

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche  
Scientifique

Université 08 Mai 1945 – Guelma

Faculté de Mathématiques et de l'Informatique et Sciences de  
la Matière

Département de Mathématiques



Polycopié de cours:

# Mathématiques I

1<sup>ère</sup> Année biologie (Sciences de la Terre et de l'Univers)

Dr : Hamid Boulares

Guelma 2016

# Tables des matières

## **1 Statistiques descriptives**

1.1 Paramètre de dispersion et de position

1.2 Représentations graphiques usuelles

1.3 Méthode des moindres carrées

1.4 Droite de régression

1.5 Ajustement par des fonctions de puissances

Exercices

## **2 Probabilités**

2.1 Vocabulaire de base

2.2 Probabilités élémentaires

2.3 Probabilités conditionnelles

2.4 Variables aléatoires discrètes

2.5 Variables aléatoires continues

Exercices

## **3 Tests d'adéquation du Khi-deux**

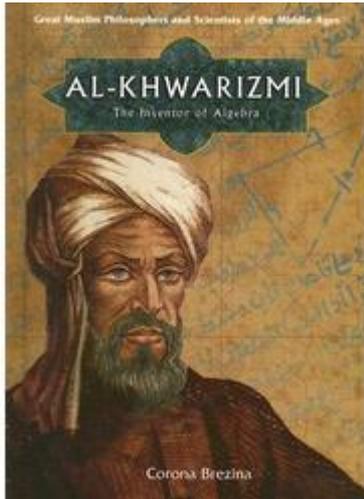
3.1 Test de comparaison de deux échantillons

Exemples

## **4 Tables**

# Chapitre 1

## LA STATISTIQUE DESCRIPTIVE



Muhammad Ibn Mussa Al Khawarizmi, mathématicien, philosophe perse né à Khawa et mort à Bagdad en 850. Il publia de grands ouvrages en astronomie, en mathématiques par exemple le fameux livre ``Al jabr wal muqabalah`` qui signifie ``l'algèbre par la preuve``. Il est l'inventeur des algorithmes chers aux informaticiens. Il est le premier à parler de l'aléatoire dans son manuscrit ``Azzahr`` qui deviendra plus tard ``Hasard`` qui signifie chance.

La statistique est une branche des mathématiques qui s'intéresse à l'étude des phénomènes aléatoires, en contre opposé aux mathématiques déterministes plus familières aux étudiant(e)s, que sont le calcul différentiel, calcul intégral, géométrie, algèbre,... Le mot statistique dont le nom est dérivé de ``state`` en référence à tout ce qui est étatique, est relativement nouveau, puisqu'il a été introduit en Allemagne au XVIIème siècle. Par contre la pratique de la statistique est plus ancienne, elle fut utile aux grands empires en Mésopotamie, dans l'Égypte ancienne, ainsi que chez les romains et les empires indiens et chinois. Il s'agissait de bien connaître la population pour administrer sa répartition sur les territoires, collecter les impôts et gérer les aspects militaires.

De nos jours, on ne peut trouver un domaine qui peut être compris, analysé sans les méthodes statistiques. Que ce soit dans le domaine des sciences sociales, sciences de la vie ou sciences de l'ingénieur, les méthodes statistiques sont omniprésentes pour mettre de l'ordre dans le protocole de travail, elles permettent quand on est devant un chaos

apparent des données, de déterminer par où commencer et quelles sont les étapes à suivre selon le contexte pour analyser ces données.

La statistique grosso-modo est formée de trois grandes classes : la statistique descriptive, la statistique inférentielle et la nouvelle branche qu'est la statistique exploratrice. Ce chapitre est consacré à la statistique descriptive. La statistique descriptive comme son nom l'indique, se propose de décrire les données, de les classer et de les présenter sous des formes claires et compréhensibles. Elle est à la base par exemple de toute organisation du système d'information d'une entreprise : statistiques de la production ou des ventes, statistiques financières, statistiques des ressources humaines... Elle est aussi une importante composante en sciences humaines de ce qu'on appelle les méthodes quantitatives. On va commencer par définir le lexique qu'on va utiliser tout le long de ce chapitre et même de ce polycopiées (voir [1], [2], [3], [6]).

## 1.1 : Les concepts et le vocabulaire de base.

Au début de tout travail statistique, il faut cerner avec précision sur quoi va porter l'étude. L'ensemble de tous les éléments sur lesquels porte l'étude s'appelle **population**. Une population peut être un ensemble d'êtres vivants (humains, oiseaux, poissons, bactéries,...) ou un ensemble de choses (maisons, voitures, rivières,...) ou un ensemble de faits (pannes, accidents, divorces,...). Chaque élément d'une population s'appelle **individu** ou **unité statistique**. Une population peut être finie (population d'un pays) ou presque infinie (population des insectes), on considère généralement les populations comme finies mêmes si elles sont très grandes. Le nombre d'unités statistiques dans une population s'appelle **taille de la population** et on le note par **N**.

Quand une étude porte sur toute la population, on dit qu'on fait un **recensement**. Mais pour des raisons techniques ou économiques, il n'est généralement pas possible de collecter des données sur tous les éléments d'une population. Alors on se contente d'extraire une partie de la population appelée **échantillon** et restreindre l'étude à cet échantillon. On verra dans le chapitre V, qu'il existe des méthodes spécifiques permettant de s'assurer que l'échantillon soit représentatif de la population, c'est-à-dire une réplique en miniature de ce qui se passe dans la population. Pour l'instant, on suppose qu'on dispose d'un échantillon sur lequel porte l'étude (sans savoir comment il a été extrait). Le nombre d'éléments dans l'échantillon s'appelle **taille de l'échantillon** et sera noté par **n**.

On appelle **variable** tout caractère observé ou mesuré sur chacun des éléments de l'échantillon. On va réserver les dernières lettres de l'alphabet pour noter les variables : X, Y, Z, U...

Les différentes valeurs que prend une variable s'appellent **modalités**. Afin que le classement d'une unité statistique soit toujours possible sans ambiguïté, les différentes modalités doivent être à la fois incompatibles (un individu ne peut avoir plusieurs modalités à la fois) et exhaustives (tous les cas doivent être prévus). Il existe deux types de variables : Les **variables qualitatives** et les **variables quantitatives**. Une variable est dite qualitative si elle ne peut être mesurée ou quantifiée, mais peut être classée en catégories comme le sexe, la race, l'espèce, le niveau scolaire,... Une variable est de type quantitatif si elle peut être mesurée ou quantifiée, comme le poids, la hauteur, le revenu, le nombre d'enfants, le nombre de pannes.

Les variables qualitatives sont constituées de deux sous-classes :

- Les variables qualitatives **nominales** : ce sont celles dont les modalités ne peuvent qu'être constatées, nommées.  
**Exemple** : Le sexe (masculin, féminin), la nationalité (Canadienne, Française, Marocaine,...), les cours suivis durant une session (mathématiques, anglais, philosophie,...) ...
- Les variables qualitatives **ordinales**. ce sont les variables qualitatives dont les modalités appellent naturellement un ordre dans leur rangement. **Exemple** : Le niveau scolaire (primaire, secondaire, collégial, universitaire), le comportement lors d'une réception (incongru, correct, parfait,...), ...

Les variables quantitatives sont elles aussi subdivisées en deux sous-classes :

- Les variables quantitatives **discrètes** : ce sont celles dont les modalités sont des valeurs isolées.  
**Exemple** : Le nombre de pannes, le nombre d'accidents, le nombre d'enfants,...
- Les variables quantitatives **continues**, ce sont celles dont les modalités forment un continuum. Ce sont celles qui peuvent prendre n'importe quelle valeur dans un intervalle raisonnable.  
**Exemple** : La taille, le poids, le revenu,...

## 1.2 Les échelles de mesures.

Pour les variables qualitatives, il existe deux échelles de mesure. **L'échelle nominale** qui s'adresse aux variables qualitatives nominales, elle ne sert qu'à coller une étiquette aux unités statistiques, elle ne les classe pas sur une échelle à une dimension.

### Exemple 1.3.1 :

- $X = \text{sexe}$ , alors  $X$  est une variable qualitative nominale et son échelle est nominale.
- $Y = \text{le numéro du dossard d'un joueur de hockey}$ . Même si  $Y$  prend des valeurs numériques, ce n'est qu'une variable nominale et son échelle est nominale. Car on peut tout aussi bien mettre des lettres sur leur dossard ou des dessins.  
L'autre **échelle est l'échelle ordinale** et s'adresse aux variables qualitatives ordinales, on l'appelle comme cela car il y a un ordre entre ses modalités.

### Exemple 1.3.2 :

- $X = \text{le niveau scolaire d'une personne adulte}$ , alors ses modalités peuvent être : primaire, secondaire, collégial, universitaire. Il y a un ordre chronologique entre ces modalités.
- $Y = \text{la note finale obtenue dans un cours de statistique}$ , ses modalités seront : F, E, D, C, B, A ou A+. Il y a un ordre de mérite entre ces modalités.

Pour les variables quantitatives, il existe aussi deux types d'échelles, la première échelle est **l'échelle d'intervalle**. On l'appelle comme ça car la seule opération possible est la différence. On reconnaît une échelle d'intervalle par l'absence du zéro absolu (c'est-à-dire que si  $X=0$ , cela ne veut pas dire absence de ce qu'on mesure).

### Exemple 1.3.3 :

- $T = \text{la température en degrés Celsius}$ . Le jour où  $T=0^\circ\text{C}$ , ça ne veut pas dire absence de température. Si on considère deux journées où la température est respectivement

égale à 10 et 30 degrés, ça veut seulement dire qu'il y a un écart de 20 degrés entre ces deux journées. Si on prend deux sots d'eau où la température est respectivement égale à 35 et 45 degrés, si on les mélange, on ne va pas obtenir une eau chauffée à 80 degrés. Alors l'échelle de cette variable est une échelle d'intervalle.

- $X$ =la date de naissance, si on est en 2010 et qu'on considère une personne née en 1950 et une autre née en 1980, tout ce qu'on peut dire est qu'il y a une différence d'âge de 30 ans entre elles. On ne peut pas dire que l'une est deux fois plus âgée que l'autre, car l'année prochaine ce ne serait plus vrai. Alors l'échelle de cette variable est une échelle d'intervalle.

L'autre échelle est l'**échelle de rapports**. C'est l'échelle la plus maniable, la plus riche. Elle admet un zéro absolu, c'est-à-dire si la variable est nulle, cela signifie l'absence de ce qu'on mesure. On peut faire toutes les opérations algébriques avec une telle échelle.

### Exemple : 1.3.4 :

- $X$ =le revenu familial annuel (en dollars), si  $X=0$  cela veut dire qu'il n'y a pas eu de revenu. Si on prend deux familles dont le revenu respectif est de 30 000 et 120 000 dollars, on peut dire qu'il y a un écart de 90 000 dollars entre ces deux revenus, on peut aussi dire que la deuxième famille gagne 4 fois plus que la première. Si on additionne ces deux revenus, on aura un revenu global de 150 000 dollars. Alors l'échelle de cette variable est une échelle de rapports.
- $Y$ =le nombre d'enfants dans un ménage. Si  $Y=0$  cela veut dire que cette famille n'a pas d'enfant. On peut faire toutes les opérations algébriques avec les modalités de cette variable, donc son échelle est une échelle de rapports.

## 1.3 Les tableaux et graphiques.

Dans ce paragraphe on va détailler comment résumer l'information contenue dans une série de données soit par des tableaux ou des graphiques. On va commencer par les variables qualitatives.

### 1.3.1 Cas de variables qualitatives.

On va considérer deux exemples où on a des variables qualitatives observées sur un échantillon et suivre le traitement possible de ces données.

**Exemple 1.4.1 :** On a pris un échantillon de 50 achats de boissons non-alcoolisées achetées dans une grande surface, en notant par :

CC=Coca-Cola; S=Sprite; CL=Coke-Light; P=Perrier; PC=Pepsi-Cola. On a obtenu les résultats suivants.

CC S PC CL CC CC PC CL CC CL CC CC CC CL PC CC  
 CC P P S CC CL PC CL PC CC PC PC CC PC CC CC PC  
 P PC PC S CC CC CC S P CL P PC CC PC S CC CL

Alors ici la variable est  $X$ =Boisson non-alcoolisée, qui est une variable qualitative nominale. Pour présenter ces données sous forme de tableau, on dresse un tableau, dans la première colonne on énumère les cinq modalités de la variable, dans la seconde colonne on donne la **fréquence absolue** ou l'effectif de chacune des modalités (c'est-à-dire le nombre de fois que cette modalité se répète dans l'échantillon) et dans la troisième colonne, on donne la **fréquence relative** de chacune des modalités. La

**fréquence relative** d'une modalité étant égale à sa fréquence absolue divisée par la taille de l'échantillon. Ce qui donne :

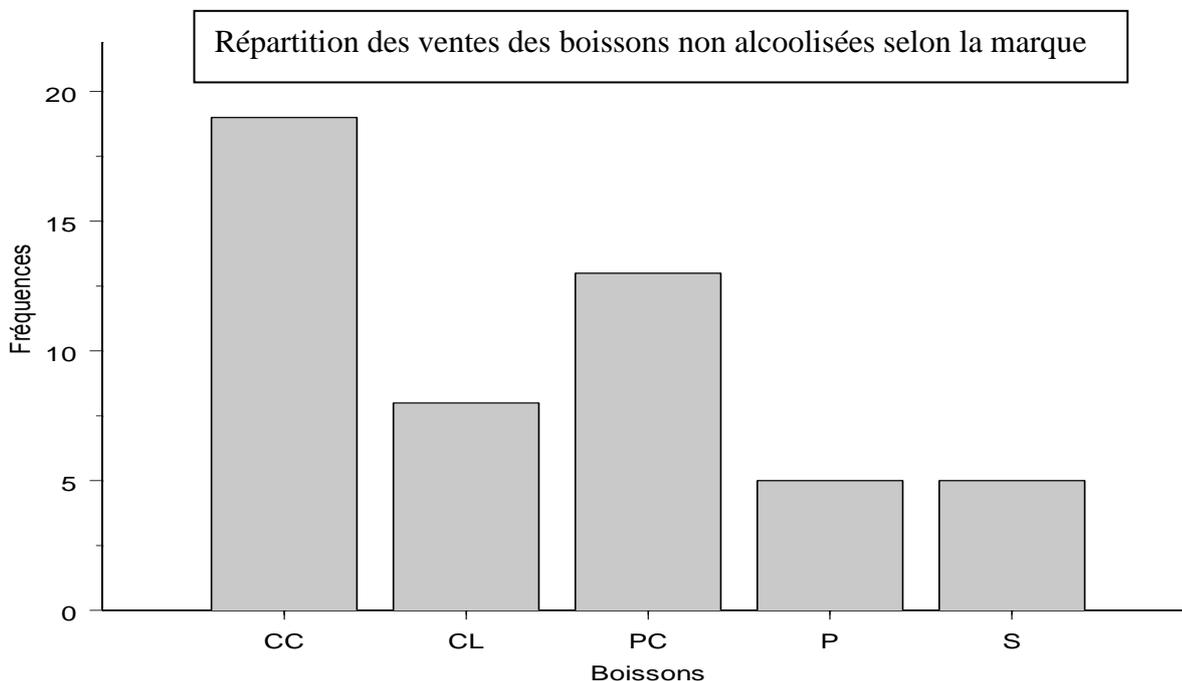
<b>Tableau des fréquences des boissons non-alcoolisées</b>		
X=Boisson	Fréquences absolues	Fréquences relatives
CC	19	0,38
CL	8	0,16
PC	13	0,26
P	5	0,10
S	5	0,10
Total	n=50	1
Source : données fictives.		

Ce tableau s'appelle **tableau de fréquences** de la variable.

**Remarque :** Pour une présentation complète des tableaux et graphiques, on doit mettre le titre en haut et la source des données en bas.

En ce qui concerne la représentation graphique, on va donner deux graphiques qui résument la même information contenue dans le tableau des fréquences.

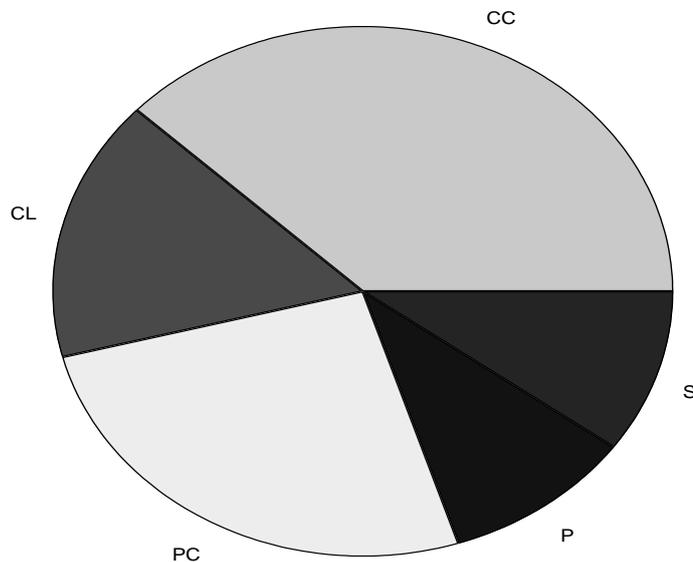
- Le diagramme à barres (horizontales ou verticales). Où on met sur un axe les modalités de la variable et sur l'autre axe les fréquences absolues ou les fréquences relatives.



**Remarque :** Les largeurs des barres doivent être les mêmes pour une belle esthétique du graphique, ainsi que la distance entre les bandes. On peut aussi ajouter les fréquences absolues au dessus des bandes.

- Le deuxième graphique qu'on peut faire est le **diagramme à secteurs** (ou circulaire) qui est une sorte de tarte où chaque modalité occupe une partie qui reflète sa fréquence relative.

Diagramme circulaire donnant la répartition des boissons non alcoolisées selon la marque



**Exemple 1.4.1.2 :** Lors d'une enquête de satisfaction de la clientèle, une compagnie de courtage a demandé à un échantillon de 60 clients d'indiquer leur degré de satisfaction vis-à-vis de leur conseiller financier, sur une échelle de 1 à 7, le 1 correspondant à « pas du tout satisfait » et le 7 correspondant à « extrêmement satisfait ». On a obtenu les résultats suivants :

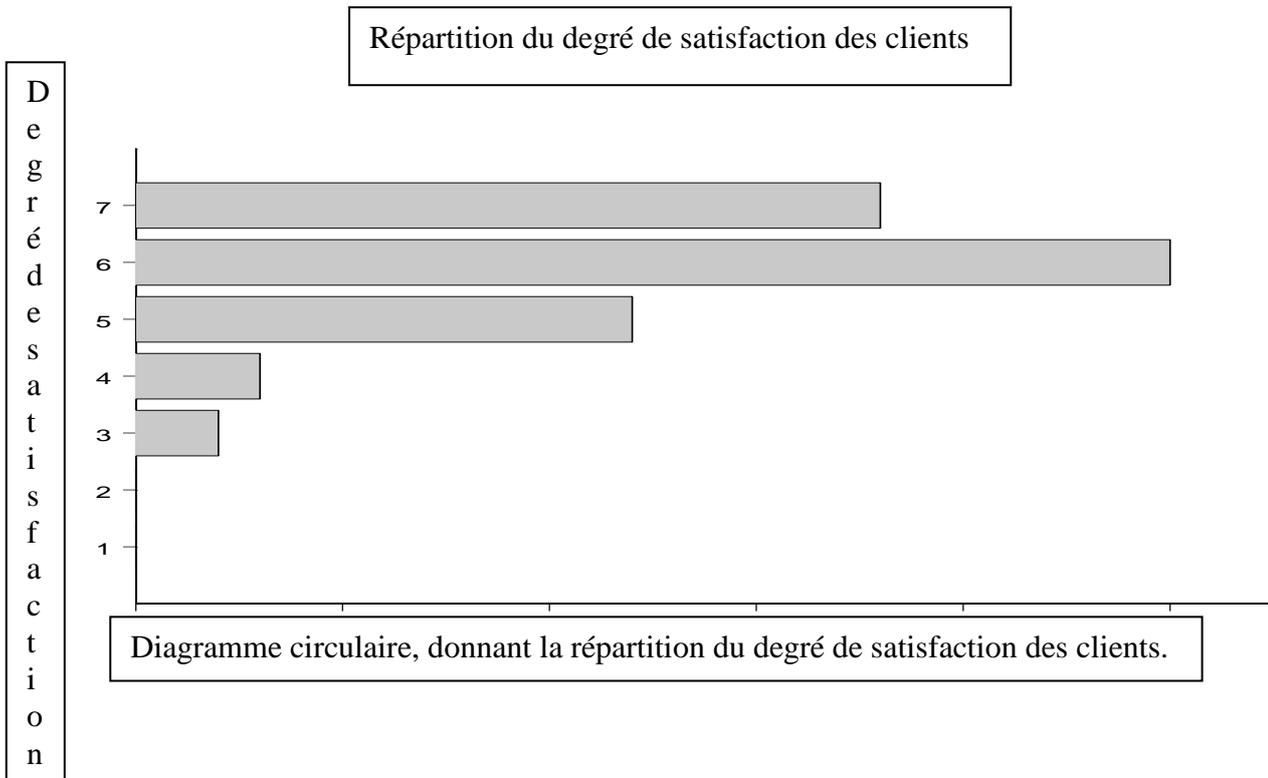
5 7 6 6 7 5 5 7 3 6 7 7 6 6 6 5 5 6 7 7  
 6 6 4 4 7 6 7 6 7 6 5 7 5 7 6 4 7 5 7 6  
 6 5 3 7 7 6 6 6 6 5 5 6 6 7 7 5 6 6 6 6

Ici la variable, « degré de satisfaction » est une variable qualitative ordinale. On peut résumer l'information contenue dans ces données sous forme d'un tableau de fréquences ce qui donne :

Tableau des fréquences du degré de satisfaction des clients.		
Degré de satisfaction	Fréquences absolues	Fréquences relatives
1	0	0,0000
2	0	0,0000
3	2	0,0333
4	3	0,0500
5	12	0,2000
6	25	0,4167
7	18	0,3000
Total	n=60	1,0000

Source : Données fictives.

En ce qui concerne la représentation graphique, les mêmes graphiques qu'on a utilisés pour une variable qualitative nominale font l'affaire. Ce qui donne :



### 1.4.2 Cas de variables quantitatives.

Le traitement des variables quantitatives discrètes étant différent de celui des variables quantitatives continues, on va donc réserver un sous paragraphe à chacune d'elles.

#### 1.4.2.1 : Cas des variables quantitatives discrètes.

Soit  $X$  une variable quantitative discrète dont le nombre de modalités n'est pas trop grand. Alors on peut dresser un tableau des fréquences comme celui utilisé pour les variables qualitatives auquel on peut ajouter une colonne supplémentaire où on met les fréquences relatives cumulées au fur et à mesure qu'on ajoute une modalité de la variable. En ce qui concerne la représentation graphique, un seul graphique s'associe avec les variables quantitatives discrètes : **le diagramme à bâtons.**

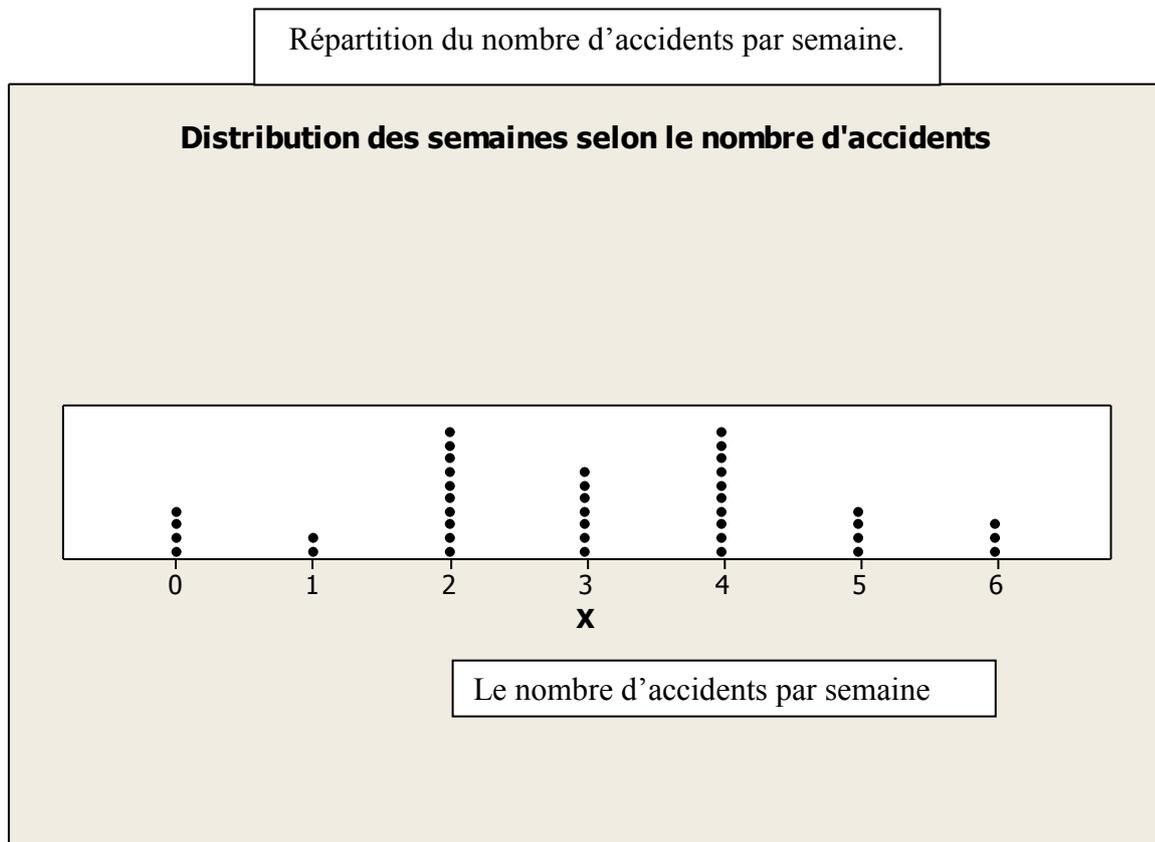
**Exemple 1.4.2.1.1** : Un inspecteur en contrôle de qualité a extrait de sa base de données, un échantillon de 40 semaines où il a noté  $X$ , le nombre d'accidents de travail enregistrés par semaine. Il a obtenu les résultats suivants :

2 0 4 2 2 1 3 2 0 5 4 3 2 4 5 6 6 4 2 0  
3 4 4 2 6 2 4 3 0 4 3 4 3 3 5 5 4 2 2 1

On peut donc dresser le tableau des fréquences suivant.

Tableau des fréquences du nombre d'accidents par semaine			
Le nombre d'accidents par semaine.	Fréquences absolues	Fréquences relatives	Fréquences relatives cumulées
0	4	0,100	0,100
1	2	0,050	0,150
2	10	0,250	0,400
3	7	0,175	0,575
4	10	0,250	0,825
5	4	0,100	0,925
6	3	0,075	1,000
Total	n=40	1,000	

Quant au diagramme à bâtons, on obtient quelque chose comme :



**Remarque :** Les bâtons ne doivent pas avoir d'épaisseur, car la variable prend exactement les valeurs 0, 1, 2,... On peut ajouter les effectifs ou les fréquences relatives sur les bâtons.

#### 1.4.2: Cas de variables quantitatives continues.

Considérons maintenant un échantillon de données provenant d'une variable quantitative continue ou discrète avec un grand nombre de modalités. Il est donc inconcevable de dresser un tableau où on énumère les modalités d'une telle variable, il

serait non analysable. Il faut donc grouper ces données en classes de valeurs. Deux questions se posent alors :

- Combien de classes faut-il former ?
- Quelles seront les largeurs de chacune des classes ?

La réponse à la première question, dépend de la taille de l'échantillon, le nombre de classe à former est donné par la formule de Sturges suivante :

*Le nombre de classes:*  $K = 1 + \frac{10}{3} \text{Log}(n)$ . Ainsi, par exemple, si  $n=150$ , il faut former  $K = 1 + \frac{10}{3} \text{Log}(150) = 8,2536 \cong 9$  (on arrondit à l'entier immédiatement supérieur). Une fois qu'on sait combien de classes à former. On essaie de former des classes de même amplitude (largeur) et cette amplitude sera égale à

$$A = \frac{\text{la plus grande observation} - \text{la plus petite observation}}{K} = \frac{X_{\max} - X_{\min}}{K}$$

On arrondit cette amplitude selon les données pour avoir des bornes de classes faciles à manipuler.

**Exemple 1.4.2.2.1 :** Soit X, les recettes quotidiennes(en dollars) d'un petit magasin. On a sélectionné un échantillon de taille  $n=40$  jours au hasard qui ont donné les résultats suivants :

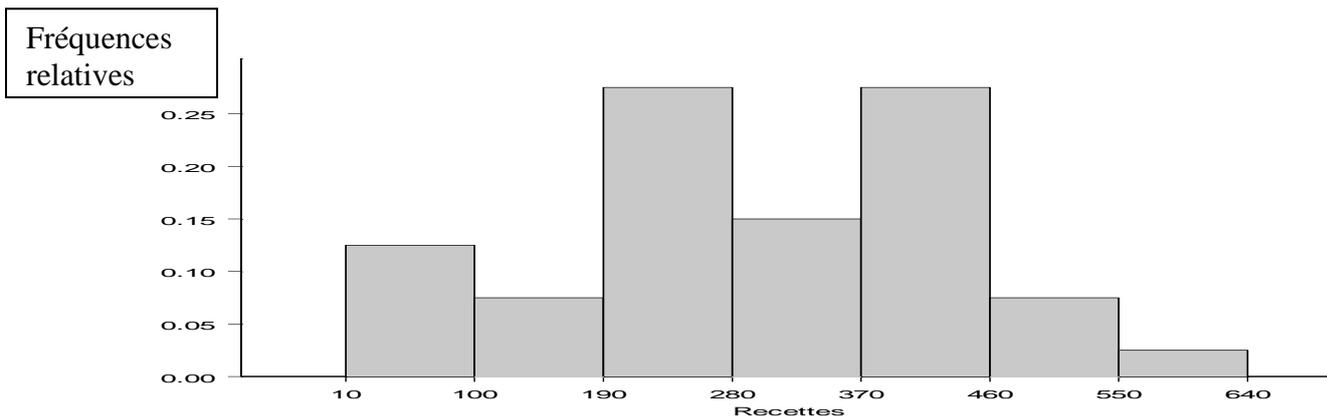
16,00 58,50 68,20 78,00 79,45 142,20 145,3 186,70 209,05 216,75  
 219,70 247,75 249,10 256,00 257,15 262,35 268,60 269,60 270,15 284,45  
 319,00 332,00 343,29 350,75 354,90 372,60 383,20 389,20 404,55 420,20  
 428,50 432,40 444,60 446,80 456,10 458,10 493,95 511,95 521,05 621,35

Le nombre de classe à former est  $K = 1 + \frac{10}{3} \log(40) = 6,34 \cong 7$  classes d'amplitude chacune égale à  $A = \frac{621,35-16}{7} = 86,48 \cong 90$ . Cette amplitude est arrondie à 90. Ce qui donne le tableau des fréquences suivant, où les classes sont des intervalles fermés à gauche et ouverts à droite sauf le dernier qui est un intervalle fermé des deux côtés.

