

## Chapitre 8

# PROBABILITÉS ET STATISTIQUES

### 8.1 Notion d'expérience, d'événement et conséquences

#### 8.1.1 *Qu'est-ce qu'une expérience ?*

Toute expérience scientifique doit être menée de façon rigoureuse en transmettant toutes les informations nécessaires pour que toute personne s'intéressant au sujet de l'expérience puisse en vérifier les résultats.

Cette transmission des informations est nommée "**protocole**" de l'expérience, c'est-à-dire la description de l'ensemble des matériels et des méthodes employés pour mener à bien l'expérience.

Lorsqu'elle ne dépend pas du temps ou n'est pas perturbée par la mesure, une expérience est généralement reproductible (par exemple, une mesure de distances, de tailles, de poids, ...). Mais il existe aussi des expériences non reproductibles (par exemple, des mesures dépendantes du temps, des caractéristiques météorologiques, transmission d'un gène par hérédité, ...).

Théoriquement, il est possible de définir deux sortes d'expériences :

- celles qui sont dites déterministes, qui suivent une loi bien précise (loi de physique par exemple) et dont les résultats sont souvent prévisibles.
- les expériences aléatoires, dont les résultats sont par définition imprévisibles. La plus connue est le résultat d'un jet de dé ou d'un jeu de "pile ou face", mais cela peut être aussi la réaction d'un individu au traitement de sa maladie ou le nombre de fruits produits par un arbre, etc.

Cette dichotomie est purement théorique. Une expérience déterministe n'existe qu'en pensée : pour mesurer une distance, l'instrument n'est jamais parfait, il peut donner des résultats légèrement différents suivant la température, le taux d'humidité ; la manière dont l'expérimentateur l'utilise peut induire aussi des résultats différents. L'ensemble des résultats se répartissent "aléatoirement" autour de la loi que l'expérience illustre. Intervenient alors les notions d'incertitudes et d'erreurs que nous avons déjà décrites lors du chapitre sur les fonctions. Mais, ici, nous allons un peu plus loin en élaborant des moyens de décrire comment les mesures se répartissent au sein de ces incertitudes et comment estimer à partir des mesures obtenues la valeur recherchée par l'expérience.

Inversement, nous avons des moyens mathématiques de décrire les résultats d'une expérience aléatoire. Cela est encore plus vrai lorsque l'expérience est reproduite un grand nombre de fois.

En résumé, toute mesure issue d'une expérience comporte une part plus ou moins grande d'aléatoire. L'objet de ce chapitre est d'apprendre des outils mathématiques qui permettent d'utiliser ces mesures pour en déduire les résultats recherchés.

Mais avant tout, nous établissons quelques définitions.

- On appelle une **"épreuve"**, le protocole d'une expérience dont le résultat est aléatoire.
- Chaque résultat d'une épreuve définit un **"événement élémentaire"**.
- L'ensemble des événements élémentaires constitue l'"univers" que nous écrivons  $[E]$ .

L'univers est ainsi l'ensemble de toutes les mesures possibles.

Exemple : pour un jet de dé :  $[E] = \{1,2,3,4,5,6\}$

- On appelle **"événement"** un sous-ensemble de  $[E]$ . Il est composé d'un ou plusieurs événements élémentaires. On note  $[\Omega]$  l'ensemble des événements.

Exemple : On considère l'événement A pour un jet de dé consistant à obtenir un chiffre pair.

$A = \{2,4,6\}$  A est composé de trois événements élémentaires.

L'univers des événements devient, dans ce cadre :

$[\Omega] = \{\text{chiffres pairs, chiffres impairs}\}$

### 8.1.2 Opérations sur les événements

Ce paragraphe nécessite, dans un premier temps, de développer une liste de définitions amenant un certain nombre de symboles de logique. Il convient de s'habituer à cette écriture qui permet d'alléger les notations.

Soient deux événements A et B correspondant à une même épreuve. Ils peuvent être élémentaires ou non.

- L'événement **{A ou B}** est réalisé si ou A, ou B ou les deux simultanément sont réalisés. On le note  $A \cup B$ .
- L'événement **{A et B}** est réalisé si A et B sont réalisés à la fois. On le note  $A \cap B$ . On remarquera que l'événement  $A \cap B$  est inclus dans l'événement  $A \cup B$  et donc que si  $A \cap B$  est réalisé alors  $A \cup B$  est réalisé (l'inverse n'étant pas systématiquement vrai).
- L'événement **"impossible"**  $\emptyset$  est celui qui ne peut pas être réalisé à la suite de l'épreuve.
- Un événement **"certain"** est celui qui est réalisé à chaque épreuve. Il contient tous les événements élémentaires et est donc identique à l'univers  $[E]$ .
- A et B sont des événements **"incompatibles"** lorsque, en effectuant l'épreuve, il est impossible de réaliser à la fois A et B. Il n'y a pas d'élément élémentaire commun à A et à B. A et B sont disjoints :  $A \cap B = \emptyset$
- L'événement **"contraire"** de A est l'événement réalisé lorsque A ne l'est pas. La liste d'événements élémentaires qui le caractérise est composée des événement de  $[E]$  n'appartenant pas à la liste de A. On l'appelle aussi parfois non-A, noté  $\bar{A}$  ou  $\overline{A}$

## 8.2 Probabilités

### 8.2.1 Définition

La notion de probabilité est liée à la notion de fréquence. Le plus simple pour les introduire est sans doute de le faire à partir d'un exemple de répétition d'une expérience aléatoire.

Exemple. La mesure d'une longueur est répétée de nombreuses fois. On a obtenu la liste des 18 résultats fiables (appelés  $\{g_k\}$ ) suivante  $\{10,4 \text{ m}, 10,7 \text{ m}, 10,9 \text{ m}, 10,9 \text{ m}, 11 \text{ m}, 11 \text{ m}, 11 \text{ m}, 11,1 \text{ m}, 11,1 \text{ m}, 11,1 \text{ m}, 11,1 \text{ m}, 11,2 \text{ m}, 11,2 \text{ m}, 11,3 \text{ m}, 11,3 \text{ m}, 11,5 \text{ m}, 11,7 \text{ m}, 12 \text{ m}\}$ . Les résultats sont représentés ci-dessous sur l'axe des  $G$  où un point représente un résultat.

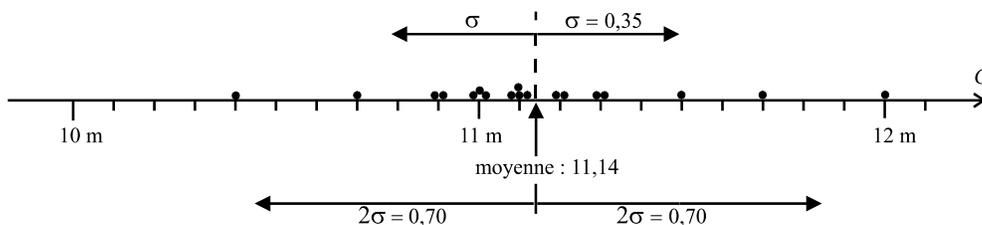


fig. 8.1 : Résultats de la mesure d'une longueur répétée 18 fois.

