ou imparfaitement connue :

- ensemble des profils pédologiques (population infinie)
- ensemble des clients d'une grande surface d'une enseigne donnée
- ensemble des parasites d'une culture de blé.

IV.2.1.2. Échantillon

L'échantillon est une fraction d'individus de la population.

On peut échantillonner dans les domaines les plus divers : consommation, opinions, sociologie, contrôle de la qualité, etc.

❖ **Raisons de l'échantillonnage**(ou de sondage) :

- Le plus souvent, on réalise un sondage pour des raisons évidentes de gain de temps et de coût.
- Dans certains domaines, l'étude de l'unité statistique exige sa destruction.

Citons par exemple les questions de "durée de vie" (aliments, produits industriels tels les piles, les ampoules électriques, les CD, etc.).

D'autres domaines (psychologie, sociologie,...) nécessitent des études très approfondies. Il est alors impossible de les réaliser sur une population (excepté les populations cibles, bien limitées).

Pour une population de taille N si n est la taille de l'échantillon le nombre total d'échantillons dans un tirage non exhaustif est égal à N^n , et ce nombre pour le tirage exhaustif est égal à $\frac{N!}{n!(N-n)!}$. Le rapport par le quotient noté t entre la taille de l'échantillon n et celle de la population N est appelé taux de sondage ou bien fraction de sondage dont l'expression est $t = \frac{n}{N}$. On appelle base de sondage la liste complète des éléments d'une population mère. Cette dernière peut être finie ou infinie.

IV.2.1.3 TYPES DE SONDAGES

Il existe deux types de sondages à savoir les sondages probabilistes et les sondages non probabilistes. Effet, les sondages probabilistes sont ceux où seul un hasard connu détermine la composition de l'échantillon, c'est-à-dire la probabilité pour que chaque élément de la population (individu) soit sélectionné dans l'échantillon est connu d'avance. Ce type de sondage

nécessite l'obtention d'une base de sondage. On appelle sondages non probabilistes ceux où il n'y a pas de hasard dans la détermination de l'échantillon.

L'obtention d'un échantillon par choix aléatoire peut se faire avec différentes méthodes d'échantillonnage dont les plus importantes sont : l'échantillonnage simple, l'échantillonnage systématique, l'échantillonnage stratifié...

Par exemple pour l'**échantillon aléatoire simple** :

Selon la définition plus traditionnelle: "un échantillon est dit aléatoire quand tous les individus de la population ont une même probabilité de faire partie de l'échantillon et il est dit aléatoire et simple ou complètement aléatoire quand, en outre, les choix successifs des différents individus qui doivent constituer l'échantillon sont réalisés indépendamment les uns des autres au sens de l'indépendance stochastique".

IV.3 DISTRIBUTIONS D'ÉCHANTILLONNAGE

IV.3.1 Distribution d'échantillonnage des moyennes et des variances

Soit la population constituée de l'ensemble des N individus, on s'intéresse à la variable aléatoire X et, soit on prélève un premier échantillon de taille n (par exemple 25). Pour chacun de ces n individus, on relève la variable aléatoire. On dispose alors d'une série statistique $x_{11}, x_{12}, \dots, x_{1n}$ de moyenne $\bar{x}_1$ et de variance s_1^2 .

Echantillons	Valeurs observées	Moyennes observées	Variances observées (empiriques)
E_1	$X_{11}, X_{12}, \dots, X_{1n}$	$\bar{X}_1$	S_1^2
E_2	$X_{21}, X_{22}, \dots, X_{2n}$	$\bar{X}_2$	S_2^2
....			
E_p	$X_{p1}, X_{p2}, \dots, X_{pn}$	$\bar{X}_p$	S_p^2
Variables aléatoires	$X_1 \ X_2 \ \dots \ X_n$	$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$	$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$

Les premières valeurs observées dans chaque échantillon (x_{pi} , où p est le numéro de l'échantillon) sont aléatoires et constituent par conséquent les réalisations d'une variable aléatoire X_1 . Un échantillon aléatoire et simple de taille n est équivalent à un ensemble de n variables aléatoires $X_1, X_2, \dots, X_n$ indépendantes.

La même traduction est faite au niveau des moyennes et des variables. Chaque moyenne observée dans un échantillon est l'observation d'une variable aléatoire Moyenne : $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$.

Chaque variance observée dans un échantillon est l'observation d'une variable aléatoire variance : $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$.