Quand aux graphiques, on va ici privilégier trois graphiques pour les variables quantitatives continues.

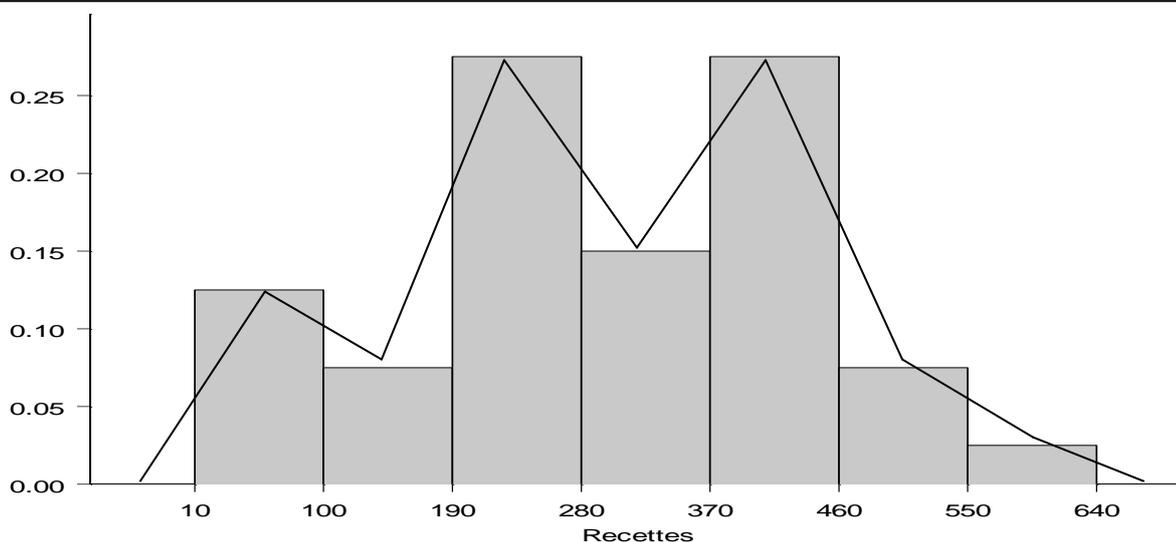
- **L'histogramme**, qui est une suite de rectangles juxtaposés les uns aux autres dressés au-dessus de chacune des classes, dont la largeur est égale à l'amplitude de la classe (prise comme unité de mesure) et dont la surface reflète la fréquence relative de la classe qu'il représente.

Histogramme donnant la répartition des 40 semaines en fonction des recettes hebdomadaires



- Le polygone des fréquences**, qui consiste à joindre le milieu des sommets des rectangles d'un histogramme par une ligne en zig-zag et cette ligne se ferme en ajoutant aux deux extrémités deux classes fictives de même amplitude que les autres, comme ça la surface délimitée par l'histogramme est identique à celle délimitée par le polygone des fréquences. Le polygone de fréquences est très utile quand on veut comparer le comportement de la même variable mesurée sur plusieurs groupes (on peut penser à comparer le revenu des hommes et des femmes) ou la même variable mesurée sur le même échantillon à différents instants (on peut comparer le poids du même groupe à différents moments d'une diète).

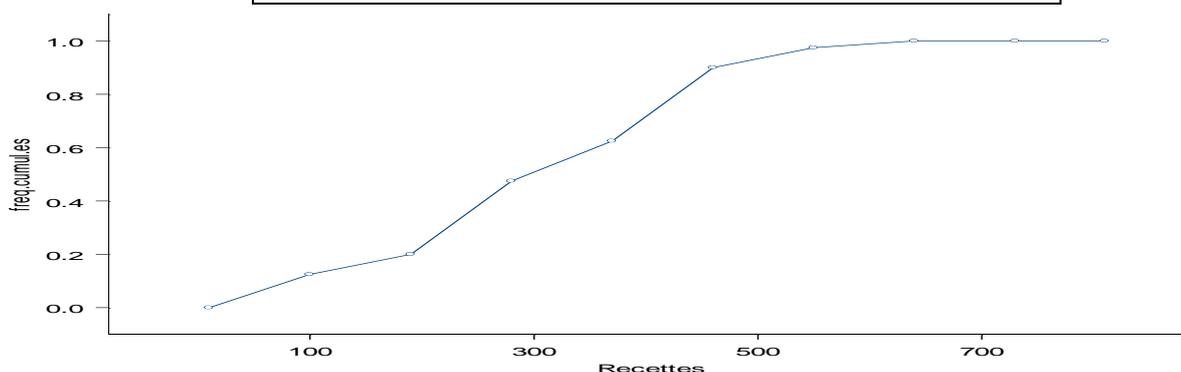
Polygone des fréquences donnant la répartition des 40 semaines selon les recettes hebdomadaires.



- La courbe des fréquences cumulées (Ogive).**

Comme son nom l'indique, elle consiste à tracer le graphique des fréquences cumulées, en mettant les limites des classes sur l'axe horizontal et les fréquences cumulées sur l'axe vertical, ces dernières se cumulant à la fin de chacune des classes. Ce graphique aura l'allure d'une courbe croissante variant entre 0 et 1.

Ogive de la répartition des 40 semaines selon les recettes



**Remarque :** Lorsque les classes ne sont pas de même amplitude, il faut se rappeler que la surface du rectangle d'un histogramme étant égale à sa fréquence relative à la classe associée à ce rectangle, alors si la largeur de cette classe par exemple est le double de l'amplitude de base, la hauteur du rectangle doit être divisée par deux.

### 1.5 : Les mesures de tendance centrale

On appelle mesures de tendance centrale, des valeurs de la variable susceptibles de nous donner une idée sur la donnée qui occupe le centre d'une série statistique. On va décrire dans ce paragraphe, les trois plus importantes mesures de tendance centrale que sont **le mode, la moyenne et la médiane**.

#### 1.5.1: Le mode

On appelle le mode d'une variable  $X$ , la valeur de la variable qui a la plus grande fréquence et on le note  $Mo(X)$ . Le mode est une importante mesure de tendance centrale pour les variables qualitatives nominales.

**Remarque :** Une distribution peut avoir un seul mode et on dit qu'elle est unimodale, ou plusieurs modes et on dit qu'elle est multimodale.

**Exemple 1.5.1:** Si on reprend l'exemple des boissons non-alcoolisées, on avait le tableau des fréquences suivant :

Tableau des fréquences des boissons non-alcoolisées		
X=Boisson	Fréquences absolues	Fréquences relatives
CC	19	0,38
CL	8	0,16
PC	13	0,26
P	5	0,10
S	5	0,10
Total	n=50	1

Alors, le mode de cette variable est  $Mo(X)=Coca-Cola (CC)$ , cela signifie que dans cet échantillon, la boisson la plus fréquemment achetée est Coca-Cola.

**Exemple 1.5.2:** En reprenant l'exemple des recettes quotidiennes d'un petit magasin, où la variable est quantitative continue avec des données groupées en classes, on avait le tableau des fréquences suivant :

X=les recettes	Fréquences absolues	Fréquences relatives
[10 ; 100[	5	0,125
[100 ; 190[	3	0,075
[190 ; 280[	11	0,275
[280 ; 370[	6	0,150
[370 ; 460[	11	0,275
[460 ; 550[	3	0,075
[550 ; 640]	1	0,025
Total	n=40	1,000

Ici, on voit qu'il y a deux classes qui ont les plus hautes fréquences, on les appelle des classes modales. Alors on est en présence d'une distribution de données bimodale, et les deux modes sont les milieux des deux classes modales, à savoir  $Mo(X)=235$  et  $Mo(X)=415$ . Cela veut dire que dans cet échantillon les recettes quotidiennes les plus fréquentes sont soit de 235\$ ou de 415\$. Il y a des auteurs qui font des interpolations à l'intérieur des classes modales pour trouver le mode, on estime que c'est un effort inutile, vu que dans le cas d'une variable quantitative le mode joue un rôle très marginal. On voit que le mode d'une variable est une mesure de tendance centrale facile à déterminer et s'applique à tous les types de variables, mais sa portée comme mesure d'analyse est très limitée.

### 1.5.2 : La moyenne.

La moyenne arithmétique ou simplement la moyenne est la mesure de tendance centrale la plus connue. Elle ne s'applique qu'aux variables quantitatives. On va décrire la méthode pour calculer la moyenne d'une variable quantitative selon que les données sont en vrac, groupées par valeurs ou groupées par classes.

#### 1.5.2.1 : Les données en vrac.

Soit X une variable quantitative dont les valeurs observées sur un échantillon forment une série en vrac  $x_1, x_2, \dots, x_n$  alors la moyenne de cet échantillon est

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

**Exemple 1.5.2.1.1 :** Un commerçant a l'habitude de noter dans son registre le nombre de clients qui se présentent quotidiennement à son magasin. On a pris un échantillon de taille 10 de ce registre et on trouvé les valeurs suivantes :

120 105 90 201 196 65 88 163 103 116

Alors dans cet échantillon le nombre moyen des clients qui se présentent à ce magasin par jour est donné par la formule suivante :

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{120 + 105 + \dots + 116}{10} = 124,7 \text{ clients par jour.}$$

### 1.5.2: Les données groupées par valeurs.

Soit X une variable quantitative discrète dont les données se présentent sous forme d'un tableau où elles sont classées par valeurs, supposons que la taille de l'échantillon est n et qu'il y a k valeurs différentes pour cette variable. Alors la moyenne d'un tel échantillon de données est :

$$\bar{x} = \frac{\sum(\text{valeur}) * (\text{sa fréquence absolue})}{\text{taille de l'échantillon}} = \frac{\sum_{i=1}^k x_i f_i}{n}$$

**Exemple 1.5.3. :** Reprenons les données de l'exemple 1.4.2.1.1, où X est le nombre d'accidents de travail par semaine. On avait le tableau de données suivant :

Tableau des fréquences du nombre d'accidents par semaine	
X	Fréquences absolues
0	4
1	2
2	10
3	7
4	10
5	4
6	3
Total	n=40

Alors la moyenne de cet échantillon est égale à

$$\bar{x} = \frac{(0)*(4)+(1)*(2)+\dots+(6)*(3)}{40} = 3,025 \text{ accidents par semaine.}$$

### 1.5.3 : Les données groupées par classes.

Supposons qu'on est devant un tableau où les données provenant d'un échantillon sont groupées par classes. Alors pour calculer la moyenne de cet échantillon, on va utiliser une formule approximative, où chaque classe est assimilée à son centre et on utilise la même formule que pour le cas où les données sont groupées par valeurs. Si on note par  $m_i$ , le milieu de la ième classe et qu'on suppose que la taille de l'échantillon est n et qu'il y a k classes, alors la moyenne de l'échantillon est :

$$\bar{x} = \frac{\sum_{i=1}^k m_i f_i}{n}$$

**Exemple 1.5.2.4 :** En reprenant l'exemple 1.4.2.2.1 où X est la recette quotidienne d'un petit magasin, on avait le tableau suivant auquel on a ajouté une colonne à gauche contenant le milieu des classes :

$m_i$	X=les recettes	Fréquences absolues
55	[10 ; 100[	5
145	[100 ; 190[	3
235	[190 ; 280[	11
325	[280 ; 370[	6
415	[370 ; 460[	11
505	[460 ; 550[	3
595	[550 ; 640]	1
	Total	n=40

Alors la moyenne de cet échantillon est :

$$\bar{x} = \frac{\sum_{i=1}^k m_i f_i}{n} = \frac{(55) * (5) + \dots + (595) * (1)}{40} = 298 \text{ dollars.}$$

#### 1.5.2.4 : Les propriétés d'une moyenne échantillonnale.

Soit X une variable quantitative dont la moyenne échantillonnale est  $\bar{x}$  et soit Y une autre variable quantitative transformée linéaire de X, c'est-à-dire que  $Y = a + b * X$  où a et b sont des constantes réelles. Alors la moyenne échantillonnale de Y sera égale à  $\bar{y} = a + b * \bar{x}$ .

On dit que la moyenne conserve la transformation linéaire entre les variables.

**Exemple 1.5.2.4.1 :** Soit X, le nombre d'heures qu'un étudiant travaille à temps partiel par semaine. Supposons qu'à partir d'un échantillon d'étudiants, on a pu trouver qu'en moyenne le nombre d'heures travaillées par ces étudiants est égale à  $\bar{x} = 14,5$  heures/semaine. Si le salaire horaire est de 10\$ et que les patrons de ces étudiants leur offrent 30\$ par semaine pour leurs déplacements, quel est le gain net moyen hebdomadaire de ces étudiants ? Posons Y, le gain net hebdomadaire de ces étudiants alors  $Y = 30 + 10 * X$ , donc le gain moyen hebdomadaire de cet échantillon d'étudiants est égal à

$$\bar{y} = 30 + 10 * \bar{x} = 30 + 10 * 14,5 = 175\$.$$

#### 1.5.3 : La médiane.

La médiane est la valeur de la variable qui divise l'échantillon en deux groupes d'égal effectif. Il y a 50% des données qui sont inférieures ou égales à la médiane et 50% des données qui sont supérieures ou égales à la médiane. La médiane se calcule pour des variables qualitatives ordinales et pour des variables quantitatives. On note la médiane d'une variable X par Med(X) ou par  $\tilde{x}$ . Dans ce qui suit on va décrire les façons de calculer une médiane dans les différents cas possibles.

### 1.5.3.1 : Cas d'une variable qualitative ordinale.

Puisque les modalités d'une telle variable sont déjà ordonnées par nature, alors pour déterminer la médiane, on calcule  $l = (50\%) * n$ , et donc

$$Med(X) = \tilde{x} = \begin{cases} \frac{x_{(l)} + x_{(l+1)}}{2} & \text{si } l \text{ est un entier} \\ x_{[l]+1} & \text{si } l \text{ n'est pas un entier} \end{cases}$$

Où  $x_{[l]+1}$  signifie, l'observation occupant le rang immédiatement supérieur à  $l$ .

**Exemple 1.5.3.1.1 :** Reprenons les données de l'exemple 1.4.1.2, où X est le degré de satisfaction de la clientèle, on avait le tableau suivant :

Tableau des fréquences du degré de satisfaction des clients.	
Degré de satisfaction	Fréquences absolues
1	0
2	0
3	2
4	3
5	12
6	25
7	18
Total	n=60

Ici,  $n=60$  et  $l = (50\%) * n = 30$  est un entier, alors

$Med(X) = \tilde{x} = \frac{x_{(30)} + x_{(31)}}{2} = \frac{6+6}{2} = 6$ . Le degré de satisfaction médian de la clientèle est égal à 6. Ce qui veut dire que dans cet échantillon 50% des clients ont un degré de satisfaction de 6 ou moins et l'autre 50% un degré de satisfaction de 6 ou plus.

### 1.5.3.2 : Cas de données quantitatives en vrac ou groupées par valeurs.

On doit d'abord ordonner les données par ordre croissant avant d'appliquer la même procédure que pour les variables qualitatives ordinales. Ci-après nous donnerons un exemple pour chacun de ces deux cas.

**Exemple 1.5.3.2.1 :** Reprenons les données de l'exemple 1.5.2.1.1 où la variable est le nombre de clients qui se présentent quotidiennement au magasin. On avait des données en vrac :

120 105 90 201 196 65 88 163 103 116

En les ordonnant, on aura : 65 88 90 103 105 116 120 163 196 201.

Ici,  $n=10$  et  $l = (50\%) * n = 5$  est un entier, alors

$Med(X) = \tilde{x} = \frac{x_{(5)} + x_{(6)}}{2} = \frac{105 + 116}{2} = 110,5$ . Ce qui veut dire qu'à partir de cet échantillon, on peut affirmer que dans 50% des journées, ce magasin reçoit 110 clients ou moins par jour et dans l'autre 50% des journées, il reçoit 110 clients ou plus.

**Exemple 1.5.3.2 :** Reprenons les données de l'exemple 1.4.2.1.1, où X est le nombre d'accidents de travail par semaine. On avait le tableau de données où les modalités de la variable sont groupées par valeurs, qu'on va changer un peu en ajoutant une donnée supplémentaire :

<b>Tableau des fréquences du nombre d'accidents par semaine</b>	
Nombre d'accidents par semaine	Fréquences absolues
0	4
1	2
2	10
3	7
4	10
5	4
6	4
Total	n=41

Ici,  $n=41$  et  $l = (50\%) * n = 20,5$  n'est pas un entier, alors

$Med(X) = \tilde{x} = x_{[20,5]+1} = x_{(21)}$  = l'observation qui occupe la 21ème position = 3.

C'est-à-dire que dans cet échantillon, dans 50% des semaines, on observe 3 accidents ou moins par semaine et l'autre 50% des semaines, on observe 3 accidents ou plus par semaine.

### 1.5.3.3 : Cas de données groupées par classes.

Dans le cas où on dispose d'un tableau de fréquences complet (incluant les fréquences cumulées) des données groupées par classes. Il faut d'abord déterminer la classe médiane, qui est la classe où les fréquences cumulées dépassent pour la première fois 50%. Cette classe aura la forme :

$C_m = [b_{inf}; b_{sup}[$ , alors on obtient la médiane par interpolation à l'intérieur de cette classe médiane et on obtient la formule suivante :

$$Med(X) = \tilde{x} = b_{inf} + \frac{(0,5 - F_{(m-1)})}{f_{r,m}} * A_m \text{ où}$$

$b_{inf}$  = borne inférieure de la classe médiane.

$F_{(m-1)}$  = la fréquence cumulée avant la classe médiane.

$f_{r,m}$  = la fréquence relative de la classe médiane.

$A_m$  = l'amplitude de la classe médiane.

**Exemple 1.5.3.3.1 :** En reprenant les données où X donne la recette quotidienne d'un petit magasin, on retrouve le tableau des fréquences suivant :

X=les recettes	Fréquences absolues	Fréquences relatives	Fréquences relatives cumulées
[10 ; 100[	5	0,125	0,125
[100 ; 190[	3	0,075	0,200
[190 ; 280[	11	0,275	0,475
[280 ; 370[	6	0,150	0,625
[370 ; 460[	11	0,275	0,900
[460 ; 550[	3	0,075	0,975
[550 ; 640]	1	0,025	1,000
Total	n=40	1,000	

Alors ici, la classe médiane est  $C_m = [b_{inf}; b_{sup}[ = [280 ; 370[$

$$b_{inf} = 280 \quad F_{(m-1)} = 0,475$$

$$f_{r,m} = 0,150 \quad A_m = 90 \quad \text{ce qui donne une médiane égale à :}$$

$$\tilde{x} = b_{inf} + \frac{(0,5 - F_{(m-1)})}{f_{r,m}} * A_m = 280 + \frac{(0,5 - 0,475)}{0,150} * 90 = 295\$$$

Ce qui veut dire qu'en se basant sur cet échantillon de données, 50% des recettes quotidiennes de ce petit magasin sont inférieures ou égales à 295\$ et les autres 50% sont supérieures ou égales à 295\$.

**Remarque 1 :** Le calcul de la médiane est basé sur l'ordre des observations et non sur leur valeur. Contrairement à la moyenne, la médiane est insensible aux données extrêmes. Dans le cas où les données sont très différentes, la médiane est une meilleure mesure de tendance centrale.

**Remarque 2 :** Si pour une variable X quantitative les 3 mesures de tendance centrale sont presque égales, on dit alors que la variable est symétrique et alors n'importe laquelle de ces mesures peut être utilisée comme mesure de cette tendance centrale. S'il y a un grand écart entre ces mesures alors c'est la médiane qu'on doit privilégier.

## 1.6 : Les mesures de position.

On a déjà parlé de la médiane comme mesure de tendance centrale, mais elle est aussi une mesure de position car elle permet de diviser une série d'observations en deux groupes chacun contenant 50% de données. On va définir d'autres mesures de position qui permettent d'autres découpages d'une série d'observations.

**1.6.1 : Les quartiles.** Lorsqu'on veut diviser les données en quatre groupes, chacun contenant 25% des observations, on utilise des mesures appelées quartiles.

$Q_1$  = le 1<sup>er</sup> quartile, à sa gauche il y a 25% des observations, qu'on note  $Q_1$ .

$Q_2$  = le 2<sup>ème</sup> quartile, coïncide avec la médiane, qu'on note  $Q_2 = Med(X)$ .

$Q_3$  = le 3<sup>ème</sup> quartile, à sa gauche il y a 75% des observations, qu'on note  $Q_3$ .

On va décrire la façon de les calculer, dans les 3 cas possibles pour une variable quantitative.

**1.6.1.1 : Les données en vrac.** On suit les étapes suivantes.

Étape 1 : On ordonne les données par ordre croissant.

Étape 2 : On calcule l'indice  $l = (i\%) * (n)$  où  $i$  est le pourcentage correspondant à la mesure voulue et  $n$  est le nombre d'observations.

Étape 3 : (a) si  $l$  n'est pas un entier, alors le  $i$ ème quartile est égal à l'observation occupant la position immédiatement supérieure à  $l$ .

(b) si  $l$  est un entier, alors le  $i$ ème quartile est la moyenne des observations occupant les positions  $l$  et  $(l + 1)$ .

**Exemple 1.6.1.1.1** :  $n=12$  et les observations sont :

-2 -3 10 12 120 11 4 8 6 13 130 200.

Étape 1 : -3 -2 4 6 8 10 11 12 13 120 130 200.

Étape 2 : Si on veut déterminer  $Q_1$ , on calcule  $l_1 = (25\%) * (n) = 3$ .

Si on veut déterminer  $Q_2$ , on calcule  $l_2 = (50\%) * (n) = 6$ .

Si on veut déterminer  $Q_3$ , on calcule  $l_3 = (75\%) * (n) = 9$ .

Étape 3 : Puisque  $l_1$  est un entier alors  $Q_1 = \frac{\text{la 3ème obs} + \text{la 4ème obs}}{2} = \frac{4+6}{2} = 5$ .

Puisque  $l_2$  est un entier alors  $Q_2 = \frac{\text{la 6ème obs} + \text{la 7ème obs}}{2} = \frac{10+11}{2} = 10,5$ .

Puisque  $l_3$  est un entier alors  $Q_3 = \frac{\text{la 9ème obs} + \text{la 10ème obs}}{2} = \frac{13+120}{2} = 66,5$ .

**Exemple 1.6.1.1.2** :  $n=10$  et les observations sont :

3 10 12 8 6 100 15 6 3 14.

Étape 1 : 3 3 6 6 8 10 12 14 15 100

Étape 2 : Si on veut déterminer  $Q_1$ , on calcule  $l_1 = (25\%) * (n) = 2,5$ .

Si on veut déterminer  $Q_2$ , on calcule  $l_2 = (50\%) * (n) = 5$ .

Si on veut déterminer  $Q_3$ , on calcule  $l_3 = (75\%) * (n) = 7,5$ .

Étape 3 : Puisque  $l_1$  n'est pas un entier alors  $Q_1 = \text{la 3ème observation} = 6$ .

Puisque  $l_2$  est un entier alors  $Q_2 = \frac{\text{la 5ème obs} + \text{la 6ème obs}}{2} = \frac{8+10}{2} = 9$ .

Puisque  $l_3$  n'est pas un entier alors  $Q_3 = \text{la 8ème observation} = 14$ .

**Remarque** : La procédure décrite pour trouver les quartiles est une convention parmi d'autres. Il n'y a pas d'accord général sur la méthode à utiliser pour déterminer les quartiles. Si vous utilisez des logiciels, les valeurs trouvées diffèrent d'un logiciel à l'autre. Par exemple, si on prend la série en vrac suivantes : 1 3 6 10 15 21 28 36, alors la calculatrice TI-83 et plus et les logiciels suivants donnent :

logiciel	$Q_1$	$Q_2$	$Q_3$
SPSS	3,75	12,5	26,25
SAS	4,5	12,5	24,5
STATDISK	4,5	12,5	24,5
Excel	5,25	12,5	22,75
R	5,25	12,5	22,75
Splus	5,25	12,5	22,75
Minitab	3,75	12,5	26,25
TI-83 et plus	4,5	12,5	24,5

Heureusement, dans la pratique, les échantillons sont très grands et ces fluctuations ne changent pas grand-chose dans les analyses des données.

### 1.6.1.2 : Les données groupées par valeurs.

On suit la même démarche que dans le cas des données en vrac, sauf l'étape 1 qui devient inutile, puisque les données sont en général déjà ordonnées par ordre croissant.

**Exemple 1.6.1.2.1** : En reprenant le tableau de l'exemple 1.5.3.2.2, déterminer les 3 quartiles de la variable X=le nombre d'accidents par semaine.

X	Fréquences absolues
0	4
1	2
2	10
3	7
4	10
5	4
6	4
Total	n=41

#### Réponse :

Étape 2 : Si on veut déterminer  $Q_1$ , on calcule  $l_1 = (25\%) * (n) = 10,25$ .

Si on veut déterminer  $Q_2$ , on calcule  $l_2 = (50\%) * (n) = 20,5$ .

Si on veut déterminer  $Q_3$ , on calcule  $l_3 = (75\%) * (n) = 30,75$ .

Étape 3 : Puisque  $l_1$  n'est pas un entier alors  $Q_1 = \text{la 11ème observation} = 2$ .

Puisque  $l_2$  n'est pas un entier alors  $Q_2 = \text{la 21ème observation} = 3$ .

Puisque  $l_3$  n'est pas un entier alors  $Q_3 = \text{la 31ème observation} = 4$ .

$Q_1 = 2$  signifie que dans cet échantillon, durant 25% des semaines, on a observé 2 accidents par semaine ou moins.

$Q_2 = 3$  signifie que dans cet échantillon, durant 50% des semaines, on a observé 3 accidents par semaine ou moins.

$Q_3 = 4$  signifie que dans cet échantillon, durant 75% des semaines, on a observé 4 accidents par semaine ou moins.

### 1.6.1.3 : Les données groupées par classes.

On suit la même démarche utilisée pour calculer la médiane quand les données sont groupées par classes. On détermine la classe où on a dépassé le pourcentage relatif à chaque quartile et on fait une interpolation à l'intérieur de cette classe. On aboutit à la même formule que celle de la médiane où seul le pourcentage est à adapter.

**Exemple 1.6.1.3.1** : En reprenant les données de l'exemple 1.5.3.3.1, déterminer les 3 quartiles de la variable X, soit les recettes quotidiennes d'un petit dépanneur, et interpréter ces mesures.

X=les recettes	Fréquences absolues	Fréquences relatives	Fréquences relatives cumulées
[10 ; 100[	5	0,125	0,125
[100 ; 190[	3	0,075	0,200
[190 ; 280[	11	0,275	0,475
[280 ; 370[	6	0,150	0,625
[370 ; 460[	11	0,275	0,900
[460 ; 550[	3	0,075	0,975
[550 ; 640]	1	0,025	1,000
Total	n=40	1,000	

### Réponse :

(a) Pour déterminer le premier quartile, les fréquences relatives cumulées ont dépassé 25% pour la première fois au niveau de la classe [190 ; 280[, donc

$Q_1 = 190 + \frac{(0,25-0,20)}{0,275} * 90 = 206,36\$$ . Ce qui signifie que dans cet échantillon de données, 25% des journées, les recettes quotidiennes de ce petit magasin ont été de 206,36\$ ou moins.

(b) Pour déterminer le deuxième quartile (on refait ce qu'on a déjà fait pour calculer la médiane), les fréquences relatives cumulées ont dépassé 50% pour la première fois au niveau de la classe [280 ; 370[, donc

$Q_2 = 280 + \frac{(0,50-0,475)}{0,150} * 90 = 295\$$ . Ce qui signifie que dans cet échantillon de données, 50% des journées, les recettes quotidiennes de ce petit magasin ont été de 295\$ ou moins.

(c) Pour déterminer le troisième quartile, les fréquences relatives cumulées ont dépassé 75% pour la première fois au niveau de la classe [370 ; 460[, donc

$Q_3 = 370 + \frac{(0,75-0,625)}{0,275} * 90 = 410,91\$$ . Ce qui signifie que dans cet échantillon de données, 75% des journées, les recettes quotidiennes de ce petit magasin ont été de 410,91\$ ou moins.

**Utilité des quartiles.** Les quartiles, en plus de leur utilisation comme mesures de position, s'utilisent pour détecter des données aberrantes dans toute série de données. Cette détection se fait à l'aide d'un graphique, appelé graphique en boîte (box-plot) ou hamac ou diagramme à moustache selon les auteurs. Son principe consiste à calculer les quartiles de la série et deux limites acceptables. Soient une limite inférieure  $L_{inf} = Q_1 - 1,5 * (Q_3 - Q_1)$  et une limite supérieure  $L_{sup} = Q_3 + 1,5 * (Q_3 - Q_1)$ . Toute observation qui ne se trouve pas entre ces deux limites est jugée aberrante et doit être exclue de la série avant toute analyse des données (on essaye de faire une interprétation de la présence des données aberrantes éventuelles en fin d'analyse).