La valeur  $g = 11m$  y apparaît trois fois, on dit que sa "fréquence" est  $3/18 \simeq 0,17$  (on l'appelle parfois "fréquence relative" par opposition à la "fréquence absolue" qui est le nombre d'observation soit ici 3)

Nous devons distinguer clairement la liste des  $\{g_k\}$  où apparaissent des répétitions et la liste des valeurs différentes prises par les  $g_k$ . Le tableau ci-dessous donne les valeurs explicites des dix-huit  $g_k$

10,4 m	10,7 m	10,9 m	11 m	11,1 m	11,2 m	11,3 m	11,5 m	11,7 m	12 m
$g_1$	$g_2$	$g_3$ et $g_4$	$g_5$ à $g_7$	$g_8$ à $g_{11}$	$g_{12}$ et $g_{13}$	$g_{14}$ et $g_{15}$	$g_{16}$	$g_{17}$	$g_{18}$

Maintenant augmentons le nombre de mesures. Nous obtenons le tableau suivant

Nombre $n$ de mesures	Nombre $k$ de fois que $g = 11m$ a été obtenu	Fréquence $f = \frac{k}{n}$
18	3	$3/18 \approx 0,17$
100	17	$17/100 = 0,17$
1000	165	$165/1000 = 0,165$
10000	1662	$1662/10000 = 0,1662$

Nous pouvons remarquer que lorsque le nombre de mesures augmente, la fréquence se stabilise et tend vers la valeur limite de  $1/6$ . On appelle "probabilité objective", issue de l'expérience, cette valeur limite de la fréquence. C'est le résultat de la loi des grands nombres pour les expériences répétées de façon identique.

$$p = \lim_{n \rightarrow \infty} \left( f = \frac{k}{n} \right)$$

Il existe une définition mathématique axiomatique (c'est-à-dire non démontrable) de la probabilité :

- 1- La "probabilité"  $P(A)$  de tout événement  $A$  associé à une épreuve est un nombre compris entre 0 et 1.  
 $0 \leq P(A) \leq 1$
- 2- Si deux événements  $A$  et  $B$  sont incompatibles, la probabilité de l'événement  $(A \cup B)$  est égale à la somme des probabilités de  $A$  et de  $B$ .  
 $A \cap B = \emptyset$  alors  $P(A \cup B) = P(A) + P(B)$
- 3- La probabilité d'un événement certain est égale à 1.  
 $P(\Omega) = 1$

On en déduit que la probabilité d'un événement impossible est égale à 0.

$$P(\emptyset) = 0$$

Démonstration :  $P(\emptyset \cup \Omega) = P(\Omega) = P(\Omega) + P(\emptyset) \implies P(\emptyset) = 0$

Par ailleurs, l'axiome 2 peut se généraliser au cas de plusieurs événements incompatibles :

Soient  $A_1, A_2, \dots, A_n$  événements incompatibles 2 à 2

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

Dans le cas où une épreuve mène à un nombre fini  $n$  d'événements élémentaires pouvant être considérés comme symétriques sur le plan de leur réalisation, la probabilité de chacun d'eux est la même et vaut  $\frac{1}{n}$ . On dit alors que les événements sont "*équiprobables*".

Exemple : Les lancers de dé non-pipé sont équiprobables et la probabilité d'obtenir une valeur particulière est  $\frac{1}{6}$ .

**Remarque : Tout calcul conduisant à des valeurs de probabilités négatives ou supérieures à 1 est faux.**

### 8.2.2 Probabilités conditionnelles et combinées

– La probabilité du contraire d'un événement est égale à 1 moins la probabilité de cet événement

$$P(\bar{A}) = 1 - P(A)$$

Démonstration : Un événement et son contraire sont incompatibles :  $A \cap \bar{A} = \emptyset$

$$P(A) + P(\bar{A}) = P(A \cup \bar{A}) = P(\Omega) = 1$$

– Si un événement  $A$  est inclus dans un événement  $B$  (ce qui se note  $A \subset B$ ), alors la probabilité de  $A$  est inférieure à celle de  $B$ .

Démonstration : Si  $A$  est inclus dans  $B$ ,  $B$  peut se décomposer en deux événements incompatibles :  $A$  et  $(\bar{A} \cap B)$ . On a donc

$$\begin{aligned} P(B) &= P(\bar{A} \cap B) + P(A) \\ P(\bar{A} \cap B) &\geq 0 \implies P(B) \geq P(A) \end{aligned}$$

– La probabilité de l'événement  $A \cup B$  est égale à la somme des probabilités de  $A$  et de  $B$ , moins celle de  $(A \cap B)$ .

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Démonstration :  $A$  peut se décomposer en deux événements incompatibles :  $(A \cap B)$  et  $(A \cap \bar{B})$

$$P(A) = P(A \cap B) + P(A \cap \bar{B})$$

$B$  peut se décomposer en deux événements incompatibles :  $(A \cap B)$  et  $(\bar{A} \cap B)$

$$\begin{aligned} P(B) &= P(A \cap B) + P(\bar{A} \cap B) \\ \implies P(A) + P(B) &= 2 P(A \cap B) + P(\bar{A} \cap B) + P(A \cap \bar{B}) \end{aligned}$$

Or,  $A \cup B$  peut se décomposer en trois événements incompatibles :  $(A \cap \bar{B})$ ,  $(A \cap B)$  et  $(\bar{A} \cap B)$

$$\begin{aligned} P(A \cup B) &= P(A \cap B) + P(\bar{A} \cap B) + P(A \cap \bar{B}) \\ \implies P(A \cup B) &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Exemple : Pour le tirage d'une carte dans un jeu de 32 cartes, nous recherchons la probabilité d'obtenir un trèfle ou un roi.

On suppose l'équiprobabilité des cartes.

$$P(\text{trèfle}) = \frac{8}{32} = \frac{1}{4} \quad P(\text{roi}) = \frac{4}{32} = \frac{1}{8} \quad P(\text{roi} \cap \text{trèfle}) = P(\text{roi de trèfle}) = \frac{1}{32}$$

Les événements ne sont donc pas incompatibles. Nous avons donc

$$P(\text{roi ou trèfle}) = P(\text{roi} \cup \text{trèfle}) = P(\text{roi}) + P(\text{trèfle}) - P(\text{roi} \cap \text{trèfle})$$

Il faut soustraire  $P(\text{roi} \cap \text{trèfle})$  car sinon le roi de trèfle est compté à la fois dans la probabilité de tirer un roi et dans celle de tirer un trèfle.

$$P(\text{roi ou trèfle}) = \frac{1}{4} + \frac{1}{8} - \frac{1}{32} = \frac{8 + 4 - 1}{32} = \frac{11}{32}$$

- Considérons deux événements A et B, au cours d'une expérience aléatoire. La probabilité de réaliser A sachant que B est réalisé s'appelle "**probabilité conditionnelle**" de A sachant B et s'écrit  $P(A/B)$ . Elle est parfois appelée probabilité "**a posteriori**".

On remarquera que :  $P(\bar{A}/B) = 1 - P(A/B)$

- Il peut arriver que l'information apportée par la réalisation ou la non-réalisation de B ne modifie pas la probabilité de réalisation de A soit  $P(A)$

$$P(A/B) = P(A/\bar{B}) = P(A)$$

On dit que A et B sont "**indépendants**", alors

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

*Attention* : Il ne faut pas confondre indépendance et incompatibilité. Si  $P(A \cap B) \neq 0$ , les deux événements sont compatibles, mais ils peuvent être dépendants ou indépendants.

Exemple : Soient les deux événements :

A = aimer lire

B = aimer le chocolat

A et B sont compatibles (on peut à la fois aimer lire et le chocolat), mais A et B sont indépendants (le fait d'aimer lire n'apporte aucune information sur le fait d'aimer le chocolat et réciproquement).

- De façon plus générale, si l'on considère deux événements A et B quelconques, il est possible de calculer la probabilité d'avoir A et B

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

Exemple : Sur une population de 100 enfants, une enquête aboutit aux résultats suivants :

- 1- 87 enfants aiment le chocolat.
- 2- 63 enfants aiment lire.
- 3- Parmi les 63 enfants qui aiment lire, 57 aiment le chocolat.