Les distributions des variables aléatoires $\bar{X}$ (ou $\mu_{\bar{X}}$) et S^2 (ou $\sigma_{\bar{X}}^2$) sont dites distributions d'échantillonnage des moyennes et des variances.

Il existe une relation entre la variance de la population et les variances des échantillons.

Lorsqu'on prend tous les échantillons de taille n , échantillons provenant d'une population de taille N et ayant pour moyenne μ et un écart-type σ ; on calcule la moyenne de chaque échantillon, l'ensemble de ces moyennes constitue la distribution d'échantillonnage de la moyenne. A partir du calcul de moyenne de la distribution d'échantillonnage de la moyenne $\mu_{\bar{x}}$ et la variance de distribution de l'échantillonnage de la moyenne $\sigma_{\bar{x}}^2$, on établit les relations suivantes :

▪ **Pour les échantillons non exhaustifs :**

$$E(\bar{X}) = \mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \rightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

▪ **Pour les échantillons exhaustifs :**

$$E(\bar{X}) = \mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \rightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

IV.3.2 Distribution d'échantillonnage des proportions

Supposons qu'une certaine population soit infinie et que la probabilité de réalisation des événements soit p , la probabilité de non réalisation est $q = 1 - p$.

Considérons maintenant tous les échantillons extraits de cette population et pour chaque échantillon déterminons la proportion p de succès. On obtient alors une distribution des fréquences dont la moyenne et l'écart-type sont respectivement :

▪ **Pour les échantillons non exhaustifs :**

$$\mu_p = p$$

$$\sigma_p = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}$$

▪ **Pour les échantillons exhaustifs :**

$$\mu_p = p$$

$$\sigma_p = \sqrt{\frac{pq}{n}} \left(\sqrt{\frac{N-n}{N-1}} \right) = \sqrt{\frac{p(1-p)}{n}} \left(\sqrt{\frac{N-n}{N-1}} \right)$$

CHAPITRE CINQUIEME. ESTIMATION STATISTIQUE

V.1 INTRODUCTION

La notion d'estimation a été présentée lors de l'introduction de la statistique inférentielle. Rappelons que sa mission essentielle est d'obtenir une valeur approchée d'un ou plusieurs paramètres statistiques d'une variable aléatoire d'une population à partir des données observées dans un échantillon.

Le but de la théorie de l'estimation est de choisir, parmi toutes les statistiques possibles, le meilleur estimateur, c'est-à-dire celui qui donnera une estimation ponctuelle la plus proche possible du paramètre et ceci, quel que soit l'échantillon.

Par exemples :

- l'estimation du chiffre d'affaires annuel moyen de carton de savon le coq ;
- l'estimation du pourcentage de congolais de plus de cinquante ans présentant un taux de cholestérol trop élevé ;
- l'estimation du prix de vente annuel moyen d'un litre de miel.

5.2. ESTIMATION PONCTUELLE

Un aspect important de l'inférence statistique consiste à obtenir des estimations fiables des caractéristiques d'une population à partir d'un échantillon extrait de cette population. C'est un problème de décision concernant des paramètres tels que :

- l'espérance mathématique notée m ou μ (pour un caractère mesurable),
- la variance ou l'écart-type notée σ ,

– la proportion p (pour un caractère dénombrable).

Comme un échantillon ne peut donner qu'une information partielle sur la population, les estimations ainsi obtenues seront inévitablement entachées d'erreurs que l'on doit minimiser autant que possible.

Bref : Estimer un paramètre, c'est donner une valeur approchée de ce paramètre, à partir des résultats obtenus sur un échantillon aléatoire extrait de la population.

Exemple 1

On veut étudier une caractéristique X d'un phénomène économique, par exemple la proportion p des individus d'une population P d'effectif N , présentant un certain caractère : posséder un ordinateur, avoir reçu un crédit de la banque au moins deux ans passés, etc.

Pour obtenir la valeur exacte de p , il suffirait d'interroger les N individus de la population ce qui, en général, est impossible ; on interroge donc les individus d'un échantillon aléatoire, de taille n , représentatif de la population. On obtient un ensemble de n valeurs x_i telles que : $X_i = 1$ si l'individu i présente le caractère X , $x_i = 0$ sinon.

Soit f_n la proportion des individus ayant le caractère X dans l'échantillon. Cette proportion f_n converge vers p quand n tend vers l'infini (loi des grands nombres).