**Exemple 1.6.1.3.2 :** Soit la série des données déjà ordonnée suivante :

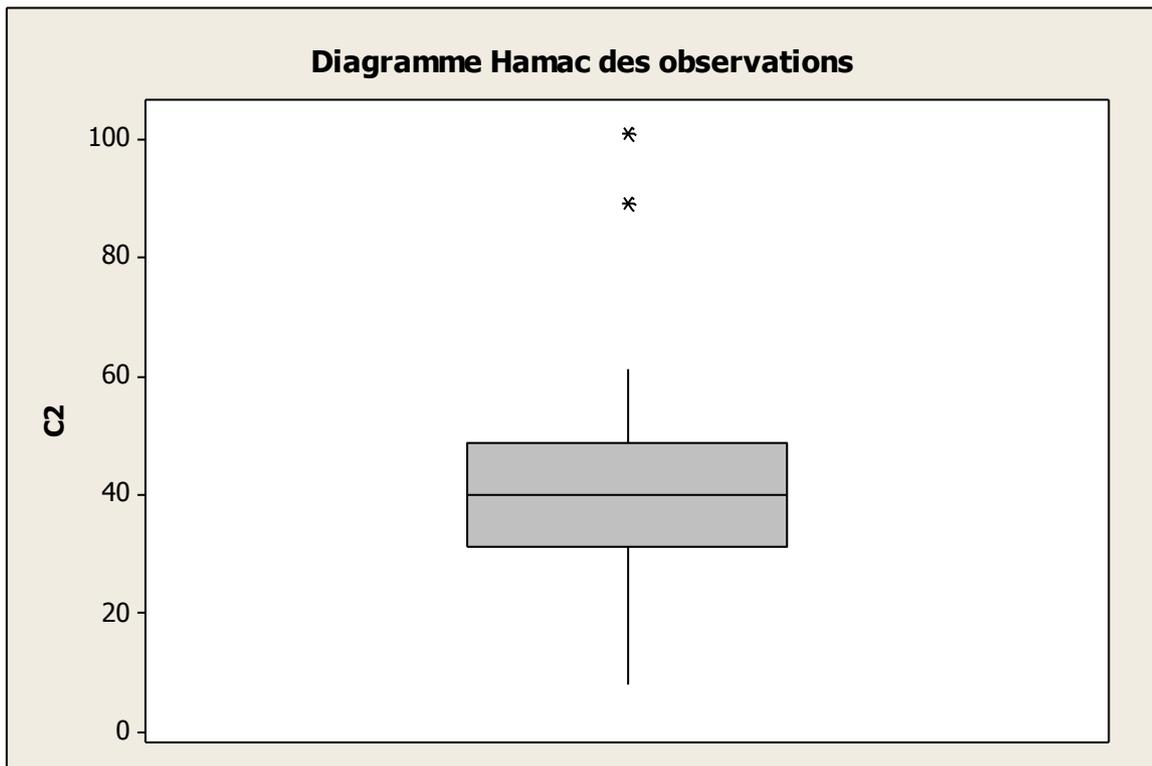
8 12 20 27 30 32 35 36 40 40 40 40 41 42 45 47 50 52 61 89 101.

(n=21 observations). Déterminer s'il y a des données aberrantes dans cette série à l'aide d'un graphique en boîte (box-plot).

**Réponse :** Les différentes mesures de cette variable sont obtenues à l'aide du logiciel Minitab:

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
C2	21	0	42.29	4.72	21.63	8.00	31.00	40.00	48.50	101.00

Ce qui signifie que  $Q_1 = 31$ ;  $Q_3 = 48,5$  et donc  $L_{inf} = Q_1 - 1,5 * (Q_3 - Q_1) = 4,75$  et  $L_{sup} = Q_3 + 1,5 * (Q_3 - Q_1) = 74,75$ . Donc, il y a 2 données aberrantes dans cette série ce sont 89 et 101 (qui sont signalées par \*), ce qui est illustré dans le diagramme en boîte ci-dessous.



**Remarque :** Une donnée aberrante peut avoir un effet catastrophique sur la moyenne, sur l'écart type et même sur l'allure générale de la distribution des données.

### 1.6.2. Les autres de position.

Quelques fois, on doit découper une série d'observations en cinq, en dix ou en cents groupes contenant chacun le même pourcentage d'observations. Dans le cas de cinq groupes, on parle alors des quintiles  $V_1, V_2, V_3$  et  $V_4$ . Entre deux quintiles consécutifs, il y a 20% d'observations. Dans le cas de dix groupes, on parle des déciles  $D_1, D_2, \dots, D_9$  et entre deux déciles consécutifs, il y a 10% d'observations. Dans le cas de cent groupes, on parle des centiles  $C_1, C_2, \dots, C_{99}$  et entre deux centiles consécutifs, il y a 1% des observations. Le calcul de ces différentes mesures de position est identique à ce qu'on a fait pour déterminer les quartiles, il n'y a que le pourcentage de la mesure à adapter à chaque fois. On va donner un exemple dans le cas où les données sont groupées par classes.

**Exemple 1.6.2.1 :** En reprenant les données de l'exemple 1.6.1.3.1, déterminer le deuxième quintile, le septième décile et le quatre vingt quinze centile de la variable

X=les recettes	Fréquences absolues	Fréquences relatives	Fréquences relatives cumulées
[10 ; 100[	5	0,125	0,125
[100 ; 190[	3	0,075	0,200
[190 ; 280[	11	0,275	0,475
[280 ; 370[	6	0,150	0,625
[370 ; 460[	11	0,275	0,900
[460 ; 550[	3	0,075	0,975
[550 ; 640]	1	0,025	1,000
Total	n=40	1,000	

X, les recettes quotidiennes d'un petit dépanneur et interprétez chacune de ces mesures.

Réponse :

- (a) Les fréquences cumulées dépassent pour la première fois 40% au niveau de la classe [190 ; 280[ ainsi le deuxième quintile est égal à

$V_2 = 190 + \frac{(0,40-0,20)}{0,275} * 90 = 255,45\$$ . Ceci signifie que dans cet échantillon de données, 40% des journées, les recettes quotidiennes de ce petit magasin ont été de 255,45 \$ ou moins.

- (b) Les fréquences relatives cumulées dépassent pour la première fois 70% au niveau de la classe [370 ; 460[, ainsi le septième décile est égal à

$D_7 = 370 + \frac{(0,70-0,625)}{0,275} * 90 = 394,55\$$ . Ce qui signifie que dans cet échantillon de données, 70% des journées, les recettes quotidiennes de ce petit magasin ont été de 394,55\$ ou moins.

- (c) Les fréquences relatives cumulées dépassent pour la première fois 95% au niveau de la classe [460 ; 550[, ainsi le quatre vingt quizième centile est égal à

$C_{95} = 460 + \frac{(0,95-0,90)}{0,075} * 90 = 520\$$ . Ce qui signifie que dans cet échantillon de données, 95% des journées, les recettes quotidiennes de ce petit magasin ont été de 520\$ ou moins.

## 1.7 : Les mesures de dispersion.

Rappelons qu'on travaille sur des données issues d'un échantillon et que le choix de cet échantillon est fait au hasard mais sensé refléter ce qui se passe dans la population. Ce qui fait que le comportement d'une variable diffère d'un échantillon à l'autre mais on espère qu'il correspond au profil de cette variable dans la population. Ce qui fait que lorsqu'on manipule une variable mesurable et qu'on se base seulement sur ses mesures de tendance centrale, on perd de vue la variabilité des données autour de ces mesures centrales. D'où l'utilité des mesures de dispersion qui, jumulées avec les mesures de tendance centrale, vont nous donner une idée plus exacte sur l'ensemble de ce qu'on a observé dans une série échantillonnale. Dans ce paragraphe, on va décrire quelques unes de ces mesures de dispersion.

### 1.7.1 : L'étendue.

C'est la mesure de dispersion la plus simple à calculer. Lorsqu'on a une variable quantitative  $X$ , mesurée sur un échantillon de taille  $n$ . Alors l'étendue est égale à  $E = \text{la plus grande donnée} - \text{la plus petite donnée} = X_{max} - X_{min}$ .

Puisque l'étendue est basée seulement sur les deux observations extrêmes, alors elle est très peu utilisée dans les applications.

### 1.7.2 : La variance.

La variance d'une variable mesurée sur un échantillon est égale à la moyenne des carrés des écarts qui séparent chaque observation de la moyenne échantillonnale, son calcul diffère selon la nature des données.

#### 1.7.2.1 : Les données en vrac.

Soit  $X$  une variable quantitative mesurée sur un échantillon de taille  $n$ , et dont les valeurs sont :  $x_1, x_2, \dots, x_n$  alors la variance de l'échantillon est

$$s_X^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

La sommation ci-dessus est divisée par  $(n-1)$  pour que cette variance échantillonnale soit une bonne estimation de la variance de toute la population. Ce qu'on verra plus en détails dans le chapitre VI. La variance se prête mal à l'interprétation car vue son calcul, son unité est égale au carré de l'unité de la variable  $X$ . Si par exemple  $X$  est égal au nombre d'enfants par ménage alors l'unité de la variance serait *(nombre d'enfants)<sup>2</sup>* qui n'a aucune signification.

La variance est surtout utile lorsqu'on a une variable mesurée dans plusieurs groupes (analyse de la variance) ou dans le cas où on veut comparer plusieurs variables mesurées sur le même échantillon ou comme étape de calcul pour calculer d'autres mesures.

**Exemple 1.7.2.1.1 :** Soit  $X$  une variable quantitative mesurée sur un échantillon de taille  $n=6$  et les valeurs suivantes ont été obtenues : -2 5 10 7 8 8

Alors  $\bar{x} = 6$  et la variance de cet échantillon sera égale à

$$s_X^2 = \frac{(-2 - 6)^2 + (5 - 6)^2 + \dots + (8 - 6)^2}{6 - 1} = 18.$$

#### 1.7.2.2 Les données groupées par valeurs.

Soit  $X$  une variable quantitative mesurée sur un échantillon de taille  $n$ , et dont les  $k$  valeurs sont :  $x_1, x_2, \dots, x_k$  avec des fréquences absolues respectivement égales à  $f_1, f_2, \dots, f_k$ . Alors la variance de  $X$  dans cet échantillon est égale à

$$s_X^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 f_i}{n-1}.$$

**Exemple 1.7.2.2.1 :** En reprenant le tableau de l'exemple 1.5.2.2.1, déterminer la variance de la variable  $X$ =le nombre d'accidents par semaine.

Réponse :  
trouvé  
moyenne  
variable

Tableau des fréquences du nombre d'accidents par semaine	
X	Fréquences absolues
0	4
1	2
2	10
3	7
4	10
5	4
6	3
Total	n=40

On avait  
trouvé  
la  
moyenne  
de  
cette  
variable

$\bar{x} = 3,025$  donc sa variance sera égale à :

$$s_{\bar{X}}^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 f_i}{n - 1} = \frac{(0 - 3,025)^2 * 4 + \dots + (6 - 3,025)^2 * 3}{39}$$

$$= 2,74 \left( \frac{\text{accidents}}{\text{semaine}} \right)^2 \text{!!!!}$$

### 1.7.2.3 : Les données groupées par classes.

Soit maintenant X, une variable quantitative mesurée sur un échantillon de taille n, et dont les observations sont groupées en k classes avec des fréquences absolues respectivement égales à  $f_1, f_2, \dots, f_k$  et dont les milieux des classes sont respectivement égaux à  $m_1, m_2, \dots, m_k$ . Alors la variance échantillonnale de cette variable est :

$$s_{\bar{X}}^2 = \frac{\sum_{i=1}^k (m_i - \bar{x})^2 f_i}{n - 1}$$

**Exemple 1.7.2.3.1 :** En reprenant les données de l'exemple 1.5.2.3.1, déterminer la variance de la variable X, les recettes quotidiennes d'un petit dépanneur.

Réponse : On avait trouvé que la moyenne de la variable est  $\bar{x} = 298\$$

$m_i$	X=les recettes	Fréquences absolues
55	[10 ; 100[	5
145	[100 ; 190[	3
235	[190 ; 280[	11
325	[280 ; 370[	6
415	[370 ; 460[	11
505	[460 ; 550[	3
595	[550 ; 640]	1
	Total	n=40

Alors la variance de cet échantillon est égale à :

$$s_{\bar{X}}^2 = \frac{\sum_{i=1}^k (m_i - \bar{x})^2 f_i}{n - 1} = \frac{(55 - 298)^2 * 5 + \dots + (595 - 298)^2 * 1}{39}$$

$$= 20021,54 (\$)^2 \text{!!!!}$$

### 1.7.3 : L'écart type.

L'écart type d'une variable quantitative mesurée sur un échantillon est égal à la racine carrée de sa variance. Son unité de mesure étant la même que celle de la variable, l'écart type se prête alors aisément à l'interprétation et est considéré comme la mesure de dispersion par excellence. La variance n'est donc qu'une étape de calcul pour déterminer l'écart type, quand on faisait les calculs à la main. Maintenant que tout est programmé, aucune calculatrice et aucun logiciel ne parle de variance comme telle.

**Exemple 1.7.3.1 :** L'écart type échantillonnal pour les 3 précédents exemples où on a calculé les variances échantillonnales est respectivement égal à :

$s_X = \sqrt{18} = 4,24$ . Pour les données de l'exemple 1.7.2.1.1 où les données sont en vrac.

$s_X = \sqrt{2,74} = 1,655$ . Pour les données de l'exemple 1.7.2.2.1 où les données sont en groupées par valeurs.

$s_X = \sqrt{20021,54} = 141,497$ . Pour les données de l'exemple 1.7.2.3.1 où les données sont groupées par classes.

#### Interprétation de l'écart type échantillonnal.

L'écart type mesure la dispersion entre toutes les valeurs observées. Des valeurs proches donneront un plus petit écart type, alors que des données très séparées donneront un plus grand écart type.

Lorsque la distribution des données (histogramme ou polygone des fréquences ou autre) a une forme en cloche et que la taille de l'échantillon est supérieure à 100, on doit s'attendre à avoir 68% des données observées comprises entre la moyenne plus ou moins un écart type et 95% des données observées soient comprises entre la moyenne plus ou moins deux écarts types. Si on se trouve dans les mêmes conditions on peut estimer l'écart type par la formule suivante :

$$s_X \approx \frac{\text{Étendue de } X}{4}$$

#### 1.7.3.1 : Propriétés de l'écart type échantillonnal.

Soit  $X$  une variable quantitative dont l'écart type échantillonnal est  $s_X$  et soit  $Y$  une autre variable quantitative telle que  $Y = a + b * X$  où  $a$  et  $b$  sont des constantes réelles. Alors l'écart type échantillonnal de  $Y$  sera égal à

$$s_Y = |b| * s_X$$

**Exemple 1.7.3.1.1 :** Reprenons le contexte de l'exemple 1.5.2.4.1, où  $X$  est le nombre d'heures qu'un étudiant travaille à temps partiel par semaine. Supposons qu'à partir d'un échantillon d'étudiants, on ait pu trouvé que l'écart type du nombre d'heures travaillées par ces étudiants est égal à  $s_X = 3,2$  heures/semaine. Si le salaire horaire est de 10\$ et que les patrons de ces étudiants leur offrent 30\$ par semaine pour leurs déplacements, quel est l'écart type du gain net hebdomadaire de ces étudiants ? Posons  $Y$ , le gain net hebdomadaire de ces étudiants alors  $Y = 30 + 10 * X$ , donc l'écart type du gain net de cet échantillon d'étudiants sera égal à  $s_Y = 10 * s_X = 32$  \$/semaine.

#### 1.7.4 : Le coefficient de variation.

On avait dit que l'unité de l'écart type d'une variable est la même que celles des données et qu'alors il s'interprète mieux que la variance. Mais si on veut comparer la

dispersion de deux variables ou plus ayant des unités différentes mesurées sur le même échantillon ou sur des échantillons différents, il nous faut une mesure de dispersion sans unité. Cette mesure est le coefficient de variation. Pour un échantillon de données dont la moyenne est non négative, on définit le coefficient de variation d'une variable X par :

$$CV_X = \frac{s_X}{\bar{x}} 100\%.$$

Si on a un seul échantillon de données, alors si le coefficient de variation de X est inférieur à 15%, on dit que la variable est homogène, sinon elle est dite hétérogène.

Si on a deux échantillons (sur une ou deux variables) ou plus, alors celui (ou celle) qui a le plus petit coefficient de variation est le (ou la) plus homogène.

**Exemple 1.7.4.1 :** On a pris un échantillon de taille  $n=50$  d'hommes d'âge adultes, on a mesuré leur poids et leur taille. Les résultats sont résumés dans le tableau suivant :

Variable	Moyenne	Écart type
X=taille	$\bar{x}=173,59$ cm	$s_X = 7,86$ cm
Y=poids	$\bar{y} = 78,42$ kg	$s_Y=11,98$ kg

Pour comparer l'homogénéité de ces deux variables, on utilise leur coefficient de variation.

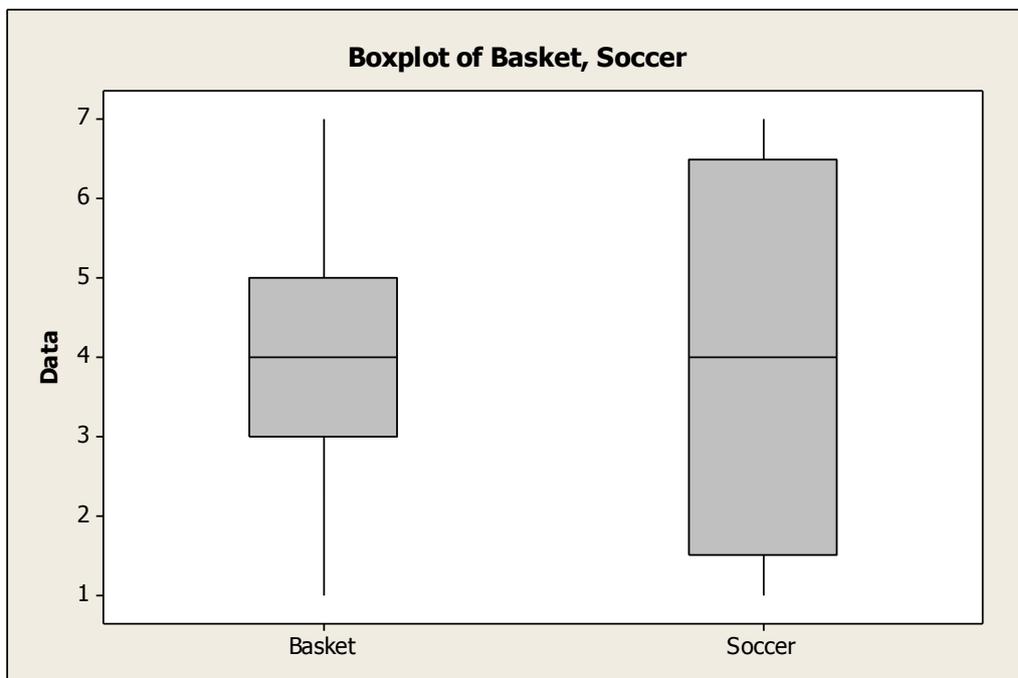
$$CV_X = \frac{7,86}{173,59} 100\% = 4,53\%$$

$$CV_Y = \frac{11,98}{78,42} 100\% = 15,28\%$$

Donc la taille des hommes adultes est plus homogène que leur poids. Ce qui correspond à l'intuition. Par exemple il est très rare de voir deux hommes adultes dont l'un serait deux fois plus grand que l'autre, alors qu'il est fréquent de voir un homme adulte dont le poids est le double d'un autre.

On voit que leur moyenne et leur médiane sont toutes égales à 4, donc si on se limitait aux mesures de tendances centrales, on aurait conclu à une similitude de ces deux distributions.

Mais en comparant leur écart type et donc leur coefficient de variation, on voit que les données sur le soccer sont plus dispersées. Ce qu'on peut aussi confirmer par des graphiques suivants :



### Ayez un esprit critique.

Maintenant qu'on est armé d'outils pour examiner la tendance centrale, la dispersion, la distribution des données, les valeurs extrêmes ou aberrantes, on pourrait être tenté de développer une procédure mécanique et aveugle, mais penser de façon critique est d'une importance primordiale dans toute analyse de données. En plus de l'utilisation des outils présentés dans ce chapitre, il est important de ne pas négliger tout autre facteur qui s'y rapporte et qui pourrait être crucial pour les conclusions de l'étude. On pourrait penser par exemple à la représentativité des données, à la source des données qui pourrait affecter leur qualité. En résumé, en plus des outils présentés dans ce chapitre, on devrait aussi penser.

### Outil technologique :

- I. **La calculatrice TI-83 :** On va présenter à partir des données de l'exemple 1.4.2.2.1 comment tracer un histogramme avec une calculatrice TI-83 ou plus performante.
- II.

<b>HISTOGRAMME</b>	
<b>Les touches de la calculatrice</b>	<b>Ce qu'on fait et ce qu'on obtient</b>
1. STAT	On accède à l'éditeur
1	
2. 16 Enter	On entre les données
....	
....	
621.35 Enter	
3. WINDOW	On accède aux fenêtres pour mettre les limites des classes et leur amplitude.
4. 10 Enter	La borne minimale
640 Enter	La borne maximale
90 Enter	L'amplitude des classes
Enter	
5. 2ND	Tracer l'historgramme

Y= Enter	
Curseur sur ON	Avec ces flèches on va chercher parmi les graphiques celui de l'histogramme
	
	
	
ENTER	
GRAPH	
6. TRACE	On détermine les fréquences et les limites en déplaçant le curseur.

Pour calculer toutes les mesures avec une calculatrice TI-83 ou plus performante, il suffit d'accéder à l'éditeur de STAT et saisir les données dans une liste et le calcul se fait automatiquement en appuyant sur la touche calcul (une seule variable) et donner le nom de la liste qui contient les données.

# Exercices

**Exercice 1 :** Soit  $x$  une série statistique. Démontrer la formule de Koenig pour la variance :

$$s_x^2 = \overline{x^2} - \bar{x}^2.$$

**Exercice 2 :** Soit une série statistique de taille  $n$ , classée suivant la partition

$[d_1, d_2[, \dots, [d_k, d_{k+1}[, \dots, [d_{m-1}, d_m[$ . On note  $n_k, N_k, a_k$  respectivement l'effectif, l'effectif cumulé et l'amplitude de la classe  $[d_k, d_{k+1}[$ . Soit  $[d_j, d_{j+1}[$  la première classe contenant au moins 50% des effectifs cumulés. Démontrer que l'on peut approcher la médiane par interpolation linéaire :

$Me \approx d_j + \frac{n/2 - N_{j-1}}{n_j} \cdot a_j$ . De façon analogue, trouver des formules approchées pour les premier et troisièmes quartiles.

**Exercice 3 :** Au poste de péage, on compte le nombre de voitures se présentant sur une période de 5mn. Sur 100 observations de 5mn, on obtient les résultats suivants :

Nombre de voitures	1	2	3	4	5	6	7	8	9	10	11	12
Nombre d'observations	2	8	14	20	19	15	9	6	2	3	1	1

- 1) Construire la table des fréquences et le diagramme en bâtons en fréquences de la série du nombre de voitures.
- 2) Calculer la moyenne et l'écart-type de cette série.
- 3) Déterminer la médiane, les quartiles et tracer le box-plot.
- 4) Etudier la symétrie de la série.

**Exercice 4 :** On donne la série unidimensionnelle suivante, correspondant à la répartition des entreprises du secteur automobile en fonction de leur chiffre d'affaire en millions d'euros.

chiffres d'affaires	moins de 0,25	[0,25 ; 0,5[	[0,5 ; 1[	[1 ; 2,5[	[2,5 ; 5[	[5 ; 10[
nombre d'entreprises	137	106	112	154	100	33

- a) Calculer le chiffre d'affaire moyen et l'écart-type de la série.
- b) Construire l'histogramme des fréquences
- c) Construire les deux polygones des fréquences cumulées
- d) Calculer la médiane et la proportion d'entreprises dont le chiffre d'affaire est supérieur à 3 millions d'euros.

**Exercice 5 :** La distribution des demandeurs d'emploi selon le sexe et la classe d'âge dans une localité est la suivante :

âge	Hommes	Femmes
[16 ; 26[	280	160
[26 ; 40[	310	360
[40 ; 50[	240	120
[50 ; 60[	420	530
[60 ; 65[	70	50

- a) Tracer les deux courbes de fréquences cumulées croissantes.
- b) Déterminer les quartiles de la variable X associant à chaque demandeur d'emploi masculin son âge. Même question pour les demandeurs d'emploi de sexe féminin.
- c) Conclusions.

**B- Statistiques descriptives bidimensionnelles**

**Exercice 6 :** On cherche à étudier la relation entre le nombre d'enfants d'un couple et son salaire. On dispose de la série bidimensionnelle suivantes :

Salaire en euros (Y)	Nombre d'enfants (X)
510	4
590	3
900	2
1420	1
2000	0
600	5
850	6
1300	7
2200	8

- a) Calculer le coefficient de corrélation linéaire entre ces deux variables statistiques. Conclusion ?
- b) Un expert en démographie affirme que les deux caractéristiques sont indépendantes. Qu'en pensez-vous ?

**Exercice 7 :** L'indice moyen d'un salaire a évolué de la façon suivante :

année	1	2	3	4	5	6	7
indice	165	176	193	202	222	245	253

- a) Représenter cette série statistique par un nuage de points.
- b) En utilisant la méthode des moindres carrés, calculer l'équation de la droite représentant l'indice en fonction de l'année.
- c) Comment pourrait-on prévoir l'indice à l'année 9 ?

**Exercice 8 :** Soit X une variable statistique qualitative à k modalités et Y une variable statistique quantitative. Chaque modalité de X définit une sous-population : celle des individus ayant cette modalité. On note  $n_j$  l'effectif correspondant à la modalité j de X,  $\bar{y}_j$  (resp.  $s^2_j(y)$ ) la moyenne (resp. la variance) des valeurs de la variable Y pour les individus de la modalité j. Montrer que  $s^2_Y = s^2_E + s^2_R$  où  $s^2_E = \frac{1}{n} \sum_{j=1}^k n_j (\bar{y}_j - \bar{y})^2$  et  $s^2_R = \frac{1}{n} \sum_{j=1}^k n_j s^2_j(y)$ . On les appelle respectivement variances inter et intra-catégories.

**Exercice 9 :** On observe le nombre d'enfants Y sur un ensemble de 12 individus répartis entre les sexes (variable X) :

F	3	4	5	4	2	5
H	10	7	6	3	4	2

- 1) Représenter graphiquement cette série.
- 2) Calculer les moyennes arithmétiques dans chaque classe
- 3) Calculer les variances inter et intra-catégories.
- 4) Calculer et interpréter le rapport de corrélation entre X et Y. Conclusion ?

**Exercice 10 :** Soient x et y deux séries statistiques de taille n. On note  $r_x$  et  $r_y$  les séries des rangs correspondantes.

a) Montrer que  $\overline{r_x} = \frac{n+1}{2}$ .

b) Montrer que  $s_{r_x}^2 = \frac{n^2-1}{12}$ .

c) En posant  $d_i = r_{x_i} - r_{y_i}$ , montrer que  $2s(r_x, r_y) = s_{r_x}^2 + s_{r_y}^2 - \frac{1}{n} \sum_{i=1}^n d_i^2$ .

d) En déduire l'expression du coefficient linéaire entre ces deux séries, appelé coefficient

$$\text{de corrélation des rangs de Spearman : } r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)}.$$

**Exercice 11 :** Dix échantillons de cidre ont été classés par ordre de préférence par deux gastronomes. On obtient les classements suivants :

A	1	2	3	4	5	6	7	8	9	10
B	3	1	4	2	6	5	9	8	10	7

- 1) Calculer le coefficient de corrélation des rangs de Spearman. Conclusion ?
- 2) Une autre façon d'évaluer le lien entre les rangs de deux séries consiste à utiliser le coefficient de corrélation des rangs de Kendall. Ce coefficient est défini par :  $\tau = \frac{2S}{n(n-1)}$ , où S est obtenue de la façon suivante : on considère tous les couples d'individus de la série. On note 1 si les individus i et j sont dans le même ordre pour les deux variables considérées (ici  $a_i < a_j$  et  $b_i < b_j$ ). On note -1 si les deux classements discordent (ici  $a_i < a_j$  et  $b_i > b_j$ ). S est la somme des valeurs obtenues pour les  $\frac{n(n-1)}{2}$  couples distincts. Montrer que  $\tau$  est compris entre -1 et 1 et qu'il est d'autant plus proche de 1 que les classements sont semblables. Calculer  $\tau$  pour les données dont on dispose.

**Exercice 12 :** On considère un échantillon de 797 étudiants d'une université ayant obtenu le DEUG. On étudie le lien entre l'âge d'obtention du Bac (variable Y), à 4 modalités (moins de 18 ans, 18 ans, 19 ans, plus de 19 ans), et la durée d'obtention du DEUG (variable X), à 3 modalités (2 ans, 3 ans, 4 ans). On a la table de contingence ci-dessous :

X	Y	Moins de 18 ans	18 ans	19 ans	Plus de 19 ans
2 ans		84	224	73	19
3 ans		35	137	75	27
4 ans		14	59	34	16

- 1) Déterminer le tableau des profils colonnes en pourcentage
- 2) Représenter graphiquement le diagramme en barre de ces profils
- 3) Déterminer le tableau des effectifs théoriques
- 4) Calculer l'indice du Chi2 et les contributions de chaque case. Conclusion ?

# Chapitre 2

## PROBABILITES - VARIABLES ALEATOIRES

Dans le cours sur les probabilités nous avons introduit la notion d'univers  $U$  et lui avons attaché une fonction–probabilité  $P$ . Dans beaucoup d'applications pratiques la notion d'univers, pour des raisons de commodité, cède sa place à celle de variable aléatoire. La fonction–probabilité attachée à la variable aléatoire est alors une application des parties de  $\rho$  dans  $[0,1]$  (plutôt que des parties de  $U$  dans  $[0,1]$ ). Cette manière de faire courante n'enlève rien à l'importance de la notion d'univers qui reste malgré tout, implicitement, à la base du calcul des probabilités (voir la référence [2], [8], [10]).

Considérons un univers des possibles  $U$  correspondant à une certaine expérience aléatoire. Les éléments de  $U$  (appelés issues : résultats possibles de l'expérience) ne sont généralement pas des nombres. Il est cependant utile de faire correspondre un nombre à chaque élément de  $U$ , en vue de faire ensuite des calculs.

Pour un jet de dé, il semble naturel de faire correspondre à la face obtenue par le jet, le nombre de points qu'elle porte, mais ce n'est pas une obligation. Si on jette 2 dés, on s'intéressera par exemple à la somme des points obtenus. Dans les situations de jeux ou de paris, comme dans la vie réelle, les éventualités se voient affecter d'un nombre : aux diverses cartes que l'on peut tirer d'un jeu bien battu, on associe leur valeur convenue : 0 pour une carte de 2 à 10, un point pour un valet, deux pour une dame, trois pour un roi et quatre pour un as. On définit ainsi une fonction  $X$ . Cette fonction  $X$  permet à son tour de définir des événements : ainsi pour le tirage d'une carte :

$X = 1$  représente le tirage d'un valet ,  $X \leq 3$  représente le tirage de toute carte autre qu'un as ,  
 $X = 0$  représente le tirage d'une carte de 2 à 10.