Prenons un enfant au hasard dans cette population. Nous voulons déterminer les probabilités pour que cet enfant corresponde aux trois événements suivants :

- A : l'enfant aime le chocolat
- B : l'enfant aime lire
- $A \cap B$  : l'enfant aime lire et manger du chocolat

D'après les résultats de l'enquête, nous déduisons :

$$P(A) = 87/100 \quad P(B) = 63/100 \quad P(A \cap B) = 57/100$$

Si l'on sait au départ que l'enfant aime lire, alors l'ensemble de référence est restreint à 63 enfants. La probabilité que l'enfant aime le chocolat sachant qu'il aime lire est donc

$$P(A/B) = 57/63$$

Or, en utilisant le théorème précédent, nous avons

$$P(A/B) = P(A \cap B) / P(B) = (57/100) / (63/100) = 57/63$$

Nous retrouvons bien le résultat que nous avons obtenu directement.

De la même façon, la probabilité que l'enfant aime lire s'il aime le chocolat est

$$P(B/A) = 57/87 = (57/100) / (87/100) = P(A \cap B) / P(A)$$

### 8.2.3 Rappels d'analyse combinatoire

Les notions présentées ici sont des révisions du lycée. Elles sont donc présentées rapidement comme aide-mémoire pour faciliter la résolution des exercices. Cette partie permet aussi d'évaluer la nécessité (ou non) de revoir ces notions dans des documents issus du secondaire.

- *Notion de factorielle* : Si  $n \in \mathbb{N}$

$$n! = n (n-1) (n-2) \dots 1$$

Par convention :  $0! = 1$

- *Permutations de  $n$  objets discernables* : toute disposition ordonnée des  $n$  ( $\in \mathbb{N}$ ) objets.

$P_n$  nombre de permutations d'un ensemble de  $n$  éléments discernables :  $P_n = n!$

- *Arrangements de  $p$  éléments parmi  $n$  ( $p \leq n$ )* : toute disposition ordonnée sans répétition (sans remise, exhaustif) de  $p$  ( $\in \mathbb{N}$ ) objets parmi  $n$  ( $\in \mathbb{N}$ ) objets discernables.

$A_n^p$  nombre d'arrangements de  $p$  éléments pris parmi  $n$  :  $A_n^p = \frac{n!}{(n-p)!}$

- *Combinaisons de  $p$  éléments parmi  $n$  ( $p \leq n$ )* : toute disposition non-ordonnée de  $p$  ( $\in \mathbb{N}$ ) objets pris parmi  $n$  ( $\in \mathbb{N}$ ) objets discernables, chaque objet ne pouvant intervenir qu'une fois.

$C_n^p$  nombre de combinaisons de  $p$  éléments pris parmi  $n$  :  $C_n^p = \frac{n!}{p!(n-p)!}$

Propriétés des  $C_n^p$  :

- Si  $n \geq 1$   $C_n^0 = 1$   $C_n^n = 1$   $C_n^1 = n$

-  $C_n^p = C_n^{n-p}$

- Si  $1 \leq p \leq n$   $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$

- Si  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}$   $(a+b)^n = \sum_{p=0}^n C_n^p a^{n-p} b^p$   
 $= C_n^0 a^n + C_n^1 a^{n-1} b + \dots + C_n^{n-1} a b^{n-1} + C_n^n b^n$

### 8.3 Variables aléatoires discrètes

#### 8.3.1 Définitions

Une variable "**aléatoire**" est une épreuve menant à des événements élémentaires qui sont des nombres.

Par convention, nous allons désigner la variable aléatoire par une majuscule et le résultat de la mesure de cette variable aléatoire par une minuscule.

Ainsi, une variable aléatoire  $X$  est définie dans son univers  $\Omega$ , en associant à chaque résultat de l'expérience aléatoire un nombre réel  $x$  caractéristique de ce résultat. L'événement  $X = x$  prend en compte tous les résultats pour lesquels la variable  $X$  (en majuscule) prend la valeur  $x$  (en minuscule).

Une variable aléatoire  $X$  est dite "**discrète**", si l'ensemble des réalisations possibles  $x_1, x_2, \dots, x_n$  est fini ou dénombrable.

Dans le cas d'une variable discrète, la probabilité, associée à la réalisation de l'événement  $X = x$ , s'écrit  $P(X = x)$

*Remarque* : Soit  $g$  une fonction mathématique quelconque. Si l'on pose  $Y = g(X)$ , on définit une nouvelle variable aléatoire  $Y$  fonction de  $X$ .

La "**loi de probabilité**" d'une variable aléatoire  $X$  discrète est définie en donnant l'ensemble des valeurs  $p_1, p_2, \dots, p_n, \dots$  correspondant aux probabilités des différentes éventualités  $x_1, x_2, \dots, x_n, \dots$

On a donc  $p_i = P(X = x_i)$  avec  $0 \leq p_i \leq 1$  et  $\sum_i p_i = 1$

Exemple : Dans le cas d'un jet de dé, on a :  $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$

La "**fonction de répartition**" associée à la variable aléatoire  $X$  est la fonction, notée  $F$  ou  $F_X$ , définie dans  $\mathbb{R}$  dans l'intervalle  $[0,1]$  par

$$F(x) = F_X(x) = P(X \leq x) = \sum_{x_i \leq x} P(X=x_i)$$

On peut montrer aisément que :

$$0 \leq F(x) \leq 1$$

$F(x)$  est croissante : si  $a \leq b$  alors  $F(a) \leq F(b)$

$$\lim_{x \rightarrow -\infty} F(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow \infty} F(x) = 1$$

### 8.3.2 Paramètres caractéristiques

Soient une variable aléatoire  $X$  et sa loi de probabilité  $P(X = x)$

- La **"moyenne"**  $\bar{X}$  de la variable aléatoire  $X$  est la moyenne des résultats qu'on obtiendrait en répétant indéfiniment l'épreuve associée à  $X$ .

$$\bar{X} = x_1 \times p_1 + x_2 \times p_2 + \dots + x_n \times p_n = \sum_{i=1}^n x_i \times p_i$$

Exemple : La moyenne des tirs d'un dé est :

$$1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3,5$$

- L'**"espérance mathématique"** d'une variable aléatoire  $X$  est notée  $E(X)$ . C'est une question de vocabulaire. Cette expression est utilisée en calcul des statistiques et probabilités comme un synonyme de moyenne. Pour être plus précis, l'espérance mathématique désigne la moyenne issue de la loi de probabilité alors que, quand on parle de la moyenne, on peut tout aussi bien désigner la moyenne "théorique" issue de la loi de probabilité que la moyenne "expérimentale" des résultats obtenus pendant l'expérience.

$$\bar{X} = E(X)$$

L'espérance mathématique, et donc la moyenne, présente quelques propriétés.

Soient  $X$  et  $Y$  deux variables aléatoires

$$E(aX+b) = a E(X) + b \quad \text{avec } a \in \mathbb{R} \text{ et } b \in \mathbb{R}$$

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

Si, de plus,  $X$  et  $Y$  sont des variables aléatoires indépendantes

$$E(X \times Y) = E(X) \times E(Y)$$

- La **"variance"**  $\sigma^2(X)$  de la variable aléatoire  $X$  est la moyenne des carrés des écarts entre  $X$  et sa valeur moyenne  $E(X)$ .

$$\sigma^2 = E[(X-E(X))^2] = \sum_i p_i (x_i - \bar{X})^2 = E(X^2) - E(X)^2 = E(X^2) - \bar{X}^2$$

L'**"écart-type"**  $\sigma(X)$  est la racine carrée de la variance.