On estimera donc p par f_n . C'est une estimation ponctuelle qui donne une seule valeur pour p .

Par un choix adéquat de n , on peut essayer de minimiser l'erreur due à cette approximation.

Exemple 2

Un autre type de problème relevant de la théorie de l'estimation consiste en l'estimation

des paramètres de la loi suivie par une variable aléatoire X , loi dont on connaît la forme.

Ainsi, par exemple, le nombre d'accidents dans un atelier pendant une semaine suit probablement une loi de Poisson. Cette loi dépend d'un paramètre λ dont on ne connaît pas la valeur. Pour donner une valeur à ce paramètre, on note le nombre d'accidents survenus pendant n semaines, c'est l'échantillon aléatoire de taille n .

Le nombre moyen d'accidents est une estimation ponctuelle du paramètre λ (propriété de la loi de Poisson).

En général les estimations peuvent revêtir deux formes :

- soit une valeur unique, l'estimation ponctuelle, ou valeur la plus probable que prendra le paramètre,
- soit un ensemble de valeurs appartenant à un intervalle, l'estimation par intervalle de confiance. Un intervalle de confiance doit avoir de « grandes chances » de contenir la vraie valeur du paramètre, il est toujours associé à un risque d'erreur α . La notion de l'estimation fait intervenir des fonctions ou statistiques particulières, appelées estimateurs :

❖ Estimateur de l'espérance mathématique

La statistique X déjà étudiée est un estimateur sans biais pour l'espérance mathématique $E(X)$. En effet, $E(\bar{X}) = E(X)$.

❖ Estimateur de la variance

La statistique $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ est un estimateur biaisé pour la variance : $E(S^2) = \frac{n-1}{n} \sigma^2$ le biais est égal à $\frac{\sigma^2}{n}$.

En revanche, la statistique S^{*2} définie par $S^{*2} = \frac{n}{n-1} S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ est un estimateur sans biais pour la variance. En effet : $E(S^{*2}) = \frac{n}{n-1} E(S^2) = \sigma^2$

La fréquence f_n est un estimateur sans biais du paramètre p . De plus, sa variance tend vers 0 quand n tend vers l'infini, il est donc convergent.

En résumé les estimateurs convergents et sans biais sont :

- d'une moyenne $E(X)$: la statistique : $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$
- d'une variance : la statistique : $S^{*2} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

V.3 ESTIMATION PAR INTERVALLE DE CONFIANCE

L'estimation ponctuelle d'un paramètre θ donne une valeur numérique unique à ce paramètre, mais n'apporte aucune information sur la précision des résultats, c'est-à-dire qu'elle ne tient pas compte des erreurs dues aux fluctuations d'échantillonnage, par exemple.

Pour évaluer la confiance que l'on peut avoir en une estimation, il est nécessaire de lui associer un intervalle qui contient, avec une certaine probabilité, la vraie valeur du paramètre, c'est l'estimation par intervalle de confiance.

L'estimation par intervalle de confiance d'un paramètre θ consiste donc à associer à un échantillon, un intervalle aléatoire I , choisi de telle façon que la probabilité pour qu'il contienne la valeur inconnue du paramètre

soit égale à un nombre fixé à l'avance, aussi grand que l'on veut. On écrit : $\Pr(\Theta \in I) = 1 - \alpha$

$(1 - \alpha)$ est la probabilité associée à l'intervalle d'encadrer la vraie valeur du paramètre, c'est le seuil de confiance ou la quasi-certitude.

V.3.1 Estimation et intervalle de confiance de la moyenne

L'estimateur sans biais de la moyenne μ est la statistique $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ qui suit la loi normale, $N(\mu, \sigma/\sqrt{n})$.

Deux cas sont à distinguer, selon que l'écart-type est connu ou estimé.

Cas 1 : l'écart-type σ est connu

Étant donné un seuil α , l'intervalle de confiance pour la moyenne μ :

$$\Pr(\bar{x} - Z_c \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_c \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

N.B : dans le cas d'un échantillonnage exhaustif on multiplie l'erreur standard par le facteur d'exhaustivité

Cas 2 : l'écart-type s n'est pas connu

La variable aléatoire :

$T(n - 1) = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{\bar{x} - \mu}{s/\sqrt{n-1}}$ suit une loi de Student à $(n - 1)$ degrés de liberté.

Le seuil α étant donné, on lit sur la table (loi de Student) le nombre $t_{1-\alpha/2}(n - 1)$ tel que :