Il est très important de bien prendre conscience du sens de ces notations, relativement inhabituelles : il n'est pas fréquent qu'on écrive le signe " = " entre une fonction et un nombre ! Et pour achever de nous désorienter, l'usage est de donner à une telle fonction le nom de **variable aléatoire** : quelques auteurs ont bien tenté de réagir, en utilisant le nom d'*aléa numérique*, tellement ils souffraient de voir une fonction baptisée variable ! Mais l'usage est le plus fort.

Une variable aléatoire  $X$ , sur un univers des possibles  $U$ , est une application de  $U$  dans  $\rho$  : à tout résultat possible de l'expérience (à tout élément de  $U$ ), la variable aléatoire  $X$  fait correspondre un nombre.

Lorsque  $U$  est fini ( $\text{Card } U = n$  et  $n \in \mathbf{v}^*$ ) ou infini dénombrable ( $\text{Card } U = \aleph_0$ ), toute application de  $U$  dans  $\rho$  est une variable aléatoire. Lorsque  $U$  est non dénombrable ( $\text{Card } U = \aleph_1$ ), il existe certaines applications de  $U$  dans  $\rho$  qui ne sont pas des variables aléatoires. En effet, la définition rigoureuse d'une variable aléatoire  $X$  impose que tout intervalle de  $\rho$  soit l'image d'un événement  $A$  de  $U$  ( $A \subset U$  ou  $A \in \mathcal{P}(U)$ ) par l'application  $X$ .

Cette condition est vérifiée pour toute application  $X$  si  $U$  est fini ou dénombrable, puisque toute partie de  $U$  est un événement. Ce n'est plus vrai si  $U$  est non dénombrable.

Les applications choisies ultérieurement seront des variables aléatoires.

On parle de variable aléatoire **discrète** lorsque la variable est une application de  $U$  dans un sous-ensemble  $X(U)$  discret de  $\rho$ , le plus souvent  $\mathbf{v}$  ou une partie de  $\mathbf{v}$ . On parle sinon de variable aléatoire **continue**.

Pour un nombre réel  $a$  donné, l'événement constitué de toutes les issues  $i$  ( $i \in U$ ) de l'expérience telles que  $X(i) = a$  est noté  $[X(i) = a]$ , ou, en abrégé,  $X = a$ .

Pour deux nombres réels  $a$  et  $b$  ( $a \leq b$ ), l'événement constitué de toutes les issues  $i$  de l'expérience telles que  $a < X(i) \leq b$  est noté  $[a < X(i) \leq b]$  ou, en abrégé,  $a < X \leq b$ .

Considérons maintenant le cas le plus simple d'une variable aléatoire finie, que nous généraliserons dans un second temps à une variable aléatoire infinie dénombrable, puis continue.

Soit  $X$  une variable aléatoire sur un univers des possibles  $U$  à valeurs finies :  $X(U) = \{x_1, x_2, \dots, x_n\}$ .

$X(U)$  devient un ensemble probabilisé si l'on définit la probabilité  $P(X = x_i)$  pour chaque  $x_i$ , que l'on note  $f(x_i)$ .

Cette fonction  $f$ , définie par  $f(x_i) = P(X = x_i)$  est appelée **distribution (ou loi) de probabilité** de  $X$ .

Puisque les  $f(x_i)$  sont des probabilités sur les événements  $\{X=x_1, X=x_2, \dots, X=x_n\}$ , on a par conséquent :

$$1) \quad \forall i, f(x_i) \geq 0 \quad 2) \quad \sum_{i=1}^n f(x_i) = 1$$

$$3) \quad P(a < X \leq b) = \sum_i f(x_i), \text{ la somme s'étendant à tous les indices } i \text{ tels que } a < x_i \leq b.$$

## Définition - Exemples

### 1.1 Définition

Soit  $U$  l'univers associé à une expérience aléatoire

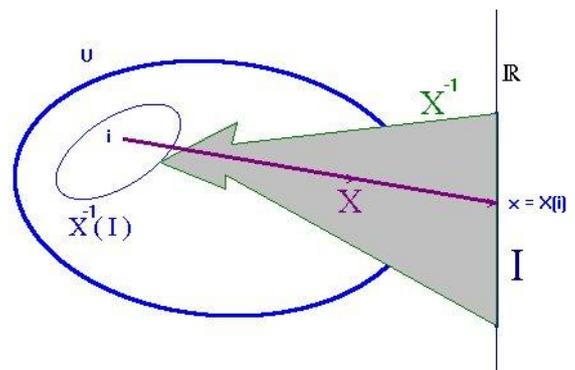
et  $\mathcal{P}(U)$  l'ensemble des événements.

On appelle variable aléatoire toute application  $X$

de  $U$  dans  $\mathcal{R}$  telle que  $X^{-1}(I) \in \mathcal{P}(U)$ , pour tout

intervalle  $I$  de  $\mathcal{R}$ .  $X(U)$  est alors l'image de  $U$ .

On a  $X^{-1}(I) = \{i \in U \mid \exists x \in I \text{ tel que } X(i) = x\} \subset U$ .



### 1.2 Exemples :

a) On jette un dé cubique et on s'intéresse au jeu suivant : si on obtient un numéro  $\leq 4$  on perd 1 F, sinon on gagne 2 F. L'application  $X$  qui à tout tirage associe le gain obtenu (une perte est un gain négatif) est une variable aléatoire (discrète prenant un nombre fini de valeurs).

$$\text{On a } U = \{1,2,3,4,5,6\}, X(U) = \{-1,2\} \text{ et si } i \in U, \text{ on pose } X(i) = \begin{cases} -1 & \text{si } i \leq 4 \\ 2 & \text{sinon} \end{cases}.$$

b) On jette une pièce de monnaie jusqu'à ce que le côté face apparaisse et on s'intéresse au jeu suivant : si on obtient face au  $i^{\text{ième}}$  lancer on gagne  $2^i$  F. L'application  $X$  qui à tout tirage associe le gain obtenu est une variable aléatoire (discrète prenant un nombre infini dénombrable de valeurs).

$$\text{On a } U = \mathbf{v}^*, X(U) = \{2,4,8,16,\dots\} \text{ et } X(i) = 2^i = k.$$

c) On s'intéresse à la durée de bon fonctionnement d'un équipement particulier fabriqué en grande série dans l'intervalle  $[0, +\infty[$ . L'application  $X$  qui à un appareil associe sa durée de bon fonctionnement est une variable aléatoire (continue, par opposition à discrète, prenant un nombre infini non dénombrable de valeurs).

$$\text{On a } U = \mathbf{v}^*, X(U) = [0, +\infty[ \text{ et } X(i) = t \text{ où } t \text{ est le temps de bon fonctionnement de } i \in U.$$

### 1.3 Définitions

Si  $X$  est une variable aléatoire définie sur un univers des possibles  $U$  et  $k$  un réel, on définit les variables aléatoires  $X+k$ ,  $kX$  ainsi :

- $\forall i \in U, (X+k)(i) = X(i) + k, (kX)(i) = kX(i)$
- si  $X(U) = \{x_1, x_2, \dots, x_n\}$ , alors  $(X+k)(U) = \{x_1+k, x_2+k, \dots, x_n+k\}$  et  $(kX)(U) = \{kx_1, \dots, kx_n\}$
- $P(X+k = x_i+k) = P(X=x_i)$  et  $P(kX = kx_i) = P(X=x_i)$

## Loi de probabilité

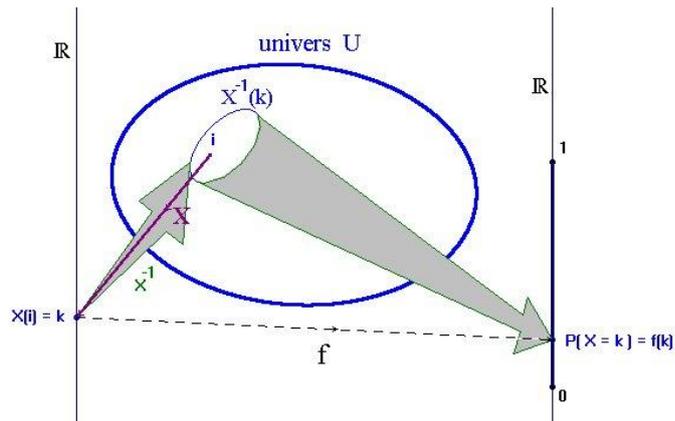
### 2.1 Définition :

La loi (ou distribution) de probabilité  $f$  de la variable aléatoire  $X$  est la fonction

$$f : X(U) \rightarrow [0,1]$$

$$k \mapsto f(k) = P(X = k)$$

Précédemment nous travaillions dans un espace probabilisé  $(U, \mathcal{P}(U), P)$ , où  $U$  est l'univers des possibles,  $\mathcal{P}(U)$  est l'ensemble des événements et  $P$  une probabilité sur  $\mathcal{P}(U)$ .



Soit  $X^{-1}(\{k\}) = \{i \in U \mid X(i) = k\} \subset U$  ; on note cet événement  $X^{-1}(\{k\}) : X = k$

Maintenant nous allons travailler dans un autre espace probabilisé  $(X(U), \mathcal{T}, f)$  où  $X(U)$  est l'ensemble des issues possibles (l'ensemble des nombres attribués à toutes les issues  $i$  possibles de  $U$ ),  $\mathcal{T} \subset \mathcal{P}(\rho)$  l'ensemble des événements (noté  $X = k, X \leq n$  ou  $a \leq X \leq b$ ), et  $f$  une loi de probabilité (densité de probabilité dans le cas d'une variable aléatoire continue) vérifiant les trois propriétés suivantes :

- 1)  $\forall i, f(x_i) \geq 0$  (correspondant à  $P(A) \geq 0, \forall A \subset U$ )
- 2)  $\sum_{i=1}^n f(x_i) = 1$  (correspondant à  $P(U) = 1$ )
- 3)  $P(a < X \leq b) = \sum_i f(x_i)$ , la somme s'étendant à tous les indices  $i$  tels que  $a < x_i \leq b$ .  
(correspondant à  $P(A \cup B) = P(A) + P(B) \Leftrightarrow A \cap B = \emptyset$ )

## 2.2 Exemples : ( reprenons les exemples du point 1.2 )

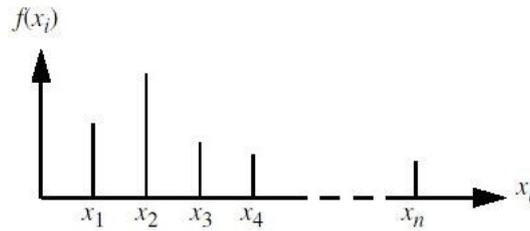
a) Le tableau suivant nous donne la loi de probabilité  $f$  de  $X$  :  
exemple de calcul :

$k$	-1	2
$f(k)=$	$\frac{2}{3}$	$\frac{1}{3}$
$P(X = k)$	$\frac{2}{3}$	$\frac{1}{3}$

$$f(2) = P(X = 2) = P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}.$$

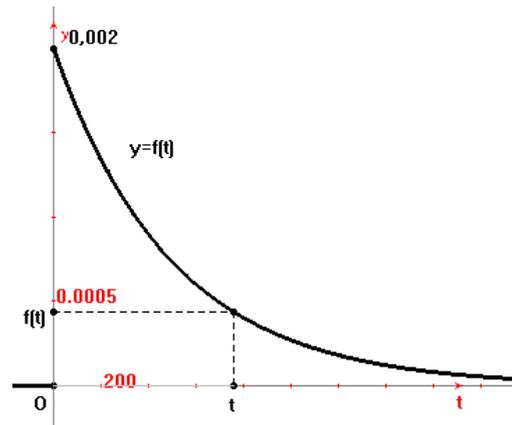
b) Ici,  $P(X = k) = \frac{1}{k} = \left(\frac{1}{2}\right)^i$ , où  $k = 2^i$ .

Remarque : on peut aussi représenter  $f$  par un diagramme en bâtons :



c) Lorsque la variable  $X$  est continue, on lui associe une densité de probabilité  $f$  ; dans cet exemple  $f$  est définie par :

$$P(X = t) = f(t) = \begin{cases} 0 & \text{si } t < 0 \\ 0,002e^{-0,002t} & \text{si } t \geq 0 \end{cases}$$



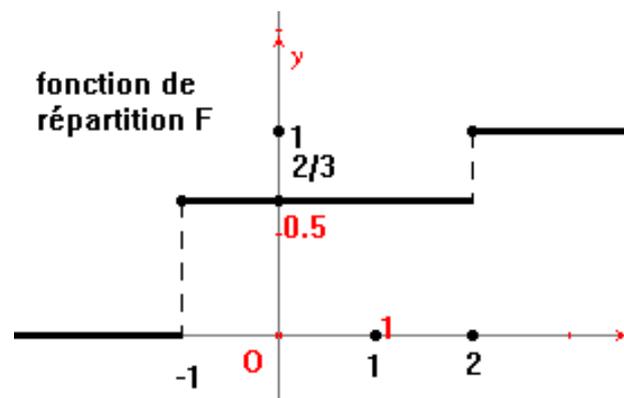
## 2.3 Fonction de répartition

### Définition :

La fonction  $F$  définie par  $F(x) = P(X \leq x)$  de  $\rho$  dans  $[0,1]$  est appelée fonction de répartition de la variable aléatoire  $X$ .

Exemples : ( cf. les exemples du point 1.2 – 2.2 )

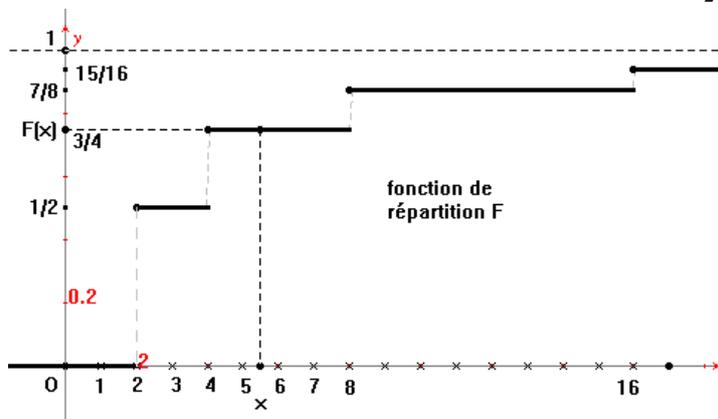
a) Ici,  $F(x) = \begin{cases} 0 & \text{si } x < -1 \\ \frac{2}{3} & \text{si } -1 \leq x < 2. \\ 1 & \text{si } x \geq 2 \end{cases}$



b) Ici,  $F(x) = \begin{cases} 0 & \text{si } x < 2 \\ 1 - \left(\frac{1}{2}\right)^n & \text{si } x \geq 2 \end{cases}$  où  $n$  est la partie entière de  $y$  si  $x = 2^y$ .

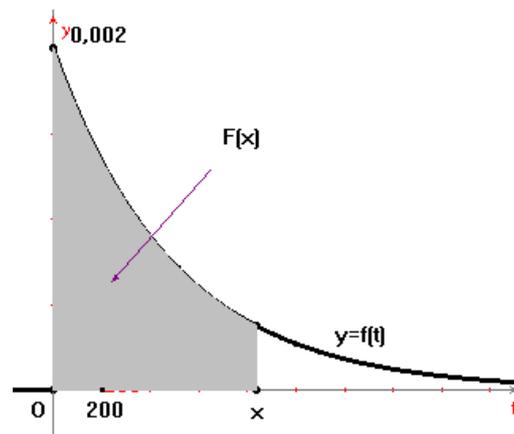
On a en effet,  $F(x) = F(n) = P(X \leq n) = \sum_{k=1}^n P(x=k) = \sum_{k=1}^n \left(\frac{1}{2}\right)^k = \frac{1}{2} \frac{1 - \left(\frac{1}{2}\right)^n}{1 - \frac{1}{2}} = 1 - \left(\frac{1}{2}\right)^n$

(somme des  $n$  premiers termes de la suite géométrique de premier terme  $\frac{1}{2}$  et de raison  $\frac{1}{2}$ ).



c) Ici, on a la densité de probabilité de  $X$ , définie par la fonction  $f$ :

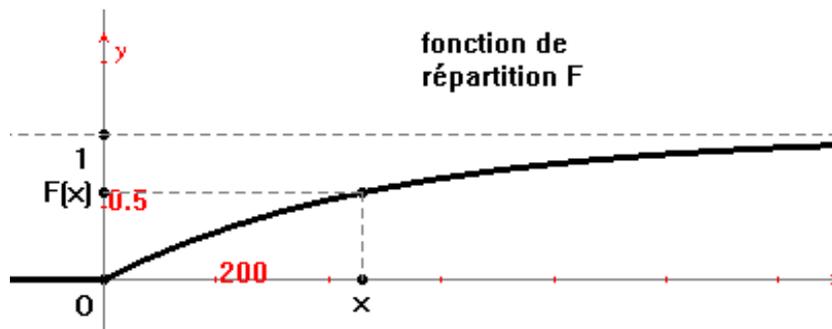
$$f(t) = \begin{cases} 0 & \text{si } t < 0 \\ 0,002e^{-0,002t} & \text{si } t \geq 0 \end{cases}$$



avec  $F(x) = P(X \leq x) = \int_0^x f(t) dt = \int_0^x 0,002e^{-0,002t} dt = [-e^{-0,002t}]_0^x = 1 - e^{-0,002x}$

On a alors  $F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-0,002x} & \text{si } x \geq 0 \end{cases}$

Dans tous les cas,  $F$  est une fonction monotone croissante ( $a \leq b \Rightarrow F(a) \leq F(b)$ ) et de plus  $\lim_{x \rightarrow -\infty} F(x) = 0$  et  $\lim_{x \rightarrow +\infty} F(x) = 1$



Plus généralement on a

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(t) dt .$$

Preuve :

Si  $a < b$ , on a :

$$(X \leq b) = (X \leq a) \cup (a < X \leq b)$$

les deux événements  $(X \leq a)$  et  $(a < X \leq b)$  étant incompatibles

$$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

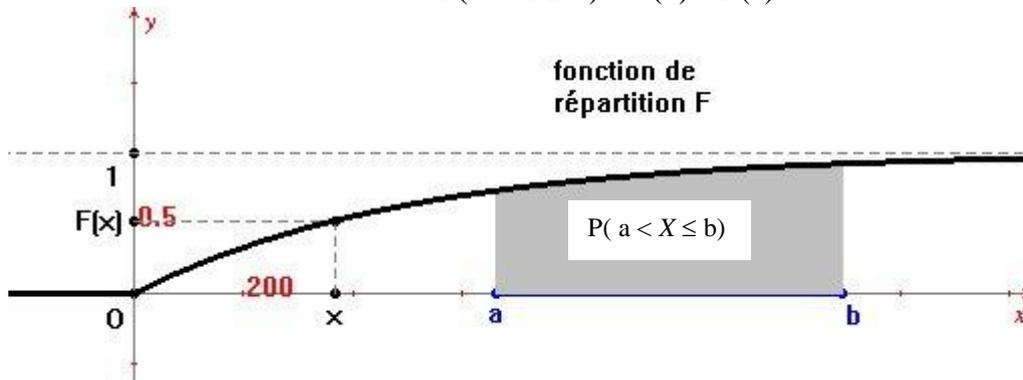
on en déduit :

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

Soit

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

Donc



## 2.4 Espérance mathématique d'une loi finie

Le terme d'**espérance mathématique** a été introduit par Pascal à propos des jeux de hasard. L'espérance mathématique peut s'interpréter comme le gain moyen du joueur ou comme le montant de la mise à engager à chaque partie si le jeu est sans frais. C'est pourquoi on dit qu'un jeu est **équitable** quand  $E(X) = 0$ , où  $X$  désigne le gain algébrique du joueur (gain – perte)

L'**espérance mathématique** cherche à traduire la tendance centrale de la loi.

Ce nom plein de poésie est celui sous lequel on désigne la **moyenne** pondérée des valeurs d'une loi de probabilité, les coefficients de pondération étant les probabilités correspondantes.

Cette **moyenne** où chacune des valeurs  $x_i$  intervient d'autant plus que sa probabilité est importante s'apparente à un barycentre ou un centre de gravité en physique. On peut dire aussi que  $E(X)$  est l'abscisse du centre de gravité des points d'abscisses  $x_1, x_2, \dots, x_n$  affectés des masses  $f(x_1), f(x_2), \dots, f(x_n)$ .

### 2.4.1 Définition

L'espérance mathématique d'une loi finie  $f$  d'une variable aléatoire  $X$  est le réel  $E(X)$  défini par :

$$\mu_X = E(X) = \sum_{i=1}^n x_i \cdot f(x_i) = x_1 f(x_1) + x_2 f(x_2) + x_3 f(x_3) + \dots + x_n f(x_n) \text{ dans le cas où } X \text{ est discrète}$$

et prend un nombre fini  $n$  de valeurs  $x_i$ , ( $1 \leq i \leq n$ ) avec  $f(x_i) = P(X = x_i)$ .

remarque :  $\mu_X$  est notée  $\mu$  s'il n'y a pas de confusion.

## 2.4.2 Interprétation géométrique

$E(x)$  est égale à l'aire algébrique comprise entre le graphique de la fonction de répartition  $F$  associée à  $X$  et l'axe des ordonnées. En effet, pour chaque valeur  $x_i$  prise par  $X$ , le graphique de  $F$  présente un saut de longueur  $f(x_i)$ . Et on peut dire que  $x_i f(x_i)$  est l'aire du rectangle de base  $[0, x_i]$  et de hauteur  $f(x_i)$  lorsque  $x_i$  est positif, et l'opposé de cette aire lorsque  $x_i$  est négatif.

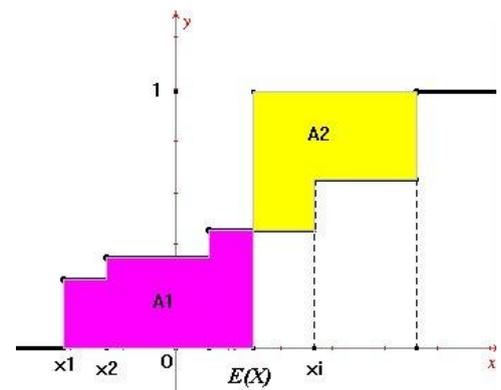
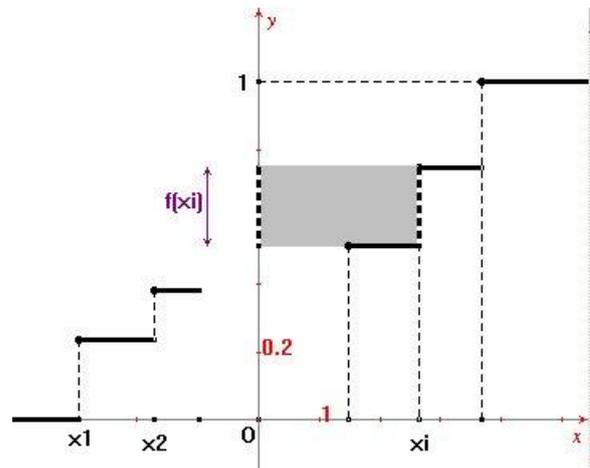
On peut vérifier que l'on a :

$$\sum_{i=1}^n f(x_i) [x_i - E(X)] = 0$$

car :

$$\sum_{i=1}^n f(x_i) [x_i - E(X)] = \sum_{i=1}^n f(x_i) x_i - E(X) \sum_{i=1}^n f(x_i) = E(X) - E(X) = 0$$

Cette relation se traduit géométriquement par le fait que la droite verticale  $x = E(X)$ , les droites horizontales  $y = 0$  et  $y = 1$ , et la courbe de  $F$  définissent deux surfaces d'aires  $A_1$  et  $A_2$  égales.



## 2.4.3 Remarque

L'espérance mathématique n'est pas seulement le résultat d'un calcul fait sur une loi de probabilité ; elle a une signification concrète, qui a été historiquement comprise avant même que la notion claire de probabilité n'ait été formulée avec précision.

Ce concept d'espérance mathématique se rattache de façon extrêmement étroite au problème de la **décision**. Il est naturel de penser que, en face de deux paris dont l'enjeu est égal, un sujet normal choisira le pari pour lequel la probabilité de gain est la plus grande. Par exemple : si l'on m'offre de gagner 100 Fr., soit en amenant *face* au jeu de pile ou face, soit en amenant un 6 par le jet d'un dé, il est clair que je choisirai la pièce. D'un autre point de vue, entre deux paris pour lesquels le hasard est égal, je choisirai naturellement celui dont l'enjeu est le plus grand : entre deux parties de pile ou face, l'une m'offrant 1000 Fr. et l'autre 100 Fr. en cas de succès, je préférerai la première.

La situation est légèrement plus complexe lorsqu'on quitte le domaine de ces événements extrêmement simples, pour comparer entre eux deux paris dont, ni les probabilités de gain, ni les enjeux ne sont égaux.

## 2.4.5 Exemples

- 1) Imaginons une loterie pour laquelle 10'000 billets sont mis en vente. Les lots promis sont :  
1 gros lot de 10'000 Fr.  
10 lots de 1'000 Fr. 50 lots de 300 Fr.

Bien que ce ne soit pas indispensable au raisonnement, nous supposons autorisé le cumul des lots, ce qui facilite le calcul des probabilités, car les divers tirages sont alors indépendants.

La probabilité  $p_1$  de gagner le gros lot est de  $\frac{1}{10'000}$  ;

la probabilité  $p_2$  de gagner un lot de 1'000 Fr. est de  $\frac{10}{10'000}$  ;

La probabilité  $p_3$  de gagner un lot de 300 Fr. est de  $\frac{50}{10'000}$  .

Par suite, l'espérance mathématique attachée à un billet est :

$$10'000 \times p_1 + 1'000 \times p_2 + 300 \times p_3 = 3,50 \text{ Fr.}$$

Quant au juste prix du billet, c'est le prix auquel l'organisateur de la loterie devrait le vendre pour ne faire ni bénéfice, ni déficit, sans tenir compte des frais d'exploitation.

Or, quel que soit le résultat des divers tirages, l'organisateur est certain d'avoir à payer :

1 gros lot de 10'000 Fr. : 10'000 Fr.

10 lots de 1'000 Fr. : 10'000 Fr.

50 lots de 300 Fr. : 15'000 Fr.

soit 35'000 Fr. .

Comme il vend 10'000 billets, le juste prix de chacun d'eux est précisément 3,50 Fr. .

2) Dans l'exemple a) du dé cubique on a ici,  $\mu_X = E(X) = \frac{2}{3} \times (-1) + \frac{1}{3} \times 2 = 0$  .

3) On considère l'expérience qui consiste à jeter deux dés parfaitement équilibrés. L'univers des possibles  $U$  est alors  $U = \{(1,1), (1,2), (1,3), \dots, (5,6), (6,6)\}$  avec  $\text{Card } U = 36$ . Les événements élémentaires sont équiprobables.

Soit la variable aléatoire  $X$  définie comme suit : soit  $i = (a,b)$  une issue quelconque de  $U$ , on pose  $X(i) = X(a,b) = \max(a,b)$  ; on a alors  $X(U) = \{1,2,3,4,5,6\}$ .

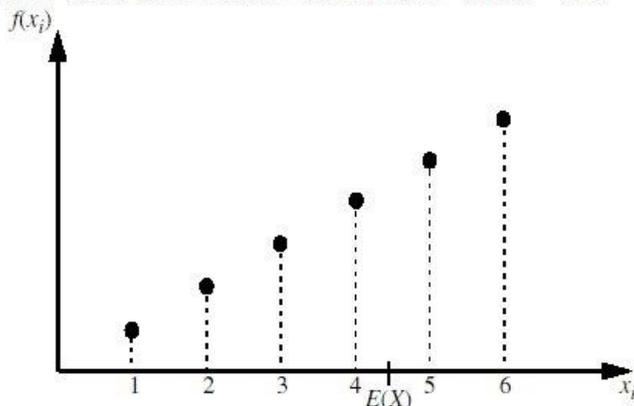
Soit la loi de probabilité  $f$  associée à  $X$ , on a

$$f(1) = P(X=1) = P(\{(1,1)\}) = \frac{1}{36} \quad ; \quad f(2) = P(X=2) = P(\{(1,2), (2,1), (2,2)\}) = \frac{3}{36} \text{ etc...}$$

Soit :

$x_i$	1	2	3	4	5	6
$f(x_i)$	1/36	3/36	5/36	7/36	9/36	11/36

$$E(X) = 1/36 + 6/36 + 15/36 + 28/36 + 45/36 + 66/36 = 161/36 = 4,47$$



## 2.4.6 Théorème

Soit  $X$  une variable aléatoire et  $k$  une constante réelle. On a :

a)  $E(kX) = k E(X)$

b)  $E(X+k) = E(X) + k$

preuve :

a)  $E(kX) = \sum_{i=1}^n (k x_i) \cdot f(x_i) = \sum_{i=1}^n k [x_i \cdot f(x_i)] = k \sum_{i=1}^n x_i \cdot f(x_i) = k E(X)$

b)  $E(X+k) = \sum_{i=1}^n (x_i + k) \cdot f(x_i) = \sum_{i=1}^n \{ [x_i \cdot f(x_i)] + [k \cdot f(x_i)] \} = \sum_{i=1}^n x_i \cdot f(x_i) + k \cdot \sum_{i=1}^n f(x_i)$   
 $= E(X) + k \cdot 1 = E(X) + k$

## 2.5 Espérance mathématique d'une loi infinie

### 2.5.1 Variables aléatoires infinies dénombrables

Tout ce qui a été vu précédemment dans le cas où  $U$  est fini se généralise au cas où  $U$  est infini dénombrable ; on aura par exemple :

$$\mu_X = E(X) = \sum_{i=1}^{\infty} x_i \cdot f(x_i) = x_1 f(x_1) + x_2 f(x_2) + x_3 f(x_3) + \dots . \text{ La somme converge à l'infini vers } E(X),$$

toutes les autres propriétés sont conservées, les sommes devenant des séries.

#### Exemple :

Dans l'exemple b) plus haut du jet d'une pièce de monnaie on a,

$$E(X) \text{ n'existe pas car } \sum_{i=1}^n p_i x_i = \frac{1}{2} \times 2 + \left(\frac{1}{2}\right)^2 \times 2^2 + \dots + \left(\frac{1}{2}\right)^n \times 2^n = 1 + 1 + \dots + 1 = n$$

et donc 
$$\lim_{n \rightarrow +\infty} \sum_{i=1}^n p_i x_i = +\infty .$$

### 2.5.2 Variables aléatoires continues

#### Définition

La généralisation au continu est délicate et même difficile si on ne dispose pas d'outils mathématiques adéquats. Les théorèmes du cours du calcul intégral nous sont ici nécessaires.