*Attention* : **La moyenne de  $X^2$  n'est pas le carré de la moyenne de  $X$ .**

$$\text{Démonstration} : E(X^2) = \sum_i p_i x_i^2 \neq \left( \sum_i p_i x_i \right)^2 = E(X)^2$$

Propriétés de la variance :

$$\sigma^2(X) \geq 0$$

$$\sigma^2(aX) = a^2 \sigma^2(X) \quad \text{avec } a \in \mathbb{R}$$

$$\sigma^2(X + a) = \sigma^2(X) \quad \text{avec } a \in \mathbb{R}$$

Si, de plus,  $X$  et  $Y$  sont des variables aléatoires indépendantes

$$\sigma^2(X + Y) = \sigma^2(X - Y) = \sigma^2(X) + \sigma^2(Y)$$

- L'espérance d'une variable aléatoire est un indicateur de position sur la distribution de probabilité de  $X$ . Les résultats de  $X$  tombent autour de cette espérance. L'écart-type permet de caractériser comment les valeurs de  $X$  sont dispersées autour de la valeur moyenne. Si cet indicateur est nul, cela veut dire que le résultat de  $X$  est certain.

- La "**médiane**" est la valeur de la variable aléatoire  $X$  qui, pour une série de résultats étudiée, permet de séparer en deux groupes égaux en nombre les résultats. 50 % des résultats ont une valeur inférieure ou égale à la médiane et 50 % des résultats ont une valeur supérieure ou égale à la médiane.
- Le "**mode**" ou la "**valeur modale**" de la variable aléatoire  $X$ , noté  $Mo$ , est la (les) valeur(s) de la réalisation de  $X$  qui a (ont) la plus grande probabilité de réalisation

$$P(X=Mo) \geq P(X=x_i) \quad \forall x_i$$

Exemple : Reprenons l'exemple des mesures de longueur de la figure 8.1. Supposons que la répartition des mesures observées corresponde exactement à la loi de probabilité de ces mesures. La fréquence de chaque mesure correspond donc à sa probabilité. Reprenons le tableau des mesures en indiquant dessous chaque mesure sa probabilité.

10,4 m	10,7 m	10,9 m	11 m	11,1 m	11,2 m	11,3 m	11,5 m	11,7 m	12 m
$\frac{1}{18}$	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{3}{18}$	$\frac{3}{18}$	$\frac{2}{18}$	$\frac{2}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$

Pour cette loi de probabilité, l'espérance est

$$E(X) = \frac{10,4}{18} + \frac{10,7}{18} + \frac{10,9}{9} + \frac{11}{6} + \frac{11,1}{6} + \frac{11,2}{9} + \frac{11,3}{9} + \frac{11,5}{18} + \frac{11,7}{18} + \frac{12}{18} = 11,14$$

L'écart-type vaut

$$\sigma(X) = \frac{10,4^2}{18} + \frac{10,7^2}{18} + \frac{10,9^2}{9} + \frac{11^2}{6} + \frac{11,1^2}{6} + \frac{11,2^2}{9} + \frac{11,3^2}{9} + \frac{11,5^2}{18} + \frac{11,7^2}{18} + \frac{12^2}{18} - 11,14^2 \approx 0,35$$

La médiane de ces résultats est : 11,1, car la moitié des résultats est inférieure ou égale à cette valeur et l'autre moitié supérieure ou égale.

Les modes de ces résultats sont : 11 et 11,1 car, pour ces deux valeurs, il y a trois occurrences ce qui est le plus grand nombre d'occurrences de ces résultats.

*Remarque* : Suivant la forme de la loi de probabilité, la moyenne, la médiane et le mode peuvent avoir parfois des valeurs très différentes.

- Une variable aléatoire "**centrée**" est une variable aléatoire dont l'espérance est nulle.

Soit  $X$  une variable aléatoire, la variable  $X - E(X)$  est centrée.

Une variable aléatoire "**réduite**" est une variable aléatoire dont l'écart-type est égal à 1

Soit  $X$  une variable aléatoire, la variable aléatoire sans unité  $X/\sigma(X)$  est réduite.

Soit  $X$  une variable aléatoire, la variable aléatoire  $\frac{X - E(X)}{\sigma(X)}$  est centrée et réduite.

### 8.3.3 Principales lois des variables discrètes

#### Loi de Bernoulli

Une variable aléatoire  $B$  ayant deux valeurs possibles 0 et 1 qu'elle prend avec les probabilités respectives  $q$  et  $p$  (avec  $q = 1 - p$ ) suit la loi de Bernoulli. On parle alors de **variable de Bernoulli** de paramètre  $p$ .

La moyenne de la loi de Bernoulli est  $p$ .

La variance de la loi de Bernoulli est  $p(1 - p) = pq$

L'écart-type de la loi de Bernoulli est  $\sqrt{pq}$

On parle parfois de loi de Bernoulli pour des cas où la variable aléatoire prend d'autres valeurs que 0 et 1 (la moyenne et la variance peuvent alors changer de valeur). La loi de Bernoulli sert dans tous les cas où l'on traite d'une simple alternative dont on sait suivant quel pourcentage elle se produit.

#### Loi Binomiale

Une variable aléatoire Bin suit une **loi binomiale** de paramètres  $p$  et  $n$  lorsqu'elle prend les valeurs 0, 1, ...,  $n$  avec les probabilités  $p_1, p_2, \dots, p_n$  suivant la formule

$$p_i = C_n^i p^i q^{n-i}$$

avec  $q = 1 - p$

La moyenne de la loi binomiale est  $np$ .

La variance de la loi binomiale est  $npq$ .

L'écart-type de la loi binomiale est  $\sqrt{npq}$

Le mode de la loi binomiale est :  $np + p - 1 \leq Mo \leq np + p$

La démonstration de ces résultats est pénible quoique faisable, nous les admettons donc.

La somme de  $n$  variables de Bernoulli indépendantes de même paramètre  $p$  est une variable aléatoire qui suit la loi de Bernoulli. De façon générale, la loi binomiale est utilisée dès qu'il s'agit d'un décompte de succès parmi  $n$  épreuves répétées, identiques et indépendantes à deux issues : échec ou succès.

Si  $X$  et  $Y$  sont deux variables aléatoires suivant des lois binomiales de paramètres respectivement  $(n_X, p)$  et  $(n_Y, p)$ , la somme de ces deux variables aléatoires suit la loi binomiale de paramètres  $(n_X + n_Y, p)$

#### Loi de poisson

Lorsqu'une variable aléatoire  $X$  suit la **loi de Poisson**, ses valeurs possibles sont 0, 1, 2, ...,  $k$ , ... La probabilité  $p_k$  d'obtenir la valeur  $k$  est donnée par la formule

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

On appelle  $\lambda$  le paramètre de la loi de Poisson.

La moyenne de la loi de Poisson est  $\lambda$ .

La variance de la loi de Poisson est  $\lambda$ .

L'écart-type de la loi de Poisson est  $\sqrt{\lambda}$

Le mode de la loi de Poisson est :  $\lambda - 1 \leq Mo \leq \lambda$

La démonstration de ces résultats ainsi que la vérification du fait qu'il s'agit bien d'une loi de probabilité nécessite le développement en série entière de la fonction exponentielle et est donc hors-programme. Elle est présente toutefois dans la plupart des livres sur les probabilités.

La loi de Poisson a de très nombreuses applications, puisqu'elle permet de représenter la survenue d'événements rares dans le temps et/ou dans l'espace. On peut la faire intervenir, par exemple, dans les titrages des solutions bactériennes, pour comptabiliser des accidents ou pour la pharmacovigilance.

Si  $X$  et  $Y$  sont deux variables aléatoires indépendantes suivant des lois de Poisson de paramètre respectif  $\lambda$  et  $\mu$ , la somme de ces variables est une variable aléatoire suivant une loi de Poisson de paramètre  $\lambda + \mu$ .

Remarque : Si l'on a une valeur élevée pour  $n$ , en gardant  $p$  petit et  $np$  fini, alors la loi binomiale de paramètres  $(n, p)$  peut être approximée par la loi de Poisson de paramètre  $\lambda = np$ .

Suivant la précision recherchée, les limites pour ces approximations restent à discuter et dépendent du contexte. On donne souvent des valeurs de l'ordre de 3 ou 5 pour la valeur supérieure de  $np$  et des valeurs de plusieurs dizaines (typiquement 30) pour la valeur inférieure de  $n$ .

## 8.4 Variables aléatoires continues

### 8.4.1 Définitions

Une variable aléatoire est dite "**continue**" si son domaine de variation est l'ensemble des réels  $\mathbb{R}$  ou un intervalle de cet ensemble.

Une variable aléatoire continue  $X$  peut être définie par sa **fonction de répartition**  $F$  ou  $F_X$ , définie comme pour les variables aléatoires discrètes

$$F(x) = F_X(x) = P(X \leq x)$$

La fonction de répartition  $F$  d'une variable aléatoire continue a les propriétés suivantes

$$\begin{aligned} F(x) &\text{ est continue} \\ F(x) &\text{ est croissante : } x_1 \leq x_2 \implies F(x_1) \leq F(x_2) \\ 0 &\leq F(x) \leq 1 \\ \lim_{x \rightarrow -\infty} F(x) &= 0 \quad \text{et} \quad \lim_{x \rightarrow \infty} F(x) = 1 \end{aligned}$$

La fonction  $f$  telle que  $f(x) = \frac{dF(x)}{dx}$  où  $F$  est la fonction de répartition d'une variable aléatoire continue  $X$  est appelée fonction "**densité de probabilité**" de la variable aléatoire  $X$ . On parle parfois aussi de "**fonction de distribution**".

On définit aussi la densité de probabilité de la manière suivante

$$f(x) dx = P(x \leq X \leq x+dx)$$

Ces deux définitions sont équivalentes.

La densité de probabilité d'une variable aléatoire continue a les propriétés suivantes

$$\begin{aligned} f(x) &\geq 0 \\ f &\text{ est une fonction intégrable et } \int_{-\infty}^{+\infty} f(x) dx = 1 \end{aligned}$$

### 8.4.2 Paramètres caractéristiques

Les paramètres caractéristiques sont les mêmes que dans le cas des variables discrètes. Cependant, il est utile de préciser leur définition dans le cas continu.

- L'**espérance** d'une variable aléatoire continue  $X$  de densité de probabilité  $f$  est donnée par

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

si cette intégrale existe.

L'espérance mathématique, et donc la moyenne, présente quelques propriétés.

Soient  $X$  et  $Y$  deux variables aléatoires

$$E(aX+b) = a E(X) + b \quad \text{avec } a \in \mathbb{R} \text{ et } b \in \mathbb{R}$$

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

Si, de plus,  $X$  et  $Y$  sont des variables aléatoires indépendantes

$$E(X \times Y) = E(X) \times E(Y)$$

Une variable aléatoire continue **centrée** est une variable aléatoire dont l'espérance est nulle.

Soit  $X$  une variable aléatoire, la variable  $X - E(X)$  est centrée.

- La **variance** d'une variable aléatoire continue  $X$  est définie par

$$\sigma^2(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx = E(X^2) - E(X)^2$$

L'**écart-type**  $\sigma(X)$  est la racine carrée de la variance.

Propriétés de la variance :

$$\sigma^2(X) \geq 0$$

$$\sigma^2(aX) = a^2 \sigma^2(X) \quad \text{avec } a \in \mathbb{R}$$

$$\sigma^2(X + a) = \sigma^2(X) \quad \text{avec } a \in \mathbb{R}$$

Si, de plus,  $X$  et  $Y$  sont des variables aléatoires indépendantes

$$\sigma^2(X + Y) = \sigma^2(X - Y) = \sigma^2(X) + \sigma^2(Y)$$

Une variable aléatoire **réduite** est une variable aléatoire

dont l'écart-type est égal à 1

Soit  $X$  une variable aléatoire, la variable aléatoire sans unité  $X/\sigma(X)$  est réduite.

Soit  $X$  une variable aléatoire continue, la variable aléatoire  $\frac{X - E(X)}{\sigma(X)}$  est centrée et réduite

- La **médiane** d'une variable aléatoire continue est le nombre réel  $Me$  tel que

$$F(Me) = 0,5$$

La médiane correspond à la valeur de  $X$  qui divise la distribution en deux parts équiprobables

$$P(X \leq Me) = P(X \geq Me) = 0,5$$

- Le **mode**  $Mo$  d'une variable aléatoire continue est défini comme la (ou les) valeur(s) qui correspond(ent) à un maximum local de la fonction densité de probabilité.

Un fonction de densité de probabilité ne présentant qu'un seul maximum est dite "**unimodale**".

## 8.4.3 Principales lois de probabilités

**Loi uniforme**

La distribution uniforme est la plus simple loi de variable aléatoire continue. Elle correspond à une densité de probabilité constante pour tout l'ensemble de valeurs de la variable aléatoire. Elle est notamment rencontrée lors d'un tirage d'un nombre aléatoire compris dans un intervalle  $[a, b]$ .

La fonction densité de probabilité d'une "**loi uniforme**" sur l'intervalle  $[a, b]$ ,  $a < b$ , est définie par

$$f(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{si } x > b \end{cases}$$

Nous pouvons calculer la fonction de répartition  $F$  associée à une distribution uniforme

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y) dy$$

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

$F(x)$  est donc composée de trois tronçons de droites : la droite horizontale à 0 jusqu'au point  $a$ , la droite horizontale à 1 à partir du point  $b$  et la droite qui relie les deux précédentes.

Nous pouvons de même calculer les paramètres caractéristiques :

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx = \int_a^b \frac{x-a}{b-a} dx = \frac{1}{b-a} \left[ \frac{x^2}{2} \right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

c'est-à-dire le milieu du segment  $[a, b]$  ce qui est assez intuitif.

$$\sigma^2(X) = E(X^2) - E(X)^2 = \int_{-\infty}^{+\infty} x^2 f(x) dx - E(X)^2 = \frac{(b-a)^2}{12}$$

**Loi normale**

*Remarque* : Sont synonymes :

**loi normale, loi de Gauss, loi de Laplace, loi de Laplace-Gauss**

On dit aussi d'une variable aléatoire suivant une loi normale qu'elle est gaussienne, et la courbe de la distribution est souvent qualifiée de courbe en cloche (the bell curve).

Pourquoi tant de noms ? Tout simplement parce que la loi normale est la plus utilisée de toutes les fonctions de distributions. Dans la plupart des cas, lorsque l'on nous donne un intervalle d'incertitude sur une mesure d'une quantité, on considère que la distribution de ces erreurs suit une loi normale.

Une variable aléatoire  $X$  suit une "**loi normale**" si sa densité de probabilité s'écrit

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

Elle est définie pour  $-\infty < x < +\infty$ .

Les deux paramètres  $\bar{X}$  et  $\sigma$  de la densité de probabilité sont respectivement la *moyenne* et l'*écart-type*.