Nous nous contenterons de décrire la construction par analogie avec le cas discret.

Soit un univers  $U$  infini non dénombrable ( $\text{Card } U = \aleph_1$ ), dont certaines parties seulement sont des événements et peuvent donc se voir affecter une probabilité.

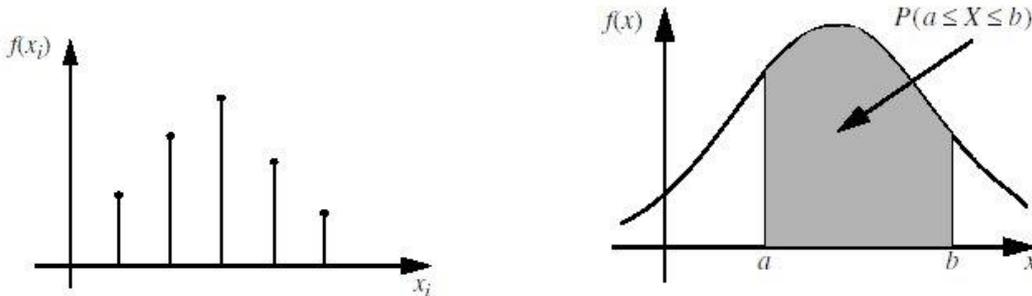
Une variable aléatoire  $X$  dont l'ensemble image  $X(U)$  est un intervalle réel est une variable aléatoire **continue** (par opposition à discrète) si l'image réciproque par  $X$  de tout intervalle réel  $I$  est un événement  $A$  de  $U$  ( $A \subset U$ ). L'événement  $a \leq X \leq b$  de  $U$  possède alors une probabilité bien définie.

On définit la loi de probabilité de  $X$  à l'aide d'une fonction  $f$ , appelée alors **densité de probabilité**

de  $X$ , telle que  $P(a \leq X \leq b) = \int_a^b f(x) dx$ .

**Remarques :**

- 1) Si  $f$  est donnée, la probabilité  $P(a \leq X \leq b)$  est l'aire de la surface délimitée par la courbe  $\Gamma_f$ , l'axe (OI) et les verticales  $x=a$  et  $x=b$  :
- 2) Le passage du discret au continu transforme les sommes  $\Sigma$  en intégrales  $\int$  et  $f(x_i)$  en  $f(x)dx$ . Ainsi, soit  $X$  une variable aléatoire discrète et  $f$  sa distribution :



La formule  $P(x_k \leq X \leq x_n) = \sum_{i=k}^n f(x_i)$  est analogue à  $P(a \leq X \leq b) = \int_a^b f(x) dx$ .

En utilisant cette analogie, on admettra les propriétés suivantes pour une variable aléatoire  $X$ , continue, de distribution  $f$  :

- 1)  $f(x) \geq 0$  (analogue à  $f(x_i) \geq 0$ )
- 2)  $\int_{-\infty}^{\infty} f(x) dx = 1$  (analogue à  $\sum_i f(x_i) = 1$ )
- 3)  $P(a \leq X \leq b) = \int_a^b f(x) dx = F(b) - F(a)$ .
- 4)  $F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$  (analogue à  $\sum_{x_i \leq x} f(x_i)$ )
- 5)  $\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$  (analogue à  $\sum_i x_i \cdot f(x_i)$ )

(l'espérance d'une variable aléatoire continue donne l'abscisse du centre de gravité de la surface comprise entre le graphique  $\Gamma_f$  de la densité  $f$  et l'axe des abscisses (OI) : cf. dans les tables

numériques à la page 91  $x_A = \frac{1}{A} \int_{-\infty}^{+\infty} x f(x) dx$  avec, ici,  $A = \int_{-\infty}^{\infty} f(x) dx = 1$ )

**exemple :** Dans l'exemple c) plus haut sur la durée de bon fonctionnement d'un équipement particulier fabriqué en grande série, on a  $\mu_X = E(X) =$

$$\int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^0 x f(x) dx + \int_0^{+\infty} x f(x) dx = \int_{-\infty}^0 x \times 0 dx + \int_0^{+\infty} x \times 0,002 e^{-0,002x} dx = \int_0^{+\infty} 0,002 x e^{-0,002x} dx = 500$$

(à l'aide d'une intégration par partie).

## La loi de Cauchy

La loi de Cauchy est définie par la fonction  $f : y = f(x) = \frac{1}{\pi(1+x^2)}$  ; on a bien  $f(x) \geq 0, \forall x \in \mathbb{R}$ ,

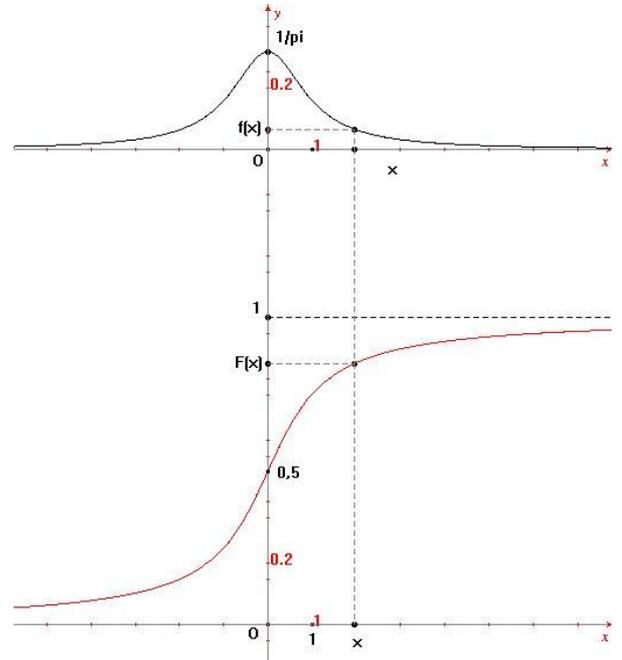
$$\text{et } \int_{-\infty}^{+\infty} f(x) dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{1}{1+x^2} dx = \frac{1}{\pi} [\arctan(x)]_{-\infty}^{+\infty} = \frac{1}{\pi} \left[ \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) \right] = 1$$

mais cette loi n'a pas d'espérance mathématique :

$$E(x) = \int_{-\infty}^{+\infty} x \cdot f(x) dx = \int_{-\infty}^{+\infty} \frac{1}{\pi} \frac{x}{1+x^2} dx = \left[ \frac{1}{2\pi} \ln(1+x^2) \right]_{-\infty}^{+\infty}$$

qui est une intégrale divergente.

Cette loi ne sert à rien, sauf de contre-exemple dans les cours de probabilités.



## Variance - Ecart type

### 3.1 Définitions

Après avoir traduit la tendance centrale par l'espérance, il est nécessaire de traduire la dispersion autour de l'espérance par les valeurs de la variance et de l'écart type.

La **variance** de  $X$ , notée  $var(X)$  ou  $\sigma^2_X$  ou  $V(X)$ , est le réel  $V(X) = E[(X - \mu_X)^2]$  où  $\mu_X = E(X)$ .

L'**écart type** de  $X$ , noté  $\sigma(X)$  ou  $\sigma_X$  ou  $S(X)$ , est le réel  $S(X) = \sqrt{V(X)}$ .

La variance et l'écart type sont des indicateurs de dispersion. L'écart type indique comment en moyenne les valeurs de la variable sont groupées autour de la moyenne (espérance mathématique).

Un faible écart type signifie que les valeurs sont peu dispersées autour de la moyenne. Statistiquement, on observe que 75 % au minimum des observations sont comprises entre  $E(X) - 2S(X)$  et  $E(X) + 2S(X)$ .

La variance traduit la dispersion de la distribution de la variable aléatoire autour de sa valeur moyenne. Etant un carré, la dimension de la variance n'est pas celle de la moyenne. C'est pourquoi on utilise plus souvent l'**écart type** qui est la racine de la variance.

On dit aussi que la variance traduit la notion d'incertitude. Plus la variance est faible, moins le résultat de l'expérience aléatoire est incertain. A la limite, une variable aléatoire de variance nulle conduit à des expériences strictement identiques (*i.e.* le phénomène est complètement déterministe, il n'y a donc plus aucune raison de garder la notion de variable aléatoire).

### 3.2 Propriétés

$$1) \quad V(X) = E[(X - \mu_X)^2] = E(X^2) - \mu_X^2 \quad (\text{théorème de Koenig})$$

$$\begin{aligned}
 \text{preuve : } V(X) &= E[(X - \mu_X)^2] = \sum_{i=1}^n (x_i - \mu_X)^2 \cdot f(x_i) = \sum_{i=1}^n (x_i^2 - 2\mu_X x_i + \mu_X^2) \cdot f(x_i) \\
 &= \sum_{i=1}^n x_i^2 \cdot f(x_i) - 2\mu_X \sum_{i=1}^n x_i \cdot f(x_i) + \mu_X^2 \sum_{i=1}^n f(x_i) = E(X^2) - 2\mu_X^2 + \mu_X^2 = E(X^2) - \mu_X^2 \\
 &\quad \text{(car } \sum_{i=1}^n f(x_i) = 1)
 \end{aligned}$$

2)  $V(X) = E[(X - \mu_X)^2] \geq 0$ , par définition ;

3) Si  $k$  est une constante, on montre que  $V(X+k) = V(X)$  et  $S(X+k) = S(X)$

preuve :

$$\begin{aligned}
 V(X+k) &= \sum_{i=1}^n [x_i + k - E(x+k)]^2 \cdot f(x_i) = \\
 &= \sum_{i=1}^n [x_i + k - E(x) - k]^2 \cdot f(x_i) = \\
 &= \sum_{i=1}^n [x_i - E(x)]^2 \cdot f(x_i) = V(X)
 \end{aligned}$$

et

$$V(kX) = k^2 \cdot V(X) \text{ et } S(kX) = |k| S(X)$$

preuve :

$$\begin{aligned}
 V(kX) &= \sum_{i=1}^n [kx_i - E(kx)]^2 \cdot f(x_i) = \\
 &= \sum_{i=1}^n [kx_i - kE(x)]^2 \cdot f(x_i) = \\
 &= \sum_{i=1}^n k^2 [x_i - E(x)]^2 \cdot f(x_i) = \\
 &= k^2 \sum_{i=1}^n [x_i - E(x)]^2 \cdot f(x_i) = k^2 \cdot V(X)
 \end{aligned}$$

4) **Variable aléatoire centrée et réduite**

Soit  $X$  une variable aléatoire d'espérance  $\mu$  et d'écart-type  $\sigma$ .

Si l'on définit la variable aléatoire  $Y_1 = X - \mu$ , appelée variable **centrée**, alors  $E(Y_1) = 0$

preuve :

$$E(Y_1) = \mu_{Y_1} = E(X - \mu) = E(X) - \mu = 0 \text{ (cf. thms 2.4.3-1) ;}$$

Si l'on définit la variable aléatoire  $Y_2 = \frac{X}{\sigma}$ , appelée variable **réduite**, alors  $S(Y_2) = 1$ ,

preuve :

$$V(Y_2) = V\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma^2} V(X) = \frac{\sigma^2}{\sigma^2} = 1 ; \text{ et donc l'écart-type } \sigma = 1 ,$$

Si l'on définit la variable aléatoire  $Y = \frac{X - \mu}{\sigma}$ , appelée variable **centrée-réduite**, alors

$$\mu_Y = E(Y) = \frac{1}{\sigma} E(X - \mu) = 0 \quad \text{et} \quad V(Y) = V\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma^2} V(X - \mu) = \frac{1}{\sigma^2} V(X) = \frac{\sigma^2}{\sigma^2} = 1$$

et donc  $\sigma_Y = 1$ .

Cette variable  $Y$ , notée parfois  $Y = X^*$ , s'appelle la **variable centrée-réduite** associée à  $X$ .

On a ainsi "normalisé" la variable  $X$  par un changement d'origine et d'échelle de manière à obtenir une variable d'espérance 0 et d'écart-type 1.

### 3.3 Exemples

Reprenons les exemples du 1.2

a) On avait  $\mu_X = 0$ , d'où  $V(X) = E(X^2) - \theta^2 = \left(\frac{2}{3} \times (-1)^2 + \frac{1}{3} \times (2)^2\right) = 2$  et  $S(X) = \sqrt{2}$ .

b) Puisque  $E(X)$  n'existe pas,  $V(X)$  et  $S(X)$  non plus.

c) On avait  $\mu_X = E(X) = 500$ , d'où

$$V(X) = E(X^2) - \mu_X^2 = \int_{-\infty}^{+\infty} t^2 f(t) dt - 500^2 = 250'000$$

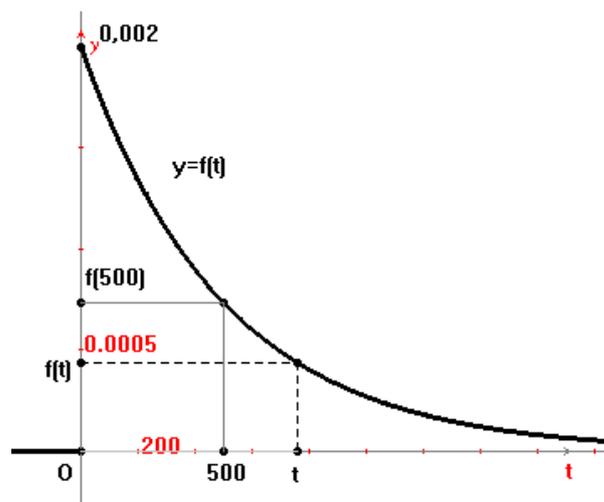
(à l'aide d'une intégration par parties)

et  $S(X) = 500$ .

Calculons la probabilité pour que la durée de bon fonctionnement d'un équipement particulier fabriqué en grande série soit comprise entre  $E(X) - S(X)$  et  $E(X) + S(X)$  :

$$P(0 \leq X \leq 1000) = F(1000) =$$

$$\int_0^{1000} f(t) dt = \int_0^{1000} 0,002 e^{-0,002t} dt = \left[-e^{-0,002t}\right]_0^{1000} = -e^{-2} + 1 \cong 0,865$$



d) On jette une pièce de monnaie une seule fois. Soit  $X$  la variable aléatoire indiquant le nombre de faces qui se présentent. On a :

$x_i$	$f(x_i)$	$x_i f(x_i)$	$[x_i - E(X)]$	$[x_i - E(X)]^2$	$f(x_i)(x_i - E(X))^2$
0	$\frac{1}{2}$	0	$-\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$
1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$
total :		$E(X) = \frac{1}{2}$			$V(X) = \frac{1}{4}$

D'où  $E(X) = \frac{1}{2}$ ,  $V(X) = \frac{1}{4}$  et  $S(X) = \frac{1}{2}$ .

e) On jette un dé. Soit  $X$  la variable aléatoire indiquant le nombre de points sortis lors du jet.  
On a :

$x_i$	$f(x_i)$	$x_i f(x_i)$	$(x_i)^2$	$(x_i)^2 \cdot f(x_i)$
1	$\frac{1}{6}$	$\frac{1}{6}$	1	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{2}{6}$	4	$\frac{4}{6}$
3	$\frac{1}{6}$	$\frac{3}{6}$	9	$\frac{9}{6}$
4	$\frac{1}{6}$	$\frac{4}{6}$	16	$\frac{16}{6}$
5	$\frac{1}{6}$	$\frac{5}{6}$	25	$\frac{25}{6}$
6	$\frac{1}{6}$	$\frac{6}{6}$	36	$\frac{36}{6}$
total :		$E(X) = \frac{21}{6} = \frac{7}{2} = 3,5$		$E(X^2) = \frac{91}{6}$

D'où  $E(X) = \frac{21}{6} = 3,5$  ,  $V(X) = E(X^2) - [E(X)]^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12} \cong 2,92$  et  $S(X) = \sqrt{\frac{35}{12}} \cong 1,71$ .

Calculons la probabilité que le nombre de points soit compris entre  $E(X) - S(X)$  et  $E(X) + S(X)$  :

c'est-à-dire  $P(1,79 \leq X \leq 5,21) = P(\{2,3,4,5\}) = \frac{4}{6} \cong 0,67$ .

f) Un arrêt de tramway est desservi toutes les 10 minutes.

Soit  $X$  la variable aléatoire indiquant le temps d'attente (en minutes) jusqu'au passage du prochain convoi lorsqu'on se rend à cet arrêt sans tenir compte de l'horaire.

Il est raisonnable d'admettre que  $X$  admet pour densité de probabilité la fonction  $f$  suivante :

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \text{ ou si } x > 10 \\ \frac{1}{10} & \text{si } 0 \leq x \leq 10 \end{cases}$$

$$\text{On a par exemple : } P(6 \leq x \leq 10) = \int_6^{10} \frac{1}{10} dx = \left[ \frac{x}{10} \right]_6^{10} = \frac{4}{10} = 0,4$$

Calculons l'espérance :

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^0 x f(x) dx + \int_0^{10} x f(x) dx + \int_{10}^{+\infty} x f(x) dx = 0 + \int_0^{10} \frac{x}{10} dx + 0 = \left[ \frac{x^2}{20} \right]_0^{10} = \frac{100}{20} = 5$$

l'attente dure en moyenne 5 minutes (logique ici).

$$\text{Calculons la variance } V(X) = \int_0^{10} x^2 f(x) dx - 25 = \int_0^{10} \frac{x^2}{10} dx = \left[ \frac{x^3}{30} \right]_0^{10} - 25 = \frac{1000}{30} - 25 \cong 8,333.$$

D'où l'écart-type  $\sigma = S(X) = \sqrt{V(X)} \cong \sqrt{8,333} \cong 2,887$ .

Calculons la probabilité que l'attente dure entre  $E(X) - S(X)$  et  $E(X) + S(X)$ ,

$$\text{c'est-à-dire entre } 2,113 \text{ et } 7,887 : P(2,113 \leq x \leq 7,887) = \int_{2,113}^{7,887} \frac{1}{10} dx = \left[ \frac{x}{10} \right]_{2,113}^{7,887} \cong 0,577.$$

g) Un exemple simple : imaginons une ficelle, dont une longueur  $h$  est délimitée entre deux mâchoires, qu'on écarte avec une force croissante, jusqu'à ce qu'elle casse. L'hypothèse est faite que la ficelle est suffisamment homogène pour qu'il n'existe aucun point en lequel la rupture soit plus probable qu'en un autre. La distance  $X$  du point de rupture à la mâchoire de gauche figure une variable aléatoire continue, qui peut prendre toutes les valeurs entre 0 et  $h$ . Mais quelle valeur attribuer à la probabilité de rupture en l'un de ces points, par exemple le milieu ? Aucune non valeur nulle ne convient : en effet, puisqu'il y a une infinité de points de rupture possibles, la somme des probabilités ainsi attachées serait infinie, et non égale à 1. On résout cette difficulté en parlant d'une densité de probabilité  $f$ , la probabilité de voir la rupture se produire sur un petit segment de longueur  $s$  étant alors égale au produit de  $s$  par cette densité. Il ne s'agit pas là d'une abstraction révolutionnaire : on aurait eu exactement la même difficulté si on avait prétendu déterminer la masse d'un point de la ficelle ; aucune valeur autre que 0 n'aurait pu convenir, alors qu'on admettra volontiers que la masse d'un fragment de ficelle est proportionnelle à sa longueur, et plus précisément, égale au produit de sa longueur par sa densité.

La densité de probabilité est d'ailleurs facile à calculer : la rupture étant certaine, la probabilité attachée à la longueur  $h$ , qui est égale au produit  $hf$  de la longueur  $h$  par la densité  $f$ , est égale à 1.

La densité  $f$  est donc égale à l'inverse de la longueur :  $f = \frac{1}{h}$ .

On peut d'ailleurs aussi s'intéresser à la probabilité de voir la rupture se produire à une distance de la mâchoire gauche inférieure à une longueur donnée  $x$ . Cet événement est justement symbolisé par l'écriture  $X < x$ , et sa probabilité est  $f \cdot x$ . La fonction  $F$  telle que  $F(x) = P(X < x)$  s'appelle fonction de répartition de la variable  $X$  ; dans notre exemple, c'est une fonction très simple, la fonction linéaire définie sur l'intervalle  $[0, h]$  par :  $F(x) = f \cdot x$ .

On peut remarquer que  $f$  est la dérivée de  $F$  ou, ce qui est peut-être plus facile à interpréter,

que  $F(x)$  peut s'écrire :  $F(x) = \int_0^x f(t) dt$ . La loi de probabilité d'une variable aléatoire continue

est donc donnée soit par sa densité, soit par sa fonction de répartition.

La loi de probabilité présentée dans cet exemple est appelée **loi uniforme** sur l'intervalle  $[0, h]$ . La densité est donc nulle en dehors de l'intervalle  $[0, h]$ , égale à  $f$  sur  $[0, h]$  ; quant à la fonction de répartition, elle est nulle sur  $]-\infty, 0]$ , égale à  $f \cdot x$  de  $]0, h]$ , et égale à 1 sur  $]h, +\infty[$ .

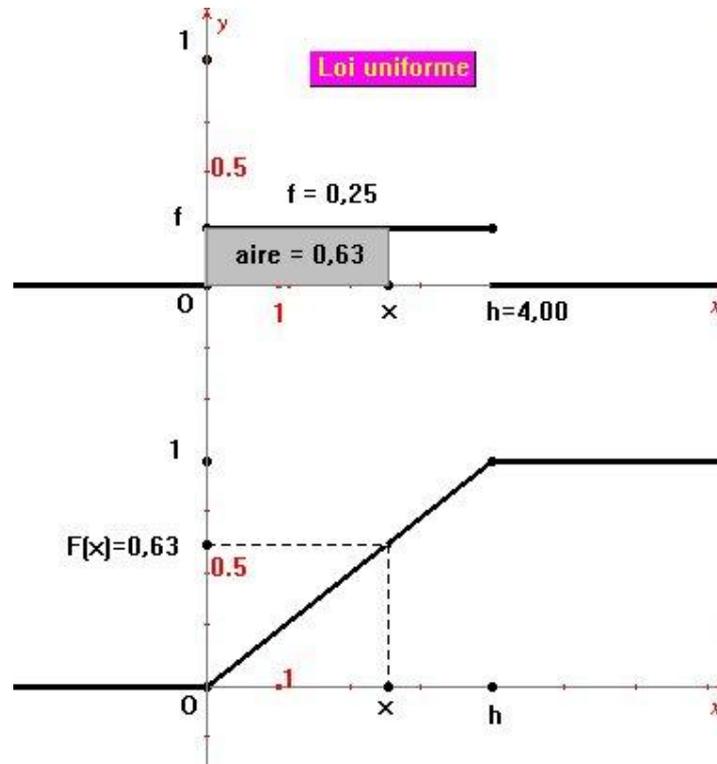
On aura l'espérance mathématique :

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^h x \cdot f dx = \left[ f \cdot \frac{1}{2} x^2 \right]_0^h = \frac{1}{2} h^2 f = \frac{h}{2} \text{ (car } hf = 1)$$

et la variance et l'écart-type :

$$E(X^2) = \int_0^h x^2 \cdot f dx = \left[ f \cdot \frac{1}{3} x^3 \right]_0^h = \frac{1}{3} h^3 f = \frac{h^2}{3} \text{ d'où } V(X) = E(X^2) - [E(X)]^2 = \frac{h^2}{3} - \frac{h^2}{4} = \frac{h^2}{12} \text{ et}$$

$$S(X) = \sqrt{V(X)} = \sqrt{\frac{h^2}{12}} = \frac{h}{\sqrt{12}} = \frac{h}{2\sqrt{3}} = \frac{\sqrt{3}}{6} h \cong 0,29 h$$



de plus  $P(E(X) - S(X) \leq X \leq E(X) + S(X)) = P\left(\frac{h}{2} - \frac{\sqrt{3}h}{6} \leq X \leq \frac{h}{2} + \frac{\sqrt{3}h}{6}\right)$

$$= P\left(\frac{3-\sqrt{3}}{6}h \leq X \leq \frac{3+\sqrt{3}}{6}h\right) = P(0,211h \leq X \leq 0,789h)$$

$$= \int_{0,211h}^{0,789h} f \, dx = [f \cdot x]_{0,211h}^{0,789h} = f \left(\frac{3+\sqrt{3}}{6} - \frac{3-\sqrt{3}}{6}\right)h = f \left(\frac{\sqrt{3}}{3}\right)h = \frac{\sqrt{3}}{3} \cong 0,58 \text{ (car } f \cdot h = 1)$$

Les exemples qui précèdent montrent que pour de nombreuses variables aléatoires la probabilité que la variable s'écarte de sa moyenne d'au plus un écart-type est à peu près égale à 0,7.

### 3.4 Inégalité de Bienaymé-Tchébychev

La signification concrète de l'écart-type (ou de la variance) en termes de probabilités est bien imagée par l'inégalité de **Bienaymé-Tchébychev**. Soit  $k$  un nombre réel positif non nul, on peut démontrer que la probabilité de l'ensemble des valeurs de la variable aléatoire qui sont à l'extérieur de l'intervalle

$[E(X) - k, E(X) + k]$  est inférieure à  $\frac{\sigma^2}{k^2}$  :

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

ou ce qui est équivalent :  $P(E(X) - k \leq X \leq E(X) + k) = P(|X - E(X)| \leq k) \geq 1 - \frac{\sigma^2}{k^2}$ .

**Démonstration** : (dans le cas discret)

Soit  $X$  une variable aléatoire discrète à valeurs positives :  $f(x_i) \geq 0, \forall i$ .

On a  $E(X) = \sum_{i=1}^n x_i \cdot f(x_i) = x_1 f(x_1) + x_2 f(x_2) + x_3 f(x_3) + \dots + x_n f(x_n)$ .

Soit  $a$  un nombre positif tel que  $0 < x_{i-1} < a \leq x_i$ .

On a les inégalités :  $E(X) \geq x_1 f(x_1) + \dots + x_n f(x_n) \geq a \cdot [f(x_1) + \dots + f(x_n)]$ .

Or la somme  $[f(x_1) + \dots + f(x_n)]$  est égale à la probabilité de l'événement  $(X \geq a)$ .

On en déduit que :  $P(X \geq a) \leq \frac{E(X)}{a}$  (appelée **inégalité de Markov**) ;

elle est encore vraie si  $x_n < a$ , car alors  $P(X \geq a) = 0$  et  $E(X) > 0$ .

Posons  $a = k^2$  ( $k > 0$ ) et remplaçons  $X$  par le carré de la variable centrée  $[X - E(X)]^2$  :

l'inégalité de Markov devient :  $P([X - E(X)]^2 \geq k^2) \leq \frac{E[X - E(X)]^2}{k^2}$ ,

puisque l'espérance de la variable  $[X - E(X)]^2$  est la variance  $\sigma^2$  de  $X$ , cette inégalité est équivalente à celle-ci :

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

Autre formulation équivalente de l'inégalité de **Bienaymé-Tchébychev** : (en posant  $k = \sigma t$ )

Soit  $X$  une variable aléatoire d'espérance  $\mu$  et d'écart-type  $\sigma$ ,

Pour tout nombre positif  $t$ , on a :  $P(|X - \mu| \geq \sigma t) \leq \frac{1}{t^2}$

Ainsi, la probabilité de l'événement  $(E(X) - 2\sigma \leq X \leq E(X) + 2\sigma)$  ( $t=2$ ) sera supérieure à 0,75. Il faut bien voir qu'il s'agit d'un résultat général, valable pour n'importe quelle variable aléatoire. Mais dans la pratique, on constate très souvent qu'une bien plus grande probabilité est attachée à un tel événement.

**Exemples :**

1) Reprenons l'exemple c) de §1.2 : On avait  $\mu = E(X) = 500$  et  $\sigma = S(X) = 500$  ;

$$d'où P(E(X) - 2\sigma \leq X \leq E(X) + 2\sigma) = P(0 \leq X \leq 1500) = F(1500) = \int_0^{1500} f(t) dt = \int_0^{1500} 0,002 e^{-0,002t} dt = [-e^{-0,002t}]_0^{1500} = -e^{-3} + 1 \cong 0,95 (\leq 1).$$

2) Dans l'exemple f) précédent : on avait  $\mu = E(X) = 5$  et  $\sigma = S(X) = 2,887$  ;

$$d'où P(E(X) - 2\sigma \leq X \leq E(X) + 2\sigma) = P(0 \leq X \leq 10,774) = 1 (\leq 1).$$

3) Dans l'exemple précédent de la ficelle on a :

$$P(E(X) - 2S(X) \leq X \leq E(X) + 2S(X)) = P\left(\frac{h}{2} - \frac{\sqrt{3}h}{3} \leq X \leq \frac{h}{2} + \frac{\sqrt{3}h}{3}\right) \\ = P\left(\frac{3-2\sqrt{3}}{6}h \leq X \leq \frac{3+2\sqrt{3}}{6}h\right) = P(-0,08h \leq X \leq 1,08h) = P(0 \leq X \leq h) = 1.$$

## Exemples de distributions

### 4.1 Lois discrètes

#### 4.1.1 Loi de Bernoulli

##### Définition

On considère une expérience n'ayant que deux résultats possibles, par exemple succès et échec (ou présence et absence d'une certaine caractéristique). On introduit la variable aléatoire  $X$  qui associe

la valeur 0 à l'échec (ou à l'absence de la caractéristique) et la valeur 1 au succès (ou à la présence de la caractéristique). Cette variable aléatoire est appelée **variable de Bernoulli**.

**Distribution de X**

Appelons  $p$  la probabilité de l'événement succès :

$$P(\{\text{succès}\}) = P(X=1) = p$$

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

$k$	0	1
$f(k) = P(X = k)$	$p$	$q$

**Espérance de X**

$$\mu_X = E(X) = \sum_{i=1}^2 x_i \cdot P(X = x_i) = 1 \cdot P(X=1) + 0 \cdot P(X=0) = 1 \cdot p + 0 \cdot q = p$$

**Variance de X**

$$\sigma^2_X = V(X) = E[(X - \mu_X)^2] = E(X^2) - \mu_X^2$$

$$= [1^2 \times P(X=1) + 0^2 \times P(X=0)] - p^2 = p - p^2 = p(1-p) = pq.$$

**4.1.2 Loi binomiale  $\mathcal{B}(n, p)$**

**Définition**

Soient les épreuves  $n$  fois répétées et indépendantes d'une même expérience de Bernoulli. Chaque expérience n'a que deux résultats possibles : *succès* ou *échec*. Comme précédemment, appelons  $p$  la probabilité de l'événement élémentaire *succès* et  $q=1-p$  celle de l'événement *échec*. A cette expérience multiple on associe une variable aléatoire  $X$  qui mesure le nombre de succès obtenus.