La fonction de répartition  $F$  de la loi normale est donnée par

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(y - \bar{X})^2}{2\sigma^2}} dy$$

Les figures 8.2 et 8.3 donnent l'allure de ces deux fonctions pour des variables centrées et réduites; dans ce cas précis, on compare parfois la fonction de répartition  $F(x)$  à une fonction mathématique classique (et tabulée dans de nombreux ouvrages), la fonction erf

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du = 2F(x\sqrt{2}) - 1$$

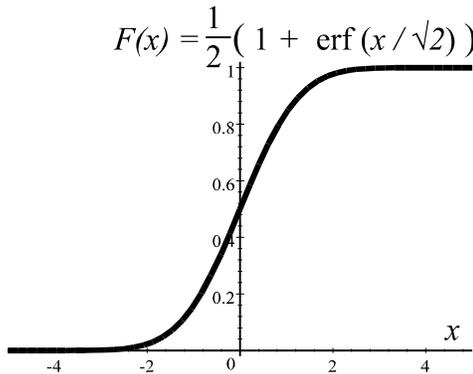


fig. 8.2 :  $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$

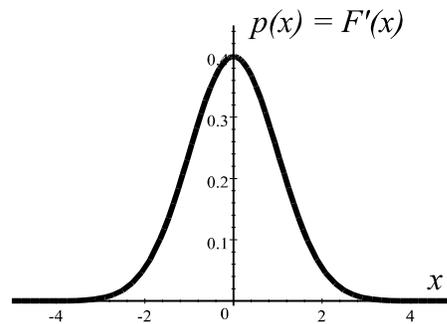


fig. 8.3 :  $p(x) = \frac{dF}{dx} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

Si  $X$  et  $Y$  sont deux variables aléatoires indépendantes distribuées suivant la loi normale, toute combinaison linéaire  $aX + bY + c$  de ces deux variables est une variable aléatoire qui suit la loi normale.

La moyenne est alors  $a\bar{X} + b\bar{Y} + c$ .

Et la variance devient  $a^2\sigma_X^2 + b^2\sigma_Y^2$ .

Cette dernière propriété nous permet de rapporter toute étude de variable aléatoire distribuée normalement de paramètres quelconques en l'étude d'une variable aléatoire suivant la loi normale centrée réduite. En effet, soit  $X$  une variable aléatoire continue suivant la loi normale de moyenne  $\bar{X}$  et d'écart-type  $\sigma$ ,  $Z = \frac{X - \bar{X}}{\sigma}$  est une variable aléatoire centrée réduite ainsi que nous l'avons déjà vu. Si on applique alors le changement de variable  $X \rightarrow Z$  (en effectuant le changement de borne correspondant  $x \rightarrow z = \frac{x - \bar{X}}{\sigma}$  dans l'intégrale), on a

$$F_X(x) = P(X \leq x) = P\left(\frac{X - \bar{X}}{\sigma} \leq \frac{x - \bar{X}}{\sigma}\right) = P(Z \leq z) = F_Z(z)$$

$F_Z(z)$  étant une fonction de répartition centrée réduite.

Pour résoudre des exercices comportant des variables suivant une ou des lois normales, il est donc précieux d'avoir sous la main une table comme celle de la figure 8.4 et de savoir l'utiliser.

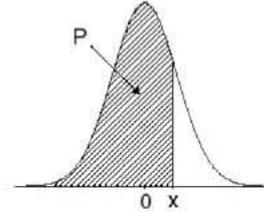
**Table de la loi normale centrée-réduite**

La table indique, pour  $x \geq 0$ , la valeur  $F(x)$  de la fonction de répartition de la loi normale centrée-réduite définie par

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

Pour  $x < 0$ ,  $F(x) = 1 - F(-x)$ .

La table retourne la valeur de  $F(x)$  pour la valeur de  $x$  lue comme la somme des valeurs figurant en tête de la ligne et de la colonne correspondantes. Ex. :  $x = 0.83 = 0.8 + 0.03 \rightarrow F(x) = 0.7967$ .



x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9 <sup>0</sup> 03	0.9 <sup>0</sup> 06	0.9 <sup>0</sup> 10	0.9 <sup>0</sup> 13	0.9 <sup>0</sup> 16	0.9 <sup>0</sup> 18	0.9 <sup>0</sup> 21	0.9 <sup>0</sup> 24	0.9 <sup>0</sup> 26	0.9 <sup>0</sup> 29
3.2	0.9 <sup>0</sup> 31	0.9 <sup>0</sup> 34	0.9 <sup>0</sup> 36	0.9 <sup>0</sup> 38	0.9 <sup>0</sup> 40	0.9 <sup>0</sup> 42	0.9 <sup>0</sup> 44	0.9 <sup>0</sup> 46	0.9 <sup>0</sup> 48	0.9 <sup>0</sup> 50
3.3	0.9 <sup>0</sup> 52	0.9 <sup>0</sup> 53	0.9 <sup>0</sup> 55	0.9 <sup>0</sup> 57	0.9 <sup>0</sup> 58	0.9 <sup>0</sup> 60	0.9 <sup>0</sup> 61	0.9 <sup>0</sup> 62	0.9 <sup>0</sup> 64	0.9 <sup>0</sup> 65
3.4	0.9 <sup>0</sup> 66	0.9 <sup>0</sup> 68	0.9 <sup>0</sup> 69	0.9 <sup>0</sup> 70	0.9 <sup>0</sup> 71	0.9 <sup>0</sup> 72	0.9 <sup>0</sup> 73	0.9 <sup>0</sup> 74	0.9 <sup>0</sup> 75	0.9 <sup>0</sup> 76
3.5	0.9 <sup>0</sup> 77	0.9 <sup>0</sup> 78	0.9 <sup>0</sup> 78	0.9 <sup>0</sup> 79	0.9 <sup>0</sup> 80	0.9 <sup>0</sup> 81	0.9 <sup>0</sup> 81	0.9 <sup>0</sup> 82	0.9 <sup>0</sup> 83	0.9 <sup>0</sup> 83
3.6	0.9 <sup>0</sup> 84	0.9 <sup>0</sup> 85	0.9 <sup>0</sup> 85	0.9 <sup>0</sup> 86	0.9 <sup>0</sup> 86	0.9 <sup>0</sup> 87	0.9 <sup>0</sup> 87	0.9 <sup>0</sup> 88	0.9 <sup>0</sup> 88	0.9 <sup>0</sup> 89
3.7	0.9 <sup>0</sup> 89	0.9 <sup>0</sup> 90	0.9 <sup>0</sup> 90	0.9 <sup>0</sup> 91	0.9 <sup>0</sup> 91	0.9 <sup>0</sup> 92	0.9 <sup>0</sup> 92	0.9 <sup>0</sup> 93	0.9 <sup>0</sup> 93	0.9 <sup>0</sup> 94
3.8	0.9 <sup>0</sup> 94	0.9 <sup>0</sup> 95	0.9 <sup>0</sup> 95	0.9 <sup>0</sup> 96	0.9 <sup>0</sup> 96	0.9 <sup>0</sup> 97	0.9 <sup>0</sup> 97	0.9 <sup>0</sup> 98	0.9 <sup>0</sup> 98	0.9 <sup>0</sup> 99
3.9	0.9 <sup>0</sup> 99									

N.B. : La notation 0.9<sup>0</sup>03, par exemple, équivaut à 0.99903.

fig. 8.4 : Table de la loi normale centrée réduite

Exemple 1 : Cherchons  $P(Z \leq 1,17)$ .

Sur la table de la figure 8.4, nous devons chercher dans la colonne de gauche le chiffre des unités et le premier chiffre après la virgule de  $Z$ . Il s'agit ici de 1,1. Il faut suivre maintenant la ligne correspondante jusqu'à la colonne correspondant au deuxième chiffre après la virgule de  $Z$ . Ici, il s'agit de 0,07. Au croisement de la ligne 1,1 et de la colonne 0,07, nous lisons la probabilité recherchée :  $P(Z \leq 1,17) = 0,8790$

### **Théorème central limite**

Ce théorème est essentiel pour comprendre l'importance de la loi normale et son omniprésence dans tous les domaines.