**Distribution de X**

La probabilité d'avoir  $k$  succès lors de  $n$  épreuves successives est

$$P(X=k \text{ pour } n \text{ essais}) = C_k^n p^k \cdot (1-p)^{n-k} = C_k^n p^k \cdot q^{n-k}$$

**Remarques**

- a) La probabilité de n'avoir aucun succès au cours de  $n$  épreuves successives ( $k=0$ ) est  $q^n$  ; la probabilité d'avoir au moins un succès est donc  $1 - q^n$  (un succès ou plus).
- b) Les  $C_k^n$ , notés aussi  $\binom{n}{k}$ , s'appellent coefficients du binôme, car ils interviennent dans le développement du binôme de Newton :  $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$ .
- c) En appliquant la formule du binôme, on montre que la somme des probabilités pour toutes les valeurs de  $X$  est égale à 1 :  $\sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = [p + q]^n = [p + (1-p)]^n = [1]^n = 1$ .

**Espérance de X**

Si  $X$  suit la loi binomiale  $\mathcal{B}(n,p)$  alors  $E(X) = np$ ,  $V(X) = npq$  et  $S(X) = \sqrt{npq}$ .

On note  $P(X=k) = b(k ; n, p) = p_k$ .

Preuve : La variable  $X$  d'une loi binômiale  $\mathcal{B}(n,p)$  est la somme de  $n$  variables  $X_i$  de Bernoulli ; par un théorème non présenté, l'espérance de la somme est égale à la somme des

espérances : d'où  $E(X) = \sum_{i=1}^n E(X_i) = np$ .

Autre démonstration :

$$E(X) = \sum_{k=0}^n k \cdot P(X=k) = \sum_{k=1}^n k \cdot C_k^n p^k (1-p)^{n-k} ;$$

$$\text{or } k \cdot C_k^n = k \cdot \frac{n!}{k!(n-k)!} = \frac{n!}{(k-1)!(n-k)!} = n \cdot \frac{(n-1)!}{(k-1)!(n-k)!} = n \cdot C_{k-1}^{n-1}, \forall k \geq 1 ;$$

$$\begin{aligned} \text{d'où } E(X) &= \sum_{k=1}^n n \cdot C_{k-1}^{n-1} p^k (1-p)^{n-k} = np \cdot \sum_{k=1}^n C_{k-1}^{n-1} p^{k-1} (1-p)^{n-k} = \\ & np \cdot \left[ C_0^{n-1} p^0 (1-p)^{n-1} + C_1^{n-1} p^1 (1-p)^{n-2} + \dots + C_{n-1}^{n-1} p^{n-1} (1-p)^0 \right] \quad (\text{formule du binôme}) \\ & np \cdot [p + (1-p)]^{n-1} = np \cdot 1 = np \end{aligned}$$

de plus, comme les  $n$  variables de Bernoulli sont indépendantes, la variance de la loi  $B(n,p)$  est égale

à la somme  $n$  des variances de Bernoulli (*théorème non présenté*):  $V(X) = \sum_{i=1}^n V(X_i) = npq$ .

Autre démonstration :

$$\text{Par le théorème de Koenig on a : } V(X) = E(X^2) - E(X)^2 = \sum_{k=0}^n k^2 \cdot p_k - (E(X))^2 ;$$

$$\text{calculons le nombre } A = \sum_{k=2}^n k(k-1)p_k = \sum_{k=2}^n k(k-1) C_k^n p^k q^{n-k} \quad (*):$$

$$\forall k \geq 2, k(k-1) C_k^n = k(k-1) \frac{n!}{k!(n-k)!} = \frac{n!}{(k-2)!(n-k)!} = n(n-1) \frac{(n-2)!}{(k-2)!(n-k)!} = n(n-1) C_{k-2}^{n-2}$$

On obtient en reportant dans (\*)

$$\begin{aligned} A &= \sum_{k=2}^n k(k-1) C_k^n p^k q^{n-k} = \sum_{k=2}^n n(n-1) C_{k-2}^{n-2} p^k q^{n-k} = n(n-1)p^2 \sum_{k=2}^n C_{k-2}^{n-2} p^{k-2} q^{n-k} = \\ & n(n-1)p^2 [C_0^{n-2} p^0 q^{n-2} + C_1^{n-2} p^1 q^{n-3} + \dots + C_{n-2}^{n-2} p^{n-2} q^0] = n(n-1)p^2 [p+q]^{n-2} = n(n-1)p^2 [1]^{n-2} \end{aligned}$$

donc  $A = n(n-1)p^2 (**)$ ;

de plus :

$$A = \sum_{k=2}^n k(k-1)p_k = \sum_{k=0}^n k(k-1)p_k \text{ car les termes correspondant à } k=0 \text{ et } k=1 \text{ sont nuls ;}$$

$$A = \sum_{k=0}^n k(k-1)p_k = \sum_{k=0}^n (k^2 p_k - k p_k) = \sum_{k=0}^n k^2 p_k - \sum_{k=0}^n k p_k = \sum_{k=0}^n k^2 p_k - E(X) = \sum_{k=0}^n k^2 p_k - np ;$$

$$\text{d'où } \sum_{k=0}^n k^2 p_k = A + np \stackrel{(**)}{=} n(n-1)p^2 + np ;$$

et ainsi :

$$V(X) = E(X^2) - E(X)^2 = \sum_{k=0}^n k^2 \cdot p_k - (E(X))^2 = n(n-1)p^2 + np - (np)^2 = n^2p^2 - np^2 + np - n^2p^2 = np - np^2 = np(1-p) = npq.$$

On obtient bien en définitive :  $V(X) = npq$

### Exemples

1) En 1975, il est né en Suisse 75'537 enfants dont 38'922 garçons. La probabilité de naissance d'un garçon est ainsi 0,51527. On prélève au hasard 1000 enfants parmi ceux-ci. Combien de garçons peut-on s'attendre à trouver dans cet échantillon (on admettra comme très improbable un écart supérieur à 3 écart-types par rapport à l'espérance) ?

On a ici  $n = 1000$  et  $p = 0,51527$  et  $q = 0,48473$  ;

d'où  $E(X) = 515,27$  et  $S(X) = \sqrt{npq} \cong 15,8$ .

Le nombre de garçons de l'échantillon sera compris fort probablement entre 467 et 563.

2) On jette 6 fois une pièce bien équilibrée ; on suppose que face est un succès. On a donc  $p = q = \frac{1}{2}$ .

On a une loi binomiale  $\mathcal{B}(6, \frac{1}{2})$ . Soit :

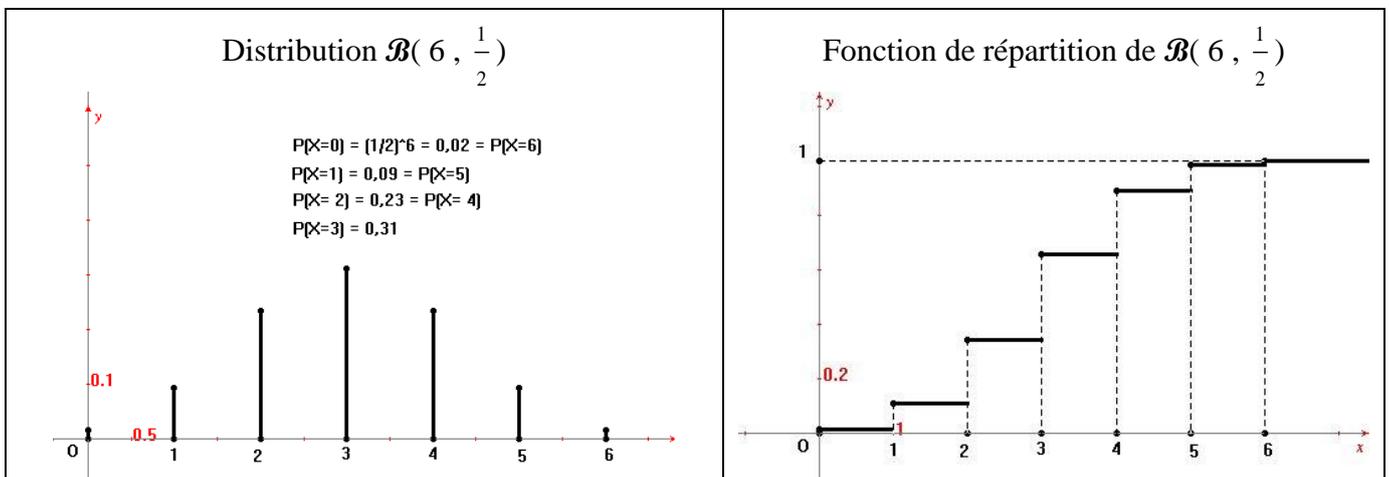
$$b(0,6, \frac{1}{2}) = C_0^6 \left(\frac{1}{2}\right)^0 \cdot \left(1 - \frac{1}{2}\right)^{6-0} = C_0^6 \left(\frac{1}{2}\right)^6 = \left(\frac{1}{2}\right)^6 = 0,015625 = b(6,6, \frac{1}{2})$$

$$\text{et } b(1,6, \frac{1}{2}) = C_1^6 \left(\frac{1}{2}\right)^6 = 6 \cdot \left(\frac{1}{2}\right)^6 = 0,09375 = b(5,6, \frac{1}{2})$$

$$\text{et } b(2,6, \frac{1}{2}) = C_2^6 \left(\frac{1}{2}\right)^6 = 15 \cdot \left(\frac{1}{2}\right)^6 = 0,234375 = b(4,6, \frac{1}{2})$$

$$\text{et } b(3,6, \frac{1}{2}) = C_3^6 \left(\frac{1}{2}\right)^6 = 20 \cdot \left(\frac{1}{2}\right)^6 = 0,3125$$

L'espérance est  $E(X) = np = 6 \cdot \frac{1}{2} = 3$  et l'écart-type est  $S(X) = \sqrt{npq} = \sqrt{6 \cdot \frac{1}{2} \cdot \frac{1}{2}} = \sqrt{\frac{3}{2}} \cong 1,225$



### 4.1.3 Loi de Poisson

Le calcul de  $P(X=k) = b(k;n,p)$  est long et fastidieux et les tables numériques habituelles ne donnent que quelques valeurs particulières pour des valeurs de n petites. Il est donc souhaitable de disposer d'approximations plus maniables, la loi de Poisson est l'une de ces approximations. Elle repose sur un résultat mathématique que nous admettrons sans démonstration :

Lorsque n augmente indéfiniment et que, simultanément, p tend vers 0 en sorte que le produit np ait une limite finie  $\lambda$ ,

$$\text{alors } P(X=k) = b(k;n,p) = C_k^n p^k \cdot (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k \cdot q^{n-k} \text{ a pour limite } e^{-\lambda} \frac{\lambda^k}{k!}.$$

Ceci suggère de définir une variable aléatoire qui prendrait pour valeurs tous les nombres entiers (puisque n augmente indéfiniment, aucune limitation n'intervient désormais sur les valeurs de k),

$$\text{la probabilité de la valeur k étant : } P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Une telle variable est appelée **variable de Poisson**,  $\lambda$  est le paramètre. Une loi de Poisson comportant un seul paramètre, il est plus facile et moins encombrant d'en présenter une table numérique.

Une telle loi intervient dans des expériences aléatoires dont les résultats futurs sont indépendants des résultats passés.

De par la façon dont nous l'avons introduite, la loi de Poisson apparaît comme une approximation de la loi binomiale pour des grandes valeurs de n et des petites valeurs de p : c'est pourquoi elle est parfois appelée loi des petites probabilités. Voici quelques exemples :

- Probabilité, dans une page d'un livre, de trouver h coquilles ;
- probabilité, dans une petite unité de temps, de voir se présenter à un guichet h clients ;
- probabilité pour un bloc d'immeubles d'une ville bombardée de recevoir h obus au cours d'un raid.

Pour nous faire une idée de la qualité de l'approximation d'un variable aléatoire binomiale par une variable de Poisson, nous allons imaginer qu'on jette au hasard et indépendamment 64 grains de blés sur un échiquier (qui comporte 64 cases), et nous examinons la variable aléatoire qui exprime le nombre de grains tombant sur une case déterminée. Le succès, pour un grain déterminé, est de tomber sur la case choisie, et il a une probabilité  $p = \frac{1}{64}$ . Il s'agit d'une variable binomiale avec n=64.

$$\text{Avec la loi binomiale : } P(X=h) = C_h^{64} p^h \cdot (1-p)^{64-h} = \frac{64!}{h!(64-h)!} \left(\frac{1}{64}\right)^h \cdot \left(\frac{63}{64}\right)^{64-h}$$

$$\text{avec la loi de Poisson } (\lambda = np = 1) : P(X=h) = \frac{e^{-1}}{h!}.$$

Le calcul numérique n'est plus qu'une affaire de machine ; en voici le résultat. Le tableau est limité à h=6, car, pour toute valeur supérieure, la probabilité calculée par l'une ou l'autre formule est inférieure à 0,000 4, si bien que l'approximation donnée, avec trois chiffres significatifs après la virgule, est nulle.

Valeur de h	binomiale	Poisson
0	0,365	0,368
1	0,371	0,368
2	0,185	0,184
3	0,061	0,061
4	0,015	0,015
5	0,003	0,003
6	0,001	0,001

## Espérance et variance de la loi de Poisson

Nous avons vu précédemment que  $\sum_{h=0}^{+\infty} \frac{\lambda^h}{h!} = e^\lambda$  ; donc  $\sum_{h=0}^{+\infty} e^{-\lambda} \frac{\lambda^h}{h!} = e^{-\lambda} \sum_{h=0}^{+\infty} \frac{\lambda^h}{h!} = e^{-\lambda} e^\lambda = e^0 = 1$  ;

ce qui est évident puisque cette somme est la somme de toutes les probabilités des diverses valeurs possibles d'une variable aléatoire. Cela permet un calcul très simple de l'espérance mathématique :

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

Ainsi l'espérance d'une loi de Poisson de paramètre  $\lambda$  est égale à  $\lambda$ . Ce résultat n'est surprenant puisque nos lois binomiales avaient toutes pour espérances  $np$ , et que la limite de  $np$  lorsque  $n$  tend vers l'infini est égale à  $\lambda$ .

Quant à la variance, elle se calcule par un procédé identique, qui la trouve sans surprise égale à la limite de  $npq$ , c'est-à-dire encore au paramètre  $\lambda$ .

### Exemple

Un certain vaccin provoque chez un individu sur 800 une réaction dangereuse.

Quelle probabilité y a-t-il, en vaccinant 3000 personnes qu'il y ait

- trois réactions dangereuses ?
- plus de deux réactions dangereuses ?

Soit  $Z$  la variable aléatoire indiquant le nombre total de réactions dangereuses.

On a une distribution binomiale avec  $p = \frac{1}{800}$ ,  $n = 3000$  et  $\lambda = np = 3,75$  ;

$$a) \quad P(Z=3) = C_3^{3000} p^3 \cdot (1-p)^{3000-3} = \frac{3000!}{3! 2997!} \left(\frac{1}{800}\right)^3 \cdot \left(\frac{799}{800}\right)^{2997} \cong \frac{3,75^3 e^{-3,75}}{3!} \cong 0,2067$$

(valeur exacte donnée par la loi binomiale est :  $P(Z=3) = 0,20678$ )

On voit que le calcul est grandement facilité par le recours à la loi de Poisson.

$$b) \quad P(Z>2) = 1 - P(Z \leq 2) = 1 - P(Z=0) - P(Z=1) - P(Z=2) \cong 0,7229$$

(valeur exacte donnée par la loi binomiale est :  $P(Z>2) = 0,7231$ ).

## 4.2 Lois continues

### 4.2.1 Loi exponentielle

Cette loi s'introduit très naturellement dans un certain nombre de problèmes d'attente. Imaginons un événement assez rare pour que la loi de Poisson lui soit applicable : l'exemple classique est celui des arrivées des clients à un guichet pendant de brefs intervalles de temps successifs, par exemple les tranches de 10 secondes.

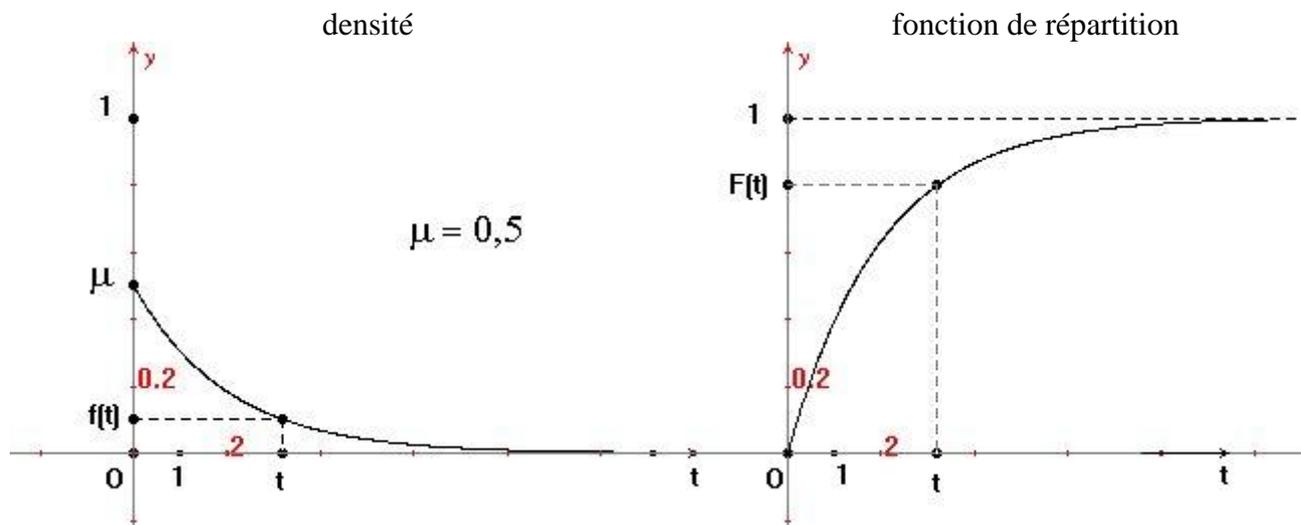
Notons  $\mu$  le nombre moyen d'arrivées par unité de temps, en sorte que  $\mu t$  soit le nombre moyen d'arrivées dans un intervalle de temps de durée  $t$ . La probabilité pour que le nombre de clients arrivés

dans cet intervalle soit  $x$  est :  $P(x) = \frac{\lambda^x}{x!} e^{-\lambda}$  où le paramètre  $\lambda$  est précisément  $\mu t$ .

Cette formule donne en particulier la probabilité pour qu'aucun client ne se présente dans cet intervalle de temps : il suffit d'y faire  $x = 0$  :  $P(0) = e^{-\lambda} = e^{-\mu t}$ .

Mais cette probabilité est donc celle pour que le temps d'attente du guichetier dépasse  $t$ . Si  $\tau$  est ce temps d'attente :  $P(\tau > t) = e^{-\mu t}$  et donc pour l'événement contraire :  $P(\tau \leq t) = 1 - e^{-\mu t}$ .

Ainsi nous avons obtenu la fonction de répartition du temps d'attente : cette loi de probabilité est connue sous le nom de loi exponentielle. Lorsque les arrivées d'un phénomène discontinu sont poissonniennes (c'est-à-dire qu'elles suivent la loi de Poisson), la durée d'attente entre deux arrivées successives suit une loi exponentielle. Le nombre  $\mu$  est le paramètre :



On voit l'intérêt d'un tel résultat : il permet, par exemple, de régler le nombre de guichets en sorte que la probabilité pour que la durée d'attente dépasse une valeur choisie soit suffisamment faible.

Le calcul de l'espérance mathématique et de la variance d'une loi exponentielle est un petit exercice de calcul intégral :

$$E(X) = \int_0^{+\infty} \mu t e^{-\mu t} dt \stackrel{\text{(en posant } x = \mu t)}{=} \frac{1}{\mu} \int_0^{+\infty} x e^{-x} dx \stackrel{\text{(par parties)}}{=} \frac{1}{\mu} [-xe^{-x} - e^{-x}]_0^{+\infty} = \frac{1}{\mu}$$

$$E(X^2) = \int_0^{+\infty} \mu t^2 e^{-\mu t} dt \stackrel{\text{(en posant } x = \mu t)}{=} \frac{1}{\mu^2} \int_0^{+\infty} x^2 e^{-x} dx \stackrel{\text{(par parties deux fois)}}{=} \frac{1}{\mu^2} [-x^2 e^{-x} - 2x e^{-x} - 2e^{-x}]_0^{+\infty} = \frac{2}{\mu^2}$$

$$V(X) = E(X^2) - E(X)^2 = \frac{2}{\mu^2} - \frac{1}{\mu^2} = \frac{1}{\mu^2}.$$

## 4.2.2 Loi normale de Laplace-Gauss

### Définition

D'autres fonctions ont été construites pour des approximations de la loi binomiale pour des grandes valeurs de  $n$ . Cette recherche est facilitée si l'on travaille pour chaque valeur de  $n$  sur la loi binomiale centrée réduite :

- centrée (c'est-à-dire, rappelons-le, que l'on a retranché à toutes les valeurs de la série leur espérance mathématique). Sur le graphique, cela évite que, au fur et à mesure que  $n$  devient plus grand, les diagrammes ne s'éloignent vers la droite, puisque l'espérance mathématique d'une loi binomiale est égale à  $np$  ;
- réduite (c'est-à-dire que l'on a divisé toutes les valeurs de la variable aléatoire par leur écart-type). Sur le graphique, cela évite que, au fur et à mesure que  $n$  devient plus grand, les diagrammes ne s'écrasent sur l'axe (OI), puisque la variance d'une loi binomiale est égale à  $np(1-p)$ .

### Exemple 1

Voici les diagrammes de la variable binomiale discrète  $Z = B(10, \frac{3}{5})$

$$\text{de loi } P(Z=z) = B(z; 10, \frac{3}{5}) = C_z^{10} \left(\frac{3}{5}\right)^z \cdot \left(\frac{2}{5}\right)^{10-z} = \frac{10!}{z!(10-z)!} \left(\frac{3}{5}\right)^z \cdot \left(\frac{2}{5}\right)^{10-z} = B(z)$$

$$\text{d'espérance } \mu = E(Z) = np = 6 \text{ et d'écart-type } \sigma = S(Z) = \sqrt{np(1-p)} = \sqrt{10 \cdot \frac{3}{5} \cdot \frac{2}{5}} = \sqrt{\frac{12}{5}} \cong 1,55$$

et de la variable centrée et réduite  $X = Z^* = \frac{Z-\mu}{\sigma}$  définie par

la loi  $f_{10}(x) = \sigma B(z; 10, \frac{3}{5})$  avec  $z = \sigma x + \mu$ . Alors que la variable  $Z$  peut prendre les valeurs  $z_i = i$

( $i = 0, 1, 2, \dots, n$ ) avec les probabilités  $B(z_i; 10, \frac{3}{5})$ , la variable (centrée et réduite)  $X = \frac{Z-\mu}{\sigma}$

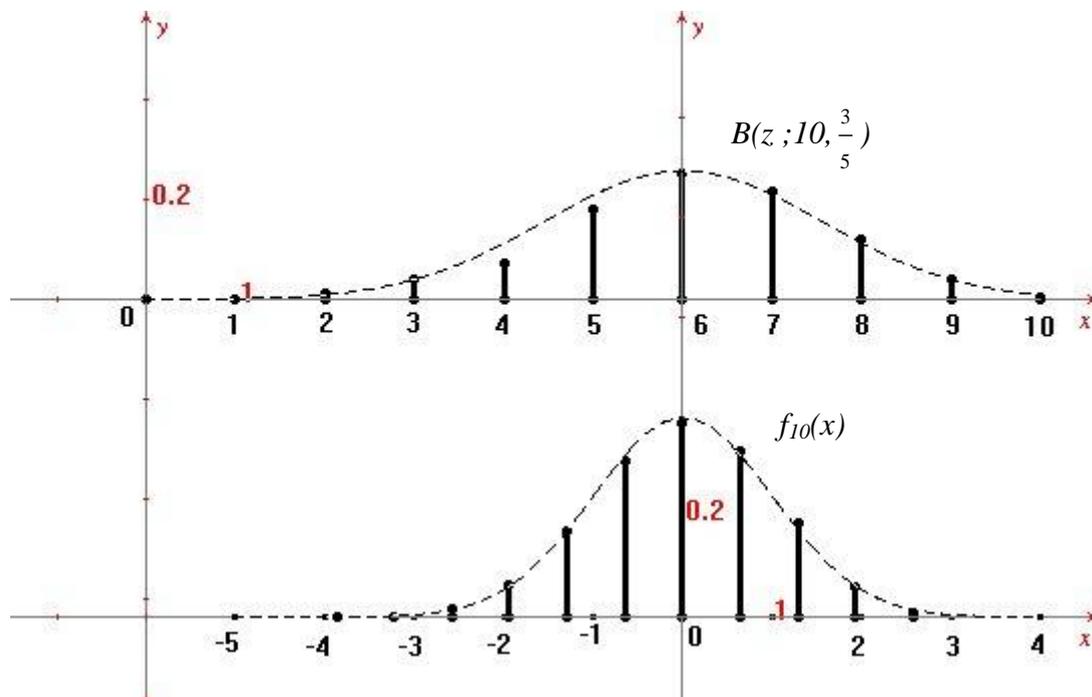
correspondante prendra les valeurs  $x_i = \frac{z_i - 6}{1,55}$  avec les probabilités respectives  $f_{10}(x_i)$ .

Voici le tableau des valeurs des deux lois de distributions des variables  $Z$  et  $X = \frac{Z-\mu}{\sigma}$  :

$z$	0	1	2	3	4	5	6	7	8	9	10
$B(z)$	0,000	0,002	0,011	0,042	0,111	0,201	0,251	0,215	0,121	0,040	0,006
$x$	-3,87	-3,23	-2,582	-1,936	-1,291	-0,645	0	0,645	1,291	1,936	2,582
$f_{10}(x)$	0,000	0,002	0,016	0,066	0,173	0,311	0,389	0,333	0,187	0,062	0,009

Sur les diagrammes ci-dessous, on distingue la distribution en cloche des  $B(z; 10, \frac{3}{5})$  avec une

espérance de 6 et un écart-type de 1,55 et celle des  $f_{10}(x)$  qui est "normalisée" avec une espérance de 0 et un écart-type de 1 :



## Exemple 2

Voici les diagrammes de la variable binomiale discrète  $Z = B(20, \frac{3}{5})$

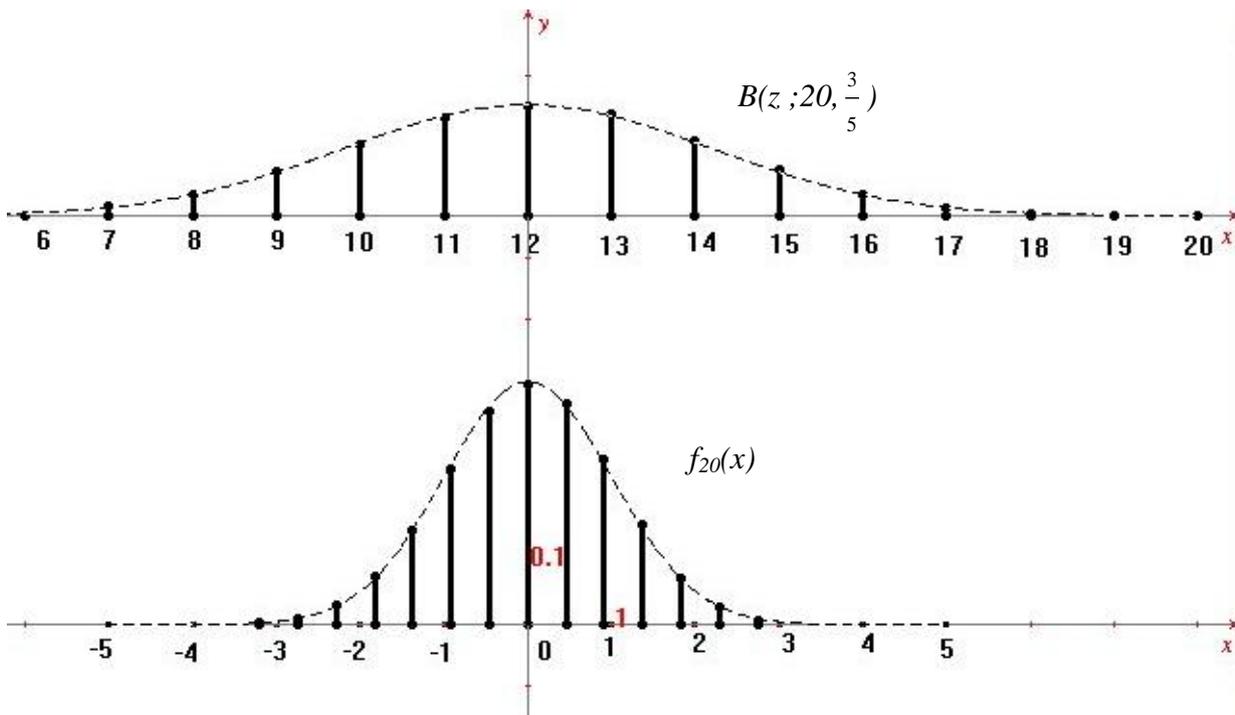
de loi  $P(Z=z) = B(z; 20, \frac{3}{5}) = B(z)$

d'espérance  $\mu = E(Z) = np = 12$  et d'écart-type  $\sigma = S(Z) = \sqrt{np(1-p)} \cong 2,19$

et de la variable centrée et réduite  $X = Z^* = \frac{Z-12}{2,19}$  définie par

la loi  $f_{20}(x) = 2,19 \cdot B(z; 20, \frac{3}{5})$  avec  $z = 2,19x + 12$  et  $x = \frac{z-12}{2,19}$

$z$	0	1	2	3	4	5	6	7	8	9	10
$B(z)$	0,000	0,000	0,000	0,000	0,000	0,001	0,005	0,015	0,035	0,071	0,117
$x$	-5,479	-5,023	-4,566	-4,110	-3,653	-3,196	-2,740	-2,283	-1,826	-1,370	-0,913
$f_{20}(x)$	0,000	0,000	0,000	0,000	0,000	0,003	0,011	0,032	0,078	0,155	0,256
$z$		11	12	13	14	15	16	17	18	19	20
$B(z)$		0,160	0,180	0,166	0,124	0,075	0,035	0,012	0,003	0,000	0,000
$x$		-0,457	0	0,457	0,913	1,370	1,826	2,283	2,740	3,196	3,653
$f_{20}(x)$		0,350	0,394	0,363	0,272	0,163	0,077	0,027	0,006	0,001	0,000



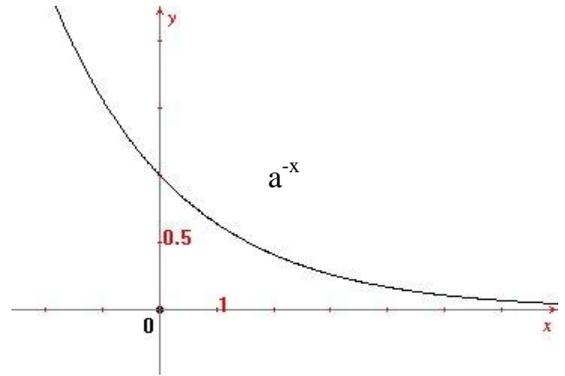
En comparant les graphiques de  $B(z; 10, \frac{3}{5})$  et de  $B(z; 20, \frac{3}{5})$ , on constate un étalement de la cloche

(en largeur,  $\sigma$  plus grand) et un déplacement vers la droite du sommet ( $\mu$  plus grand) mais aussi une diminution des ordonnées (probabilités plus petites car plus nombreuses).

Pour la variable  $X = Z^*$  de lois  $f_{10}(x)$  et  $f_{20}(x)$ , on constate par contre que la cloche a sensiblement les mêmes formes et disposition. Seuls les bâtonnets sont plus nombreux (plus serrés).

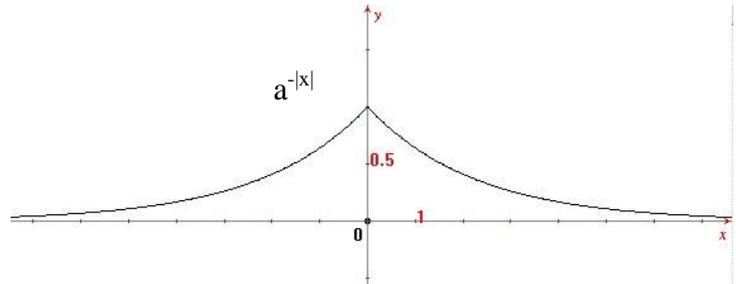
On pressent donc que la distribution  $X$  tend vers une limite lorsque  $n$  tend vers l'infini, c'est-à-dire qu'il existe une fonction  $g$  continue et dérivable qui est en quelque sorte l'enveloppe des bâtonnets lorsque  $n$  tend vers l'infini.

C'est Laplace qui, le premier, a cherché à ajuster une courbe à ces diagrammes ainsi réduits ; il a d'abord pensé à utiliser une fonction de type exponentiel d'équation  $y = a^{-x}$ , qui, pour les valeurs positives de  $x$ , présente un aspect comparable à nos diagrammes de lois binomiales :



Pour obtenir une loi symétrique, il faut associer la même valeur de  $y$  à deux valeurs opposées de  $x$ , autrement dit, utiliser une valeur absolue.

La fonction d'équation  $y = a^{-|x|}$  dans laquelle il ne reste plus qu'à déterminer  $y$ , par exemple en calculant l'écart-type, a été appelée "première fonction de Laplace" :



Laplace lui-même l'a rapidement abandonnée pour une fonction de la forme  $y = ba^{-x^2}$  qui fournit une approximation bien meilleure. Le théorème de **Laplace-De Moivre** démontre d'ailleurs par le calcul que cette formule fournit la meilleure approximation possible de la loi binomiale centrée réduite, et l'on sait calculer les valeurs  $a$  et  $b$ .

Enoncé du **théorème de Laplace-De Moivre** :

Soit  $F_n$  la fonction de répartition d'une v.a binomiale  $X$  centrée réduite, de paramètres  $n$  et  $p$ .

Pour tout nombre  $x$ , la suite des nombres  $F_n(x)$  tend vers  $g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  lorsque  $n$  tend vers l'infini ;

plus précisément, si  $a$  et  $b$  sont deux entiers, alors la probabilité  $P(a \leq X \leq b) = \sum_{z=a}^b B(z; n, p)$

est égale approximativement à  $\frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{x^2}{2}} dx = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

La formule obtenue :  $y = g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  donne la densité  $g$  de la loi limite ; celle-ci est d'un usage si

fréquent qu'on l'appelle **loi normale**, notée  $N(0,1)$ , car d'espérance 0 et d'écart-type 1.

On utilise aussi le terme "loi gaussienne" ou loi de Laplace-Gauss.

La formule ci-dessus est rarement utilisée en tant que telle ; elle est en effet simple d'apparence, mais les calculs numériques restent malaisés (il est impossible de calculer une primitive de cette fonction).

Il existe des tables numériques donnant des valeurs approchées de  $y$ .

Par un changement d'origine et d'échelle, on étend cette présentation au cas général, obtenant pour fonction de densité, notée  $N(\mu, \sigma)$ . Le paramètre  $\mu$  peut être quelconque mais  $\sigma$  est positif.

On obtient une loi de Laplace-Gauss  $N(\mu, \sigma)$  d'espérance  $\mu$  et d'écart-type  $\sigma$  en faisant subir au graphique de la loi normale  $N(0,1)$  les transformations successives suivantes :

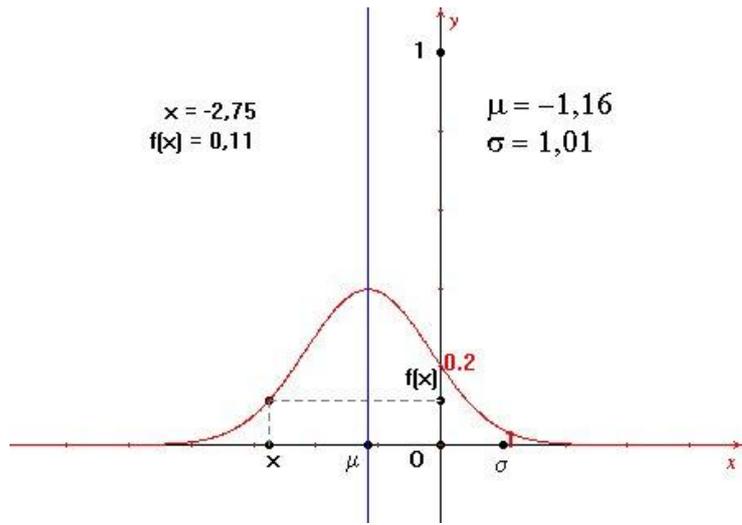
- 1) on multiplie les abscisses par  $\sigma$  (affinité) : on remplace  $x$  par  $\frac{t}{\sigma}$  ;
- 2) on ajoute  $\mu$  aux abscisses (translation) : on remplace  $t$  par  $z - \mu$  ;
- 3) on multiplie les ordonnées par  $\frac{1}{\sigma}$  (affinité) : la fonction  $g$  est multipliée par  $\frac{1}{\sigma}$ .

La densité de probabilité  $f(z)$  de cette loi  $N(\mu, \sigma)$  est ainsi donnée par  $f(z) = \frac{1}{\sigma} g(x)$  où  $x = \frac{t}{\sigma} = \frac{z - \mu}{\sigma}$ ,

on obtient :  $f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(z-\mu)^2}{\sigma^2}}$

Si l'on remplace  $z$  par  $x$ , cela donne :

$$f_{(\mu, \sigma)}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}}$$



On remarquera que par les transformations précédentes l'aire totale sous le graphique n'a pas varié.

Les transformations précédentes, effectuées en sens inverse, permettent de passer d'une variable

aléatoire  $Z$  suivant la loi normale  $N(\mu, \sigma)$  à la variable  $X = Z^* = \frac{Z - \mu}{\sigma}$  qui suit la loi normale  $N(0,1)$  :

$$P(Z \leq z_0) = P(\sigma X + \mu \leq z_0) = P(X \leq \frac{z_0 - \mu}{\sigma}).$$

En posant  $z_0^* = \frac{z_0 - \mu}{\sigma}$ , on peut écrire  $P(Z \leq z_0) = P(Z^* \leq z_0^*) = \Phi(z_0^*)$  où  $\Phi$  est la fonction de

répartition de la loi  $g$  :  $P(X \leq z_0^*) = \int_{-\infty}^{z_0^*} g(x) dx = \Phi(z_0^*)$ .

*Exemple :* Calculons  $P(Z \leq 13)$  pour une variable aléatoire  $Z$  suivant une loi  $N(10,2)$  :

On a  $Z^* = \frac{Z-10}{2}$  et si  $Z = 13$  alors  $Z^* = \frac{13-10}{2} = 1,5$

d'où  $P(Z \leq 13) = P(Z^* \leq 1,5) = \Phi(1,5) = 0,9332$ .

C'est l'une des lois les plus importantes. Une telle loi intervient dans des expériences aléatoires possédant de nombreuses causes indépendantes dont les effets s'ajoutent, sans que l'un d'eux soit dominant.

## Remarques

- La densité est maximale au point  $x = \mu$ . Abusivement, on a envie de dire que la valeur la plus probable d'une variable aléatoire normale  $N(\mu, \sigma)$  est égale à l'espérance mathématique ; c'est bien sûr un abus, puisque la probabilité d'une valeur déterminée est nulle. Ce qu'on peut dire, c'est que de tous les petits intervalles de même longueur, c'est celui qui est centré sur l'espérance mathématique qui a la plus grande probabilité.
- La courbe est symétrique par rapport à la droite verticale  $x = \mu$  ( $x = 0$  pour la loi normale centrée réduite  $N(0,1)$ ). Cela a une conséquence importante pour une table de valeurs numériques de la densité : il suffit de donner une table pour les valeurs positives de  $x$ .
- Dès que l'on s'éloigne de 0, les valeurs de la densité deviennent très petites.  
Le tableau ci-après donne quelques valeurs :

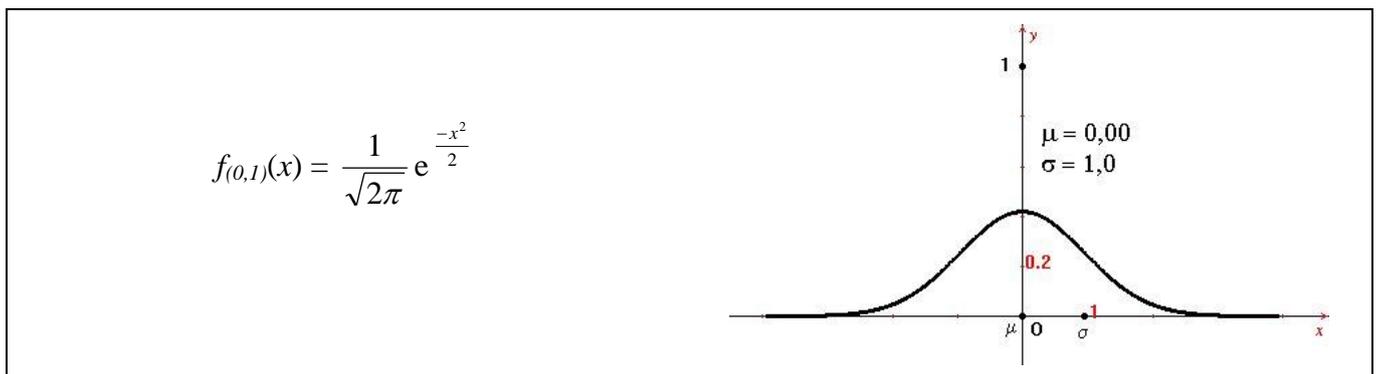
x	0	1	2	3	4
densité f(x)	0,3899	0,2420	0,0540	0,0044	0,0001

Une représentation graphique courante aura donc une graduation en abscisses qui ne va guère au-delà de l'intervalle  $[-3,3]$ .

## Propriétés

Si  $X$  suit la loi normale  $N(\mu, \sigma)$  alors  $E(X) = \mu$ ,  $V(X) = \sigma^2$ ,  $S(X) = \sigma$  et la variable aléatoire  $X^* = \frac{X - \mu}{\sigma}$  suit la loi normale centrée réduite  $N(0,1)$ .

On peut donc ramener l'étude de n'importe quelle loi normale à celle-ci :



On a alors  $P(X \leq u) = \int_{-\infty}^u f(x) dx = \Phi(u)$

avec  $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  qui correspond à l'aire

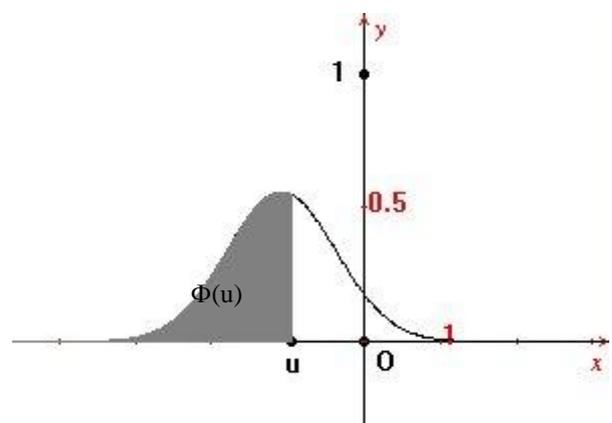
suivante :

Les valeurs usuelles de  $\Phi(u)$  (pour  $u \geq 0$ ) sont données dans la table du formulaire (page 139) et on a

$$P(X \leq -u) = \Phi(-u) = 1 - \Phi(u),$$

$$P(u_1 < X \leq u_2) = \Phi(u_2) - \Phi(u_1) \text{ et}$$

$$P(-u < X \leq u) = 2\Phi(u) - 1 .$$



Quelques exemples de valeurs pour la loi  $N(0,1)$  :

$$P(|X| < 1) = 0,6827$$

$$P(|X| < 2) = 0,9545$$

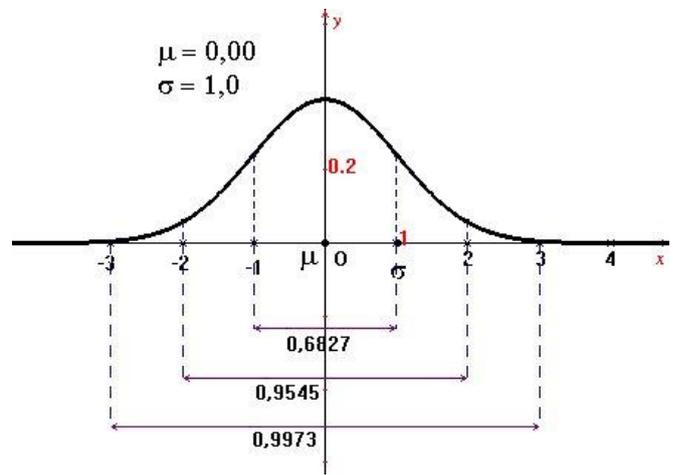
$$P(|X| < 3) = 0,9973$$

Ou, en revenant au cas général de la loi  $N(\mu, \sigma)$  :

$$P(|X - \mu| < \sigma) = 0,6827$$

$$P(|X - \mu| < 2\sigma) = 0,9545$$

$$P(|X - \mu| < 3\sigma) = 0,9973$$



## Exemple

Une entreprise de transport a un parc de 150 camions. On désigne par  $X$  la variable aléatoire qui à chaque camion associe la distance qu'il a parcourue dans la journée (en km).  $X$  suit la loi normale  $N(120, 14)$ .

On considère alors la variable  $X^* = \frac{X - 120}{14}$  qui suit la loi normale  $N(0,1)$ .

Si on veut la probabilité qu'un camion choisi au hasard parcoure dans la journée entre 110 et 130 km on procède de la manière suivante :

$$110 \leq X \leq 130 \Leftrightarrow -10 \leq X - 120 \leq 10 \Leftrightarrow -\frac{10}{14} \leq \frac{X - 120}{14} \leq \frac{10}{14} \Leftrightarrow -\frac{5}{7} \leq Y \leq \frac{5}{7}. \text{ Or, } \frac{5}{7} \approx 0,71.$$

$$\text{Donc } P(110 \leq X \leq 130) = P\left(-\frac{5}{7} \leq X^* \leq \frac{5}{7}\right) = 2\Phi\left(\frac{5}{7}\right) - 1 \approx 2\Phi(0,71) - 1.$$

Les deux premiers chiffres (0,7) nous donne la ligne de la table alors que le troisième (1) nous indique la colonne.

On obtient donc  $P(110 \leq X \leq 130) \approx 2 \times 0,7611 - 1 = 0,5222$ .

## Espérance et variance

$$\text{L'espérance est bien donnée par } E(X) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} dx = \left[ \frac{-1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} = 0 ;$$

$$\text{La variance sera : } \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx \stackrel{(\text{par parties})}{=} \left[ \frac{-1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0 + 1 = 1,$$

d'où l'écart-type est bien  $S(X) = 1$ .

## Approximation de la distribution binomiale par la loi normale N(0,1) et par la loi de Poisson

Dans la pratique, on approche la loi binomiale soit par la loi normal N(0,1), soit par la loi de Poisson.

Voici à titre de comparaison les probabilités fournies par les lois binomiale (exacte), de Poisson (approchée) et normale (approchée) dans le cas d'un événement de probabilité  $p = 0,002$  répété  $n = 1000$  fois. On a  $E(X) = np = 2$  et  $S(X) = \sigma = \sqrt{np(1-p)} \cong 1,428$ .

$z$	Binomiale	Poisson	Normale
0	0,13506	0,13534	0,14369
1	0,27067	0,27067	0,2175
2	0,27094	0,27067	0,25055
3	0,18062	0,18045	0,14767
5	0,03602	0,03609	0,06482
8	0,00085	0,00086	0,00194

On voit que la concordance entre les lois binomiale et de Poisson est excellente alors que la loi normale donne ici de mauvaises approximations du fait que  $np$  est inférieur à 5.

L'approximation d'une loi binomiale par la loi normale est correcte au centième près dès que  $np$  et  $n(1-p)$  dépasse 5.

**En résumé :** La règle d'approximation pour déterminer une probabilité dans le cas d'une loi binomiale par la loi normale centrée réduite (cf. tables page 139) est la suivante :

*Données :*

- une loi binomiale de paramètres  $n$  et  $p$  ;
- une valeur  $x$  dont on recherche la probabilité.

*Calculs :*

- espérance mathématique  $E(X) = np$  ;
- écart-type  $S(X) = \sigma = \sqrt{np(1-p)}$  ;
- écart réduit  $z$  correspondant à la valeur de  $x$  étudiée :

$$\begin{array}{ll} \text{différence} & x - np \\ \text{écart} & |x - np| \\ \text{écart réduit} & z = \frac{|x - np|}{\sigma} \end{array}$$

- consultation de la table pour une approximation de  $\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$
- division par  $\sigma$  de l'approximation obtenue.

### Applications numériques :

Cherchons une approximation de la probabilité  $B(x ; n, p)$  de la valeur  $x=84$  dans un loi binomiale  $B(n, p)$  de paramètres  $n = 400$  et  $p = 0,2$  :

- espérance mathématique :  $E(X) = np = 400 \times 0,2 = 80$
- écart-type  $\sigma = \sqrt{np(1-p)} = 8$
- écart réduit correspondant à la valeur de  $x$  observée :  $\frac{84-80}{8} = 0,5$
- lecture de la table : 0,3521

Nous adoptons donc pour valeur approchée de la probabilité  $B(x ; n, p)$  le nombre  $\frac{1}{8} \times 0,3521 \cong 0,043$ .

On peut comparer cette approximation avec la valeur " exacte " de

$$B(84 ; 400 ; 0,2) = C_{84}^{400} (0,2)^{84} \cdot (0,8)^{316} = \frac{400!}{84! (316)!} (0,2)^{84} \cdot (0,8)^{316} \cong 0,04323.$$

### Exemple :

*Données :*

- une loi binomiale de paramètre  $n$  et  $p$  ;
- deux nombres  $a$  et  $b$  ; on recherche la somme des probabilités des valeurs comprises entre  $a$  et  $b$ .

*Calculs :*

- espérance mathématique de la loi :  $\mu = np$  ;
- écart-type  $\sigma = \sqrt{np(1-p)}$  ;
- valeurs centrées réduites correspondant à chacun des nombres  $a$  et  $b$  :

$$z_a = \frac{a - \mu - \frac{1}{2}}{\sigma} \quad \text{et} \quad z_b = \frac{b - \mu + \frac{1}{2}}{\sigma}$$

- consultation de la table : si  $z$  est positif, lecture directe  $\Phi(z)$   
si  $z$  est négatif,  $\Phi(z) = 1 - \Phi(-z)$ .

*Exemple :*

On suppose que la durée de vie des ampoules de projecteurs fabriquées par une machine automatique est une variable aléatoire binomiale  $X$  d'espérance mathématique 50 heures, et d'écart-type 2 heures ; quelle est la probabilité qu'une ampoule ait une durée de vie qui dépasse 54 heures ?

Quelle est la probabilité qu'une ampoule ait une durée de vie inférieure à 47 heures ?

Calculons les durées centrées réduites :  $z_{54} = \frac{54-50}{2} = 2$  et  $z_{47} = \frac{47-50}{2} = -1,5$  ;

on lit dans la table  $\Phi(2) = 0,9772$  et donc la probabilité recherchée est

$$P(X > 54) = 1 - \Phi(2) = 0,0228$$

et de même

$$\Phi(-1,5) = 1 - \Phi(1,5) = 1 - 0,9332 = 0,0668 = P(X < 47).$$

# Séries de TD

## Série1

**Exercice N° 01 :** On veut étudier l'influence de l'ordre de prise de trois médicaments sur l'efficacité d'un traitement constitué par ces trois produits. De combien de façons peut-on organiser ce traitement ?

**Exercice N° 02 :** Une télévision privée décide d'opter pour le système de « programmes à péage » en utilisant des décodeurs commandés par des codes à huit chiffres.

Donner le nombre d'abonnés potentiels puis le nombre d'abonnés avec codes composés de huit chiffres différents.

**1-** Calculer le nombre de codes à deux chiffres différents, l'un étant utilisé une fois et l'autre sept fois.

**2-** Même question avec trois chiffres différents, dont deux sont utilisés une fois et le troisième six fois.

**Exercice N° 03 :** On dispose de cinq antibiotiques efficaces pour traiter une maladie infectieuse. On a vérifié au laboratoire que les cinq produits sont également actifs in vitro sur le microbe, mais on ne peut pas donner plus de deux antibiotiques à la fois. Combien y a-t-il de traitements possibles en associant deux antibiotiques ?

**Exercice N° 04 :** Neuf personnes se présentent à la médecine du travail pour passer la visite annuelle. Deux médecins les reçoivent. Le premier verra cinq personnes, le second quatre.

**1-** De combien de façon différentes les neufs personnes peuvent-elles être réparties entre chaque médecin ?

**2-** Il y a quatre personnes portant des lunettes. De combien de façons différentes peut-on réaliser cette répartition, sachant que chaque médecin verra deux personnes portant des lunettes ?

**3-** De plus, on veut que Mr. Durand qui porte des lunettes et Mr. Dupond qui n'en porte pas, soient examinés par le même médecin. Combien de répartitions sont possibles ?

**Exercice N° 05** Combien de nombre de cinq chiffres différents peut-on former avec les chiffres 1,2,...,9 si l'on veut que : **1-** les nombres formés soient pairs ; **2-** les deux premiers chiffres du nombres soient pairs.

**Exercice N° 06** De combien de façons différentes peut-on poser sur une même rangée six livres différents ?

**Exercice N° 07 :** On aligne sur une même étagère 5 livres tome 1, 4 livres tome 2 et 3 livres tome 3 d'un auteur. Combien d'arrangements différents y a-t-il ?

**Exercice N° 08 :** De combien de façons peut-on sélectionner 2 hommes, 4 femmes, 4 garçons et 5 filles parmi 6 hommes, 8 femmes, 4 garçons et 5 filles si : - aucune restriction n'est imposée, -un certain homme et une certaine femme doivent obligatoirement être choisis.

**Exercice N° 09 :** Une classe se compose de 6 filles et de 10 garçons. On forme un conseil de classe en prenant au hasard trois élèves.

**1-** Combien de possibilité a-t-on pour choisir : **a-** 3 garçons, **b-** 2 garçon, **c** – au moins un garçon, exactement une fille ?

**2-** Trois des filles ont les yeux bleus. En prenant 2 filles au hasard, combien de possibilités a-t-on pour que : **a-** toutes les deux aient des yeux bleus, **b-** aucune n'ait les yeux bleus, **c-** au moins une ait les yeux bleus ?

## Série 2

### Exercice N° 01 :

On sait que la probabilité pour qu'une personne soit allergique à un certain médicament est égale à  $10^{-3}$ . On s'intéresse à un échantillon de 1000 personnes. On appelle  $X$  la variable aléatoire dont la valeur est le nombre de personnes allergiques dans l'échantillon.

1 - Déterminer, en la justifiant, la loi de probabilité de  $X$ .

2 - En utilisant une approximation que l'on justifiera, calculer les probabilités des événements suivants :

a) Il y a exactement deux personnes allergiques dans l'échantillon.

b) Il y a au moins deux personnes allergiques dans l'échantillon.

### Exercice N° 02 :

On considère qu'il apparaît, en moyenne, chaque année 20 nouveaux cas de poliomyélite pour 100 000 habitants. On appelle  $X$  le nombre de nouveaux cas de poliomyélite dans l'année, dans une ville de 50 000 habitants (respectivement 5 000 habitants).

1°/ Montrer que la variable  $X$  suit une loi binomiale.

2°/ Montrer que l'on peut faire ici l'approximation de la loi binomiale par une loi de Poisson.

3°/ Calculer la probabilité pour qu'il apparaisse au plus quatre nouveaux cas de poliomyélite :

a) dans une ville de 50 000 habitants.

b) dans une ville de 5 000 habitants.

### Exercice N° 03 :

Un agriculteur a entreposé dans un local humide 12 doses d'un herbicide total et 8 doses d'un fongicide. Après plusieurs mois de séjour, les étiquettes sont indifférenciables. Chaque dose a la même probabilité d'être tirée. En vue d'un traitement, l'agriculteur prend 6 doses au hasard. Soit  $X$  la variable aléatoire égale au nombre de doses d'herbicide prises parmi ces 6 doses.

Déterminer la loi de probabilité de  $X$ . Donner la moyenne et la variance de cette distribution.

### Exercice N° 04 :

Le nombre de particules visibles dans des flacons d'une solution injectable suit une loi de Poisson. Sachant que sur 10 000 flacons contrôlés, on a décelé 2 particules dans 109 flacons et 3 particules dans 6 flacons, estimer :

Le nombre total de particules qui doivent avoir été décelées sur l'ensemble des 10000 flacons.

Le nombre de flacons sans particule visible.

### Exercice N° 05 :

Un bureau de réservations reçoit, entre 10h et 12h, en moyenne 1.2 appels téléphoniques par minute. Déterminer :

La probabilité pour qu'entre 11h et 11h01 on ait : aucun appel, un appel, deux appels.

La probabilité de recevoir 4 appels entre 11h et 11h02.

## Série 3

**Exercice N° 01 :** Une urne contient deux boules blanches et une boule noire.

On tire deux boules successivement avec remise et on appelle X le nombre de boules blanches. Donner la loi de probabilité de X.

On tire deux boules successivement sans remise et on appelle X le nombre de boules blanches. Donner la loi de probabilité de X.

On tire deux boules simultanément et on appelle X le nombre de boules blanches. Donner la loi de probabilité de X.

**Exercice N°02 :** Une machine M est constituée de deux éléments A et B. La défectuosité d'un seul élément suffit à mettre hors service la machine et on exclut toute autre éventualité de panne. Les avaries éventuelles relatives aux éléments A et B sont deux événements indépendants de probabilité respective  $a=0.1$  et  $b=0.2$ .

- 1- Calculer la probabilité pour que les deux éléments soient hors service en même temps.
- 2- Calculer la probabilité pour que la machine soit hors service.
- 3- On appelle X la v. a relative au nombre d'éléments hors service. Déterminer la loi de probabilité de la v. a X.
- 4- Calculer  $E(X)$  et  $\text{var}(X)$ .

**Exercice N° 03 :** On place un hamster dans une cage. Il se trouve face à 5 portillons dont un seul lui permet de sortir de la cage. A chaque essai infructueux, il reçoit une décharge électrique et on le replace à l'endroit initial.

1- En supposant que le hamster ne soit pas doué d'apprentissage et qu'il choisisse don de façon équiprobable entre les 5 solutions à chaque nouvel essai, déterminer la probabilité des événements : a) le hamster sort au premier essai, b) le hamster sort au troisième essai, c) le hamster sort au septième essai.

2- Le hamster mémorise maintenant les essais infructueux et choisi de façon équiprobable entre les portillons qu'il n'a pas encore essayé. On désigne par X la v. a. égale au nombre d'essais effectués.

a) Quelles valeurs peut prendre X ? Déterminer sa loi de probabilité.

b) Déterminer et interpréter l'espérance mathématique  $E(X)$ . Donner la variance.

**Exercice N° 04 :** Un certain jeu dépend du nombre de points X obtenus en jetant un dé équilibré.

1- Calculer la loi de probabilité de X. Calculer l'espérance mathématique et la variance.

2- Le gain de ce jeu est donné par la fonction linéaire  $G=2X+8$ .

a) Tabuler la distribution de G.

b) Calculer le gain espéré  $E(G)$  et sa variance.

**Exercice N° 05 :** Déterminer la loi de probabilité, le mode, l'espérance et l'écart-type de la v.a. donnée par la fonction de répartition :

$$F_x(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1/5 & \text{si } 1 \leq x < 2 \\ 4/5 & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x \geq 3 \end{cases} .$$

**Problème à domicile :** Chez une espèce végétale, la longueur des entre-noeuds séparant deux feuilles successives ayant terminé leur croissance est gouvernée par un couple d'allèles Aa.

Les longueurs des entre-noeuds des génotypes AA, Aa, aa sont respectivement égales à 3cm, 5cm et 2cm.

On considère une population composée des trois types d'individus avec les fréquences  $(1-q)^2$  pour le génotype AA,  $2q(1-q)$  pour le génotype Aa et  $q^2$  pour le génotype aa.