Soient  $X_1, X_2, \dots, X_n$ ,  $n$  variables aléatoires indépendantes suivant des lois de probabilités quelconque d'espérance  $E(X_i)$  et de variance finie  $\sigma_i^2$ . La loi suivie par la variable aléatoire :  $X = \sum_{i=1}^n X_i$  peut être approximée pour  $n$  grand (et sous certaines conditions) par une loi normale de paramètres

$$\bar{X} = \sum_{i=1}^n E(X_i) \quad \text{et} \quad \sigma = \sqrt{\sum_{i=1}^n \sigma_i^2}$$

Cela signifie concrètement que dès que l'on a un grand nombre de mesures d'une même quantité, ces mesures vont se distribuer suivant une loi normale quelles que soient les origines des variations dans les mesures. L'expérience montre que le théorème central limite s'applique à la quasi totalité des mesures effectuées dans le domaine industriel, économique ou scientifique, sciences naturelles ou sciences humaines.

## **8.5 Statistique, échantillonnage et estimation**

### *8.5.1 Définitions*

Les "**statistiques**" reposent sur des ensembles de données, d'observations : recensement, cadastre, sondage, ... (ce ne sont donc que des chiffres).

Les probabilités, ainsi que nous l'avons vu, forment une branche des mathématiques et représentent donc une description rigoureuse et exacte ; pour cela, elles travaillent sur des objets mathématiques parfaitement définis et abstraits (bien que d'origine concrète).

La statistique est donc la science qui utilise les méthodes mathématiques (dont beaucoup viennent des probabilités) pour étudier et analyser des statistiques en vue :

- d'en accroître les connaissances scientifiques
- de planifier des stratégies
- d'aider à des prises de décisions, etc.

Cette approche est valable quel que soit le domaine abordé : biologie, médecine, physique, mécanique, chimie, ...

La "**population**" est l'ensemble étudié par la statistique.

L'"**individu**" est un élément de la population.

Le "**caractère**", le "**facteur**" ou la "**variable**" qualifient toute caractéristique prise par les individus de la population.

L'"**échantillon**" est un sous-ensemble de la population.

On parle de "**séries statistiques**" lorsque l'on fait la correspondance entre les individus de la population et les valeurs des facteurs que l'on étudie sur cette population.

Sur la table de la figure 8.4, nous devons chercher dans la colonne de gauche le chiffre des unités et le premier chiffre après la virgule de  $Z$ . Il s'agit ici de 1,1. Il faut suivre maintenant la ligne correspondante jusqu'à la colonne correspondant au deuxième chiffre après la virgule de  $Z$ . Ici, il s'agit de 0,07. Au croisement de la ligne 1,1 et de la colonne 0,07, nous lisons la probabilité recherchée :  $P(Z \leq 1,17) = 0,8790$

### **Théorème central limite**

Ce théorème est essentiel pour comprendre l'importance de la loi normale et son omniprésence dans tous les domaines.

Soient  $X_1, X_2, \dots, X_n$ ,  $n$  variables aléatoires indépendantes suivant des lois de probabilités quelconques d'espérance  $E(X_i)$  et de variance finie  $\sigma_i^2$ . La loi suivie par la variable aléatoire :  $X = \sum_{i=1}^n X_i$  peut être approximée pour  $n$  grand (et sous certaines conditions) par une loi normale de paramètres

$$\bar{X} = \sum_{i=1}^n E(X_i) \quad \text{et} \quad \sigma = \sqrt{\sum_{i=1}^n \sigma_i^2}$$

Cela signifie concrètement que dès que l'on a un grand nombre de mesures d'une même quantité, ces mesures vont se distribuer suivant une loi normale quelles que soient les origines des variations dans les mesures. L'expérience montre que le théorème central limite s'applique à la quasi totalité des mesures effectuées dans le domaine industriel, économique ou scientifique, sciences naturelles ou sciences humaines.

## **8.5 Statistique, échantillonnage et estimation**

### *8.5.1 Définitions*

Les "**statistiques**" reposent sur des ensembles de données, d'observations : recensement, cadastre, sondage, ... (ce ne sont donc que des chiffres).

Les probabilités, ainsi que nous l'avons vu, forment une branche des mathématiques et représentent donc une description rigoureuse et exacte ; pour cela, elles travaillent sur des objets mathématiques parfaitement définis et abstraits (bien que d'origine concrète).

La statistique est donc la science qui utilise les méthodes mathématiques (dont beaucoup viennent des probabilités) pour étudier et analyser des statistiques en vue :

- d'en accroître les connaissances scientifiques
- de planifier des stratégies
- d'aider à des prises de décisions, etc.

Cette approche est valable quel que soit le domaine abordé : biologie, médecine, physique, mécanique, chimie, ...

La "**population**" est l'ensemble étudié par la statistique.

L'"**individu**" est un élément de la population.

Le "**caractère**", le "**facteur**" ou la "**variable**" qualifient toute caractéristique prise par les individus de la population.

L'"**échantillon**" est un sous-ensemble de la population.

On parle de "**séries statistiques**" lorsque l'on fait la correspondance entre les individus de la population et les valeurs des facteurs que l'on étudie sur cette population.

La notion d'échantillon apparaît souvent dans le langage courant : on le retrouve quotidiennement dans les journaux à propos des innombrables sondages qui visent à donner une image de l'opinion de la population globale à partir des opinions recueillies sur un nombre limité de personnes. On le retrouve aussi lorsque l'on veut acheter du papier peint pour tapisser une pièce, on cherche à déterminer le résultat final à partir d'un petit morceau de papier appelé également "échantillon". L'échantillon statistique est constitué d'un ensemble de sujets tirés au sort dans la population à laquelle on s'intéresse. Comme dans l'exemple du papier peint, l'échantillon a pour objet de donner des informations les plus fiables possibles sur l'ensemble de la population. Un des problèmes essentiels consiste à déterminer la bonne taille de l'échantillon pour répondre à une question donnée. Reprenons l'exemple du papier peint : s'il comporte un motif périodique, l'échantillon minimal doit avoir la taille de ce motif ; si, par contre, le papier peint est uni, un échantillon plus réduit peut suffire pour tester la couleur. Une fois, l'échantillon déterminé, il faut en déduire une estimation des paramètres de la population et donner la précision que l'on peut obtenir sur ces paramètres ; c'est l'objet de ce cours.

### 8.5.2 Caractéristiques d'un échantillon statistique

Un "**échantillon statistique**" de taille  $n$  d'une variable aléatoire  $X$  suivant une loi de paramètres  $\bar{X}$  et  $\sigma$  est obtenu en répétant  $n$  fois de façons indépendantes l'épreuve qui mène à  $X$ . Cet échantillon est noté  $(X_1, X_2, \dots, X_n)$ . Il s'agit donc d'une variable aléatoire à  $n$  dimensions. La variable  $X_i$  de l'échantillon suit la même loi que  $X$  par définition. On a donc  $E(X_i) = \bar{X}$  et  $\sigma_{X_i} = \sigma$ . On peut noter parfois  $(x_1, x_2, \dots, x_n)$  les valeurs des variables aléatoires une fois les mesures faites. On l'appelle alors de l'échantillon réalisé.

La loi de la variable aléatoire  $X$  qu'on étudie grâce à un échantillon est dite "**loi parente**".

La moyenne  $M_n$  d'un échantillon de  $n$  variables aléatoires  $(X_1, X_2, \dots, X_n)$  est la moyenne arithmétique de ses valeurs

$$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

$M_n$  est une variable aléatoire qui a pour espérance

$$E(M_n) = \bar{X} \text{ et pour écart-type } \sigma_{M_n} = \sigma/n.$$

Démonstration : D'après les propriétés de l'espérance et de la variance,

$$E(M_n) = \frac{E(X_1) + E(X_2) + \dots + E(X_n)}{n} = \frac{n E(X)}{n} = E(X) = \bar{X}$$

$$\sigma_{M_n}^2 = \frac{1}{n^2} \sigma^2 (X_1 + X_2 + \dots + X_n) = \frac{n \sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Si l'on ajoute à cela que la variable  $X$  suit une loi normale,  $M_n$  suit aussi une loi normale de moyenne  $\bar{X}$  et de variance  $\frac{\sigma^2}{n}$ .

D'après le théorème central limite vu à la section précédente, plus la taille de l'échantillon est grande, plus la distribution de la moyenne  $M_n$  des valeurs de l'échantillon s'approche d'une loi normale quelle que soit la loi de la variable  $X$ .  $M_n$  suit donc une loi normale de moyenne  $\bar{X}$  et de variance  $\frac{\sigma^2}{n}$  pour des valeurs de  $n$  supérieures typiquement à 30.

### 8.5.3 Estimateur

Un "**estimateur**" est une valeur calculée à partir des différentes valeurs  $x_1, x_2, \dots, x_n$  de l'échantillon. Il est destiné à donner une valeur la plus proche possible du paramètre inconnu de la population.

On appelle "**biais**" d'un estimateur la différence entre la valeur moyenne qu'il prendrait sur une infinité d'échantillons et la valeur recherchée du paramètre à estimer. Un estimateur est dit sans biais si sa valeur est en moyenne égale à la vraie valeur du paramètre de la population, pour une taille d'échantillon fixée. Dans le cas contraire, l'estimateur est dit biaisé.

Un estimateur est "**convergent**" si la valeur moyenne des carrés des différences entre les estimations et le paramètre recherché devient de plus en plus petite au fur et à mesure que la taille de l'échantillon augmente.

On démontre que :

- le meilleur estimateur de la moyenne de la population est :

$$m = M_n \text{ moyenne de l'échantillon.}$$

- le meilleur estimateur de la variance de la population est :

$$s^2 = \frac{n}{n-1} \sigma_{M_n}^2$$

- le meilleur estimateur de la probabilité dans la population est :

$$f, \text{ fréquence observée}$$

Ces trois estimateurs sont sans biais.

Par ailleurs, si la variable  $X$  suit une loi normale, la variable  $\frac{M_n - \bar{X}}{s/\sqrt{n}}$  est une variable aléatoire qui suit une loi connue appelée loi de Student à  $(n-1)$  degrés de liberté ( $\nu = n-1$ ), dont la table est présentée dans la figure 8.5.

### 8.5.4 Intervalle de confiance

Supposons que l'on réalise un sondage électoral. On dispose d'un échantillon de 850 personnes dont 422 affirment vouloir voter pour A. On fournit une estimation "**ponctuelle**" des intentions de vote dans la population si on fournit un seul chiffre : ce serait ici 49,6 %. Cette présentation est souvent utilisée dans les journaux, mais elle n'est pas satisfaisante car elle ne dit rien sur la pertinence de l'échantillon choisi et donc sur la fiabilité de ce chiffre. Il arrive que, par prudence, les journaux préfèrent donner une "**fourchette**", c'est-à-dire, un intervalle censé encadrer le pourcentage réel de la population. Par exemple, de 47% à 52%. Mais cet intervalle ne doit pas être choisi n'importe comment, il faut trouver une taille raisonnable qui ne doit pas être trop grande, sinon l'information ne devient plus intéressante (comme dans le cas : de 0% à 100%), ni trop petit car alors on prend un grand risque de se tromper. Ainsi doit-on se contenter de fournir un intervalle de confiance dans lequel on peut avoir une certaine confiance que la vraie valeur se trouve bien qu'il y ait un certain risque de se tromper.

On appelle "**intervalle de confiance**" à  $1-\alpha$  (on dit aussi intervalle de confiance à risque  $\alpha$ ) d'une variable aléatoire  $X$  (ou d'un paramètre) un intervalle  $[a, b]$  dans lequel cette variable (ou ce paramètre) a une probabilité  $\alpha$  de se trouver.

**Remarques :**

- Si la loi de répartition de la variable est symétrique, la valeur centrale  $c$  est souvent prise comme milieu de  $[a, b]$ .

L'intervalle prend alors la forme symétrique  $[c-l_\alpha, c+l_\alpha]$ ,  $l_\alpha$  dépendant de la loi de la variable, de  $\alpha$  et de la taille de l'échantillon.

Dans ce cas, on définit un *coefficient de précision* :  $\frac{l_\alpha}{c}$

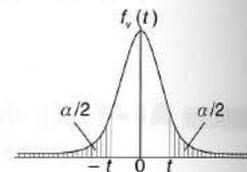
- Si la loi n'est pas symétrique, on centrera l'intervalle sur la probabilité.

- On peut aussi chercher des intervalles de confiance de la forme  $[-\infty, b]$  ou  $[a, +\infty]$ .

**A2-Table de distribution de T (loi de Student)**

$\alpha = P(|T| \geq t)$

$v =$  nombre de degrés de liberté



$\alpha \backslash v$	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,929
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
$\infty$	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

fig. 8.5 : Loi de Student

On suppose, à présent et pour toute la suite, que la variable aléatoire  $X$  suit une loi normale de moyenne  $\bar{X}$  et de variance  $\sigma^2$ .

Cela est toutefois aussi vrai pour n'importe quelle loi si  $n$  est très grand (au moins supérieur à 30).

– *Intervalle de confiance de la variable  $X$  à  $1-\alpha$*

Trois cas doivent être envisagés :

-  $\bar{X}$  et  $\sigma^2$  sont connus :

$$[\bar{X} - z_\alpha \sigma ; \bar{X} + z_\alpha \sigma]$$

-  $\bar{X}$  est inconnu mais estimé par  $M_n$  et  $\sigma^2$  est connu :

$$[M_n - z_\alpha \sigma ; M_n + z_\alpha \sigma]$$

Dans les deux cas,  $z_\alpha$  est la valeur de  $z$  trouvée dans la table de la figure 8.4 pour une probabilité  $\alpha$ .

-  $\bar{X}$  et  $\sigma^2$  sont inconnus mais estimé par  $M_n$  et  $s^2$  et  $n$  suffisamment grand :

$$[M_n - t_{\alpha, n-1} s ; M_n + t_{\alpha, n-1} s]$$

$t_{\alpha, n-1}$  étant issu de la loi de Student à  $n-1$  degré de liberté dont nous avons déjà parlé (fig. 8.5).

– *Intervalle de confiance de la variable  $M_n$  à  $1-\alpha$*

$$[\bar{X} - z_\alpha \frac{\sigma}{\sqrt{n}} ; \bar{X} + z_\alpha \frac{\sigma}{\sqrt{n}}]$$

$z_\alpha$  est la valeur de  $z$  trouvée dans la table de la figure 8.4 pour une probabilité  $\alpha$ .

– *Intervalle de confiance de la variable  $\bar{X}$  à  $1-\alpha$*

Deux cas doivent être envisagés :

-  $\bar{X}$  est inconnu mais estimé par  $M_n$  et  $\sigma^2$  est connu :

$$[M_n - z_\alpha \frac{\sigma}{\sqrt{n}} ; M_n + z_\alpha \frac{\sigma}{\sqrt{n}}]$$

$z_\alpha$  est la valeur de  $z$  trouvée dans la table de la figure 8.4 pour une probabilité  $\alpha$ .

-  $\bar{X}$  et  $\sigma^2$  sont inconnus mais estimé par  $M_n$  et  $s^2$  et  $n$  suffisamment grand :

$$[M_n - t_{\alpha, n-1} \frac{s}{\sqrt{n}} ; M_n + t_{\alpha, n-1} \frac{s}{\sqrt{n}}]$$

$t_{\alpha, n-1}$  étant issu de la loi de Student à  $n-1$  degré de liberté dont nous avons déjà parlé (fig. 8.5).