(a) i Un individu est tiré au hasard dans la population. Quelles sont les probabilités :

- de tirer un individu de génotype AA;
- de tirer un individu de génotype Aa ;
- de tirer un individu de génotype aa ?

ii Calculer, en fonction de  $q$ , l'espérance mathématique  $\mu_1$  de la variable aléatoire  $X_1$  égale à la longueur des entre-noeuds de l'individu tiré au hasard dans la population et étudier les variations de  $\mu_1$ , lorsque  $q \in [0,1]$ .

iii Pour quelle valeur de  $q$ , l'espérance mathématique de la longueur des entre-noeuds est-elle maximum ? Quelles sont alors les probabilités de tirer :

- un individu de génotype AA;
- un individu de génotype Aa ;
- un individu de génotype aa ?

iv Calculer, en fonction de  $q$ , les probabilités pour qu'un gamète issu d'un individu tiré dans la population :

- porte l'allèle A;
- porte l'allèle a.

(b) Un sélectionneur décide d'éliminer de la reproduction tous les individus dont la longueur des entre-noeuds est égale à 2 cm.

i. Calculer, en fonction de  $q$ , pour un individu tiré au hasard dans la population, la probabilité:

- d'être de génotype AA;
- d'être de génotype Aa ?

ii Calculer, en fonction de  $q$ , l'espérance mathématique de la variable aléatoire  $X_2$  égale à la longueur des entre-noeuds d'un individu tiré dans la population conservée pour la reproduction et étudier les variations de  $\mu_2$ , lorsque  $0 \leq q < 1$ .

iii. Calculer, en fonction de  $q$ , les probabilités pour qu'un gamète issu d'un individu tiré dans la population conservée pour la reproduction :

- porte l'allèle A;
- porte l'allèle a.

Comparer ces deux probabilités à celles de la question précédente.

**N.B. 1** On supposera toujours que tous les individus de la population ont la même probabilité d'être tirés lors du tirage au hasard.

**N.B. 2** On rappelle qu'un individu de génotype AA produit des gamètes portant l'allèle A avec la probabilité 1, qu'un individu de génotype aa produit des gamètes portant l'allèle a avec la probabilité 1 et qu'un individu de génotype Aa produit des gamètes portant l'allèle A avec la probabilité  $1/2$ , et portant l'allèle a avec la probabilité  $1/2$ .

## Série 4

### Exercice N° 01 :

La température  $T$  mesurée par jour dans une ville a pour moyenne de  $12^{\circ}\text{C}$  avec un écart-type de 10. On suppose que  $T$  suit une loi normale.

- 1- Quelle est la probabilité pour qu'en un jour donné la température soit positive ?
- 2- Quelle est la probabilité pour qu'en un jour donné la température soit inférieure à  $-5^{\circ}\text{C}$  ?
- 3- Quelle est la probabilité pour qu'en un jour donné la température soit comprise entre  $-5^{\circ}\text{C}$  et  $+5^{\circ}\text{C}$  ?

### Exercice N° 02 :

Dans une population masculine, la taille est une v. a.  $X \sim N(172,3)\text{cm}$ .

Dans une population féminine comparable, la taille est une v. a.  $Y \sim N(166,6)\text{cm}$ .

- 1- Y a-t-il plus d'hommes ou de femmes qui mesurent plus de 184cm ?
- 2- Quelle est la probabilité qu'une femme mesure plus de 184 cm, sachant qu'elle mesure plus de 180 cm ?

### Exercice N° 03 :

Pour un échantillon de 300 individus sains, on a étudié la glycémie ; on a constaté que 20% des glycémies sont inférieures à 0.82 g/l et que 30% des glycémies sont supérieures à 0.98 g/l. Dans ces conditions, si l'on suppose que la glycémie suit une loi normale, déterminer la moyenne et l'écart-type de cette loi.

### Exercice N° 04 :

A l'entrée d'une station de métro, un marchand de journaux remarque qu'en moyenne, entre 8h et 9h, une personne sur 10 achète un journal.

- 1- Sachant qu'il passe 400 personnes entre 8h et 9h, indiquer la loi de probabilité de  $X$ , nombre de journaux vendus pendant cette période (préciser les hypothèses).
- 2- Donner l'espérance mathématique et la variance de  $X$ .
- 3- Par quelle loi de probabilité peut-on approcher la loi de  $X$  ? Utiliser cette approximation pour calculer les probabilités des événements:  $X = 42$  ;  $X \geq 45$  ;  $35 \leq X < 45$ .

### Exercice N°05 :

Pour un certain type de graines, la probabilité de germination est  $p=0.8$ . Une personne sème 400 graines. Calculer la probabilité que 300 au moins germent.

### Exercice N°06 :

Dans une population homogène de 20 000 habitants, la probabilité pour qu'une personne quelconque demande à être vaccinée contre la grippe est de 0.4.

De combien de vaccins doit-on disposer pour que la probabilité qu'en vienne à en manquer soit inférieure à 0.1 ?

## Série 05

**Exercice N° 01 :** Soit  $X$  une v. a. continue, de fonction de répartition :

$$F_X(x) = \begin{cases} 1 - e^{-x/2} \left(1 + \frac{x}{2}\right) & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}.$$

- 1-  $F_X$  est-elle continue ? déterminer et interpréter  $\lim_{x \rightarrow +\infty} F_X(x)$ .
- 2- Déterminer la densité de probabilité de  $X$ . Calculer  $E(X)$  et  $\sigma_X$ .
- 3- Déterminer de ce qui précède les variations et le graphe de  $F_X$ .
- 4- Calculer  $P(1 \leq X < 2)$ .

**Exercice N° 02 :** Soit  $X$  une v. a. continue, de densité de probabilité :

$$f(x) = \begin{cases} \lambda(1-x^2) & \text{si } x \in [-1, 1] \\ 0 & \text{ailleurs} \end{cases}$$

- 1- Calculer  $\lambda$  et construire le graphe de  $f$ .
- 2- Donner et construire le graphe de la fonction de répartition de  $X$ .
- 3- Calculer  $P(|X| \geq \frac{1}{2})$  et représenter cette probabilité sur les deux graphes précédents.

**\*\*Exercice N° 03 :** Soit  $X$  une v. a. continue, de densité de probabilité :

$$f(x) = \begin{cases} \frac{2}{\pi} \cos^2 x & \text{si } x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ 0 & \text{ailleurs} \end{cases}$$

Calculer  $E(X)$  et  $\sigma_X$ . Donner la fonction de répartition de  $X$ . Calculer  $P\left(X \geq \frac{4\pi}{10}\right)$ .

**\*\*Exercice N° 04 :** Soit  $X$  une v. a. continue, de densité de probabilité :

$$f(x) = \begin{cases} a\sqrt{4-x^2} & \text{si } x \in [0, 2] \\ 0 & \text{ailleurs} \end{cases}$$

- 1- Donner la valeur de  $a$ . Calculer  $E(X)$  et  $\sigma_X$ .
- 2- Donner et construire le graphe de la fonction de répartition de  $X$ .

# Chapitre 3

## LOI DU CHI-DEUX

Définition (voir [3], [5], [6], [11]):

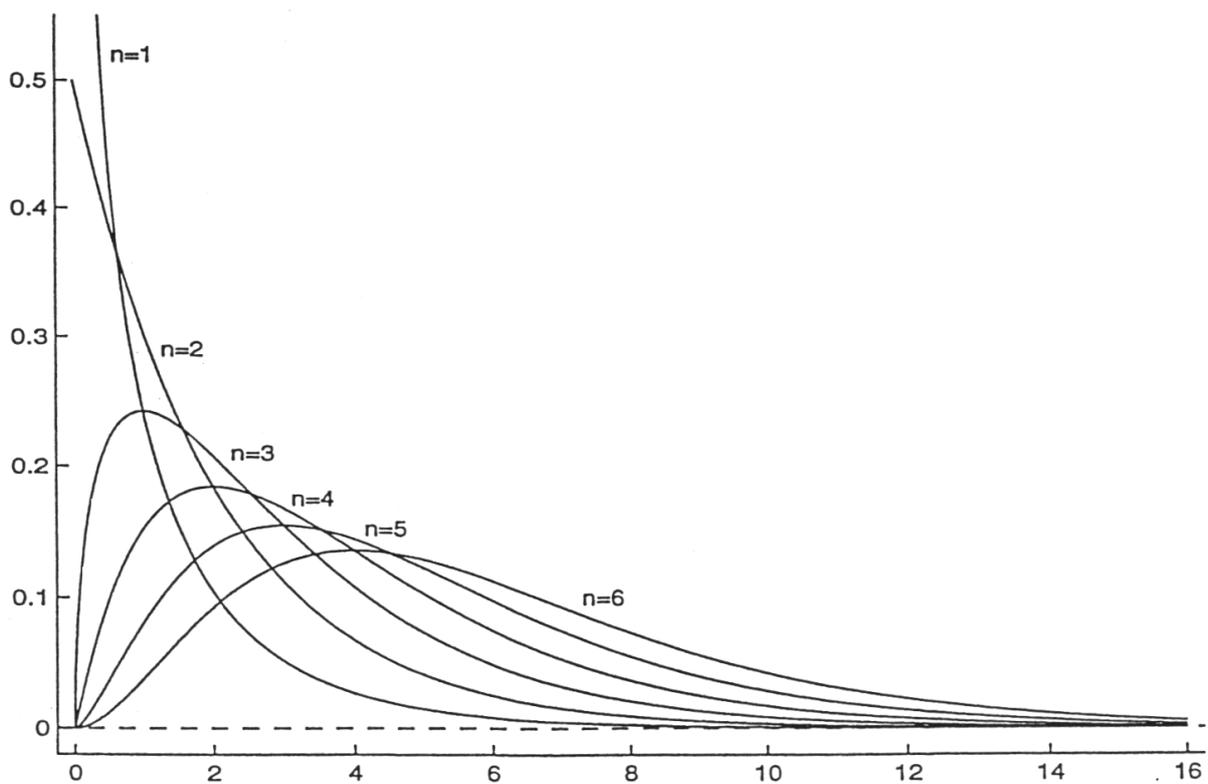
Soit  $(X_i)_{i=1, \dots, n}$  une suite de variables aléatoires indépendantes normales centrées réduites

(c'est à dire telles que  $L(X_i) = N(0,1)$ ). La somme des carrés des  $X_i$ ,  $S = \sum_{i=1}^n X_i^2$  est une variable aléatoire qui suit une loi du chi-deux à  $n$  degrés de liberté.

On note  $L(S) = X_n^2$

Propriétés : Si  $L(S) = X_n^2$  alors  $E(S) = n$  et  $Var(S) = 2n$

Si  $L(K_1) = X_{n_1}^2$  et  $L(K_2) = X_{n_2}^2$  et si  $K_1 \perp K_2$  alors  $L(K_1 + K_2) = X_{n_1+n_2}^2$



Fonctions densité de probabilité de la loi du  $X_n^2$  pour quelques valeurs de  $n$ .

Une situation où intervient la loi du chi-deux :

Soit  $(X_i)_{i=1, \dots, n}$  un  $n$ -échantillon tel que  $L(X_i) = N(\mu, \sigma)$  avec  $\mu$  et  $\sigma$  inconnus.

Dans ce cas  $\mu$  est estimé par :  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  et  $\sigma$  par  $S_c^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

On a alors :  $L\left(\frac{(n-1)S_c^2}{\sigma^2}\right) = X_{n-1}^2$  ou encore  $L\left(\frac{nS^2}{\sigma^2}\right) = X_n^2$  avec  $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$

## TEST D'ADEQUATION A UNE LOI

On considère une variable  $X$  à  $r$  modalités pour laquelle on a une série statistique d'effectifs  $(n_1, n_2, \dots, n_r)$ . On note  $n = n_1 + n_2 + \dots + n_r$  l'effectif total.

On se demande si ces observations correspondent à la loi de probabilité définie sur les mêmes  $r$  modalités par les probabilités  $(p_1, p_2, \dots, p_r)$ .

On veut tester  $H_0 : L(X) = (p_1, p_2, \dots, p_r)$  contre  $H_1 : L(X) \neq (p_1, p_2, \dots, p_r)$

On définit la statistique du  $X^2$  d'adéquation par :

$$kh = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i}$$

Les  $n_i$  sont les effectifs observés,  $n = n_1 + n_2 + \dots + n_r$  est l'effectif total et les  $np_i$  sont les effectifs « théoriques » de chaque modalité.

**Propriété :** Si l'hypothèse  $H_0$  est vraie alors la statistique  $kh$  est la réalisation lorsque  $n \rightarrow +\infty$  et  $n_i \rightarrow +\infty$  pour  $i = 1, \dots, r$  d'une variable aléatoire KH telle que  $L(KH) = X_{r-1}^2$ .

En pratique, nous utilisons cette loi asymptotique dès que  $n \geq 50$  et  $np_i \geq 5$  pour  $i = 1, \dots, r$ .

Si la condition  $np_i \geq 5$  n'est pas remplie on regroupera les modalités voisines pour qu'elle soit remplie.

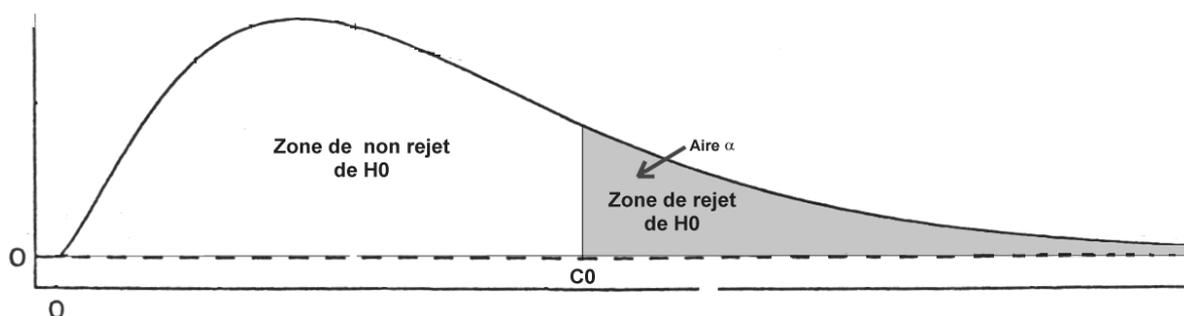
### Règle de décision :

Pour un seuil  $\alpha$  donné, les tables du  $X^2$  nous donnent une constante  $c_0$  telle que :

$$P(X_{r-1}^2 > c_0) = \alpha.$$

Si  $kh < c_0$  j'accepte l'hypothèse  $H_0$  au niveau  $1 - \alpha$

Si  $kh \geq c_0$  je rejette l'hypothèse  $H_0$  au seuil  $\alpha$ .



Exemple : souris des villes et souris des champs.

Dans une étude portant sur l'orientation spatiale chez les souris, les animaux de l'expérience ont été placés un par un au centre d'un labyrinthe radiaire comportant 8 allées orientées dans les huit directions de la rose des vents. Chaque animal s'est échappé par l'une de ces allées. Les expériences ont porté d'une part sur des souris de laboratoire et d'autre part sur des souris sauvages récemment capturées en un lieu situé au nord-est du laboratoire. Les répartitions des directions de fuite sont données dans le tableau ci-dessous.

Directions	Souris de laboratoire	Souris sauvages
N	17	26
NO	25	17
O	13	9
SO	28	2
S	19	3
SE	20	16
E	22	33
NE	16	54

On désire tester, dans le cas des souris de laboratoire d'une part puis dans le cas des souris sauvages l'hypothèse que le choix de la direction de fuite se fait au hasard.

L'effectif total (160) ainsi que les effectifs théoriques des 8 modalités (20) vérifient les conditions d'application du test d'adéquation du  $X^2$ .

Directions	Souris de laboratoire			Souris sauvages		
	$n_i$	$n_{pi}$	$(n_i - n_{pi})^2 / n_{pi}$	$n_i$	$n_{pi}$	$(n_i - n_{pi})^2 / n_{pi}$
N	17	20	0,45	26	20	1,80
NO	25	20	1,25	17	20	0,45
O	13	20	2,45	9	20	6,05
SO	28	20	3,20	2	20	16,20
S	19	20	0,05	3	20	14,45
SE	20	20	0,00	16	20	0,80
E	22	20	0,20	33	20	8,45
NE	16	20	0,80	54	20	57,80
	160		8,40	160		106,00

Choisissons le seuil de 0.05. Le nombre de degrés de liberté est  $8-1=7$ .

La lecture dans la table du  $X^2$  donne  $c_0 = 14.07$ .

Dans le cas des souris de laboratoire :  $kh \approx 8.4$  est inférieur à  $c_0 = 14.07$ . Les écarts entre les valeurs observés et les valeurs théoriques ne sont pas significatifs et on accepte l'hypothèse  $H_0$  au niveau  $1 - \alpha$ .

Dans le cas des souris sauvages :  $kh \approx 104$  est largement supérieur à  $c_0 = 14.07$ . les écarts entre les valeurs observés et les valeurs théoriques sont significatifs et on rejette l'hypothèse  $H_0$  au seuil  $\alpha$ .

**TABLE DE LA FONCTION DE REPARTITION DU  $\chi^2$**

Degré liberté	Probabilité <b>0,005</b>	<b>0,010</b>	<b>0,025</b>	<b>0,050</b>	<b>0,100</b>	<b>0,900</b>	<b>0,950</b>	<b>0,975</b>	<b>0,990</b>	<b>0,995</b>
<b>1</b>	0,00	0,00	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88
<b>2</b>	0,01	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60
<b>3</b>	0,07	0,11	0,22	0,35	0,58	6,25	7,81	9,35	11,34	12,84
<b>4</b>	0,21	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86
<b>5</b>	0,41	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75
<b>6</b>	0,68	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
<b>7</b>	0,99	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
<b>8</b>	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95
<b>9</b>	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
<b>10</b>	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
<b>11</b>	2,60	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73	26,76
<b>12</b>	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
<b>13</b>	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
<b>14</b>	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
<b>15</b>	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
<b>16</b>	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
<b>17</b>	5,70	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72
<b>18</b>	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
<b>19</b>	6,84	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58
<b>20</b>	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
<b>21</b>	8,03	8,90	10,28	11,59	13,24	29,62	32,67	35,48	38,93	41,40
<b>22</b>	8,64	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29	42,80
<b>23</b>	9,26	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64	44,18
<b>24</b>	9,89	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
<b>25</b>	10,52	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31	46,93
<b>26</b>	11,16	12,20	13,84	15,38	17,29	35,56	38,89	41,92	45,64	48,29
<b>27</b>	11,81	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96	49,65
<b>28</b>	12,46	13,56	15,31	16,93	18,94	37,92	41,34	44,46	48,28	50,99
<b>29</b>	13,12	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59	52,34
<b>30</b>	13,79	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	53,67
<b>31</b>	14,46	15,66	17,54	19,28	21,43	41,42	44,99	48,23	52,19	55,00
<b>32</b>	15,13	16,36	18,29	20,07	22,27	42,58	46,19	49,48	53,49	56,33
<b>33</b>	15,82	17,07	19,05	20,87	23,11	43,75	47,40	50,73	54,78	57,65
<b>34</b>	16,50	17,79	19,81	21,66	23,95	44,90	48,60	51,97	56,06	58,96
<b>35</b>	17,19	18,51	20,57	22,47	24,80	46,06	49,80	53,20	57,34	60,27
<b>36</b>	17,89	19,23	21,34	23,27	25,64	47,21	51,00	54,44	58,62	61,58
<b>37</b>	18,59	19,96	22,11	24,07	26,49	48,36	52,19	55,67	59,89	62,88
<b>38</b>	19,29	20,69	22,88	24,88	27,34	49,51	53,38	56,90	61,16	64,18
<b>39</b>	20,00	21,43	23,65	25,70	28,20	50,66	54,57	58,12	62,43	65,48
<b>40</b>	20,71	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69	66,77

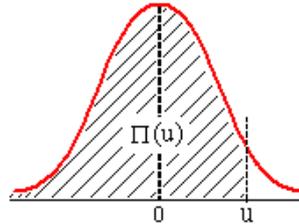
# Chapitre 4

## Tables

Voir la référence [4], [7], [9], [10]).

Table de Loi Normale

Fonction de répartition  $\Phi$  de la loi normale centrée réduite.



Probabilité de trouver une valeur inférieure à  $u$

$$\Phi(-u) = 1 - \Phi(u)$$

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169

2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520

2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997

### LOI DE POISSON ( $0.1 \leq \lambda \leq 1$ )

La table donne, pour  $k$  entier et  $\lambda$  réel,  $p_k = P(X = k)$  et  $\sum p_k = P(X \leq k)$ .

k	$\lambda = 0.1$	$\sum p_k$	$\lambda = 0.2$	$\sum p_k$	$\lambda = 0.3$	$\sum p_k$	$\lambda = 0.4$	$\sum p_k$	$\lambda = 0.5$	$\sum p_k$
	$p_k$		$p_k$		$p_k$		$p_k$		$p_k$	
0	0.9048	0.9048	0.8187	0.8187	0.7408	0.7408	0.6703	0.6703	0.6065	0.6065
1	0.0905	0.9953	0.1637	0.9825	0.2222	0.9631	0.2681	0.9384	0.3033	0.9098
2	0.0045	0.9998	0.0164	0.9989	0.0333	0.9964	0.0536	0.9921	0.0758	0.9856
3	0.0002	1	0.0011	0.9999	0.0033	0.9997	0.0072	0.9992	0.0126	0.9982
4	0	1	0.0001	1	0.0003	1	0.0007	0.9999	0.0016	0.9998
5			0	1	0	1	0.0001	1	0.0002	1
k	$\lambda = 0.6$	$\sum p_k$	$\lambda = 0.7$	$\sum p_k$	$\lambda = 0.8$	$\sum p_k$	$\lambda = 0.9$	$\sum p_k$	$\lambda = 1$	$\sum p_k$
	$p_k$		$p_k$		$p_k$		$p_k$		$p_k$	
0	0.5488	0.5488	0.4966	0.4966	0.4493	0.4493	0.4066	0.4066	0.3679	0.3679
1	0.3293	0.8781	0.3476	0.8442	0.3595	0.8088	0.3659	0.7725	0.3679	0.7358
2	0.0988	0.9769	0.1217	0.9659	0.1438	0.9526	0.1647	0.9371	0.1839	0.9197
3	0.0198	0.9966	0.0284	0.9942	0.0383	0.9909	0.0494	0.9865	0.0613	0.9810
4	0.0030	0.9996	0.0050	0.9992	0.0077	0.9986	0.0111	0.9977	0.0153	0.9963
5	0.0004	1	0.0007	0.9999	0.0012	0.9998	0.0020	0.9997	0.0031	0.9994
6	0	1	0.0001	1	0.0002	1	0.0003	1	0.0005	0.9999
7	0	1	0	1	0	1	0	1	0.0001	1

### LOI DE POISSON ( $2 \leq \lambda \leq 6$ )

La table donne, pour  $k$  entier et  $\lambda$  réel,  $p_k = P(X = k)$  et  $\sum p_k = P(X \leq k)$ .

k	$\lambda = 2$ $p_k$	$\sum p_k$	$\lambda = 3$ $p_k$	$\sum p_k$	$\lambda = 4$ $p_k$	$\sum p_k$	$\lambda = 5$ $p_k$	$\sum p_k$	$\lambda = 6$ $p_k$	$\sum p_k$
0	0.1353	0.1353	0.0498	0.0498	0.0183	0.0183	0.0067	0.0067	0.0025	0.0025
1	0.2707	0.4060	0.1494	0.1991	0.0733	0.0916	0.0337	0.0404	0.0149	0.0174
2	0.2707	0.6767	0.2240	0.4232	0.1465	0.2381	0.0842	0.1247	0.0446	0.0620
3	0.1804	0.8571	0.2240	0.6472	0.1954	0.4335	0.1404	0.2650	0.0892	0.1512
4	0.0902	0.9473	0.1680	0.8153	0.1954	0.6288	0.1755	0.4405	0.1339	0.2851
5	0.0361	0.9834	0.1008	0.9161	0.1563	0.7851	0.1755	0.6160	0.1606	0.4457
6	0.0120	0.9955	0.0504	0.9665	0.1042	0.8893	0.1462	0.7622	0.1606	0.6063
7	0.0034	0.9989	0.0216	0.9881	0.0595	0.9489	0.1044	0.8666	0.1377	0.7440
8	0.0009	0.9998	0.0081	0.9962	0.0298	0.9786	0.0653	0.9319	0.1033	0.8472
9	0.0002	1	0.0027	0.9989	0.0132	0.9919	0.0363	0.9682	0.0688	0.9161
10	0	1	0.0008	0.9997	0.0053	0.9972	0.0181	0.9863	0.0413	0.9574
11	0	1	0.0002	0.9999	0.0019	0.9991	0.0082	0.9945	0.0225	0.9799
12	0	1	0.0001	1	0.0006	0.9997	0.0034	0.9980	0.0113	0.9912
13	0	0	0	1	0.0002	0.9999	0.0013	0.9993	0.0052	0.9964
14	0	0	0	1	0.0001	1	0.0005	0.9998	0.0022	0.9986
15	0	0	0	0	0	1	0.0002	0.9999	0.0009	0.9995
16	0	0	0	0	0	1	0	1	0.0003	0.9998
17	0	0	0	0	0	1	0	1	0.0001	0.9999
18	0	0	0	0	0	0	0	1	0	1

### LOI DE POISSON ( $7 \leq \lambda \leq 11$ )

La table donne, pour  $k$  entier et  $\lambda$  réel,  $p_k = P(X = k)$  et  $\sum p_k = P(X \leq k)$ .

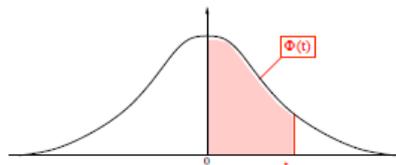
k	$\lambda = 7$	$\sum p_k$	$\lambda = 8$	$\sum p_k$	$\lambda = 9$	$\sum p_k$	$\lambda = 10$	$\sum p_k$	$\lambda = 11$	$\sum p_k$
	$p_k$		$p_k$		$p_k$		$p_k$		$p_k$	
0	0.0009	0.0009	0.0003	0.0003	0.0001	0.0001	0	0	0	0
1	0.0064	0.0073	0.0027	0.0030	0.0011	0.0012	0.0005	0.0005	0.0002	0.0002
2	0.0223	0.0296	0.0107	0.0138	0.0050	0.0062	0.0023	0.0028	0.0010	0.0012
3	0.0521	0.0818	0.0286	0.0424	0.0150	0.0212	0.0076	0.0103	0.0037	0.0049
4	0.0912	0.1730	0.0573	0.0996	0.0337	0.0550	0.0189	0.0293	0.0102	0.0151
5	0.1277	0.3007	0.0916	0.1912	0.0607	0.1157	0.0378	0.0671	0.0224	0.0375
6	0.1490	0.4497	0.1221	0.3134	0.0911	0.2068	0.0631	0.1301	0.0411	0.0786
7	0.1490	0.5987	0.1396	0.4530	0.1171	0.3239	0.0901	0.2202	0.0646	0.1432
8	0.1304	0.7291	0.1396	0.5925	0.1318	0.4557	0.1126	0.3328	0.0888	0.2320
9	0.1014	0.8305	0.1241	0.7166	0.1318	0.5874	0.1251	0.4579	0.1085	0.3405
10	0.0710	0.9015	0.0993	0.8159	0.1186	0.7060	0.1251	0.5830	0.1194	0.4599
11	0.0452	0.9467	0.0722	0.8881	0.0970	0.8030	0.1137	0.6968	0.1194	0.5793
12	0.0263	0.9730	0.0481	0.9362	0.0728	0.8758	0.0948	0.7916	0.1094	0.6887
13	0.0142	0.9872	0.0296	0.9658	0.0504	0.9261	0.0729	0.8645	0.0926	0.7813
14	0.0071	0.9943	0.0169	0.9827	0.0324	0.9585	0.0521	0.9165	0.0728	0.8540
15	0.0033	0.9976	0.0090	0.9918	0.0194	0.9780	0.0347	0.9513	0.0534	0.9074
16	0.0014	0.9990	0.0045	0.9963	0.0109	0.9889	0.0217	0.9730	0.0367	0.9441
17	0.0006	0.9996	0.0021	0.9984	0.0058	0.9947	0.0128	0.9857	0.0237	0.9678
18	0.0002	0.9999	0.0009	0.9993	0.0029	0.9976	0.0071	0.9928	0.0145	0.9823
19	0.0001	1	0.0004	0.9997	0.0014	0.9989	0.0037	0.9965	0.0084	0.9907
20	0	1	0.0002	0.9999	0.0006	0.9996	0.0019	0.9984	0.0046	0.9953
21	0	1	0.0001	1	0.0003	0.9998	0.0009	0.9993	0.0024	0.9977
22	0	1	0	1	0.0001	0.9999	0.0004	0.9997	0.0012	0.9990
23	0	1	0	1	0	1	0.0002	0.9999	0.0006	0.9995
24	0	1	0	1	0	1	0.0001	1	0.0003	0.9998
25	0	1	0	1	0	1	0	1	0.0001	0.9999
26	0	1	0	1	0	1	0	1	0	1

## LOI DE POISSON ( $12 \leq \lambda \leq 16$ )

La table donne, pour  $k$  entier et  $\lambda$  réel,  $p_k = P(X = k)$  et  $\sum p_k = P(X \leq k)$ .

k	$\lambda = 12$		$\lambda = 13$		$\lambda = 14$		$\lambda = 15$		$\lambda = 16$	
	$p_k$	$\sum p_k$								
0	0	0	0	0	0	0	0	0	0	0
1	0.0001	0.0001	0	0	0	0	0	0	0	0
2	0.0004	0.0005	0.0002	0.0002	0.0001	0.0001	0	0	0	0
3	0.0018	0.0023	0.0008	0.0011	0.0004	0.0005	0.0002	0.0002	0.0001	0.0001
4	0.0053	0.0076	0.0027	0.0037	0.0013	0.0018	0.0006	0.0009	0.0003	0.0004
5	0.0127	0.0203	0.0070	0.0107	0.0037	0.0055	0.0019	0.0028	0.0010	0.0014
6	0.0255	0.0458	0.0152	0.0259	0.0087	0.0142	0.0048	0.0076	0.0026	0.0040
7	0.0437	0.0895	0.0281	0.0540	0.0174	0.0316	0.0104	0.0180	0.0060	0.0100
8	0.0655	0.1550	0.0457	0.0998	0.0304	0.0621	0.0194	0.0374	0.0120	0.0220
9	0.0874	0.2424	0.0661	0.1658	0.0473	0.1094	0.0324	0.0699	0.0213	0.0433
10	0.1048	0.3472	0.0859	0.2517	0.0663	0.1757	0.0486	0.1185	0.0341	0.0774
11	0.1144	0.4616	0.1015	0.3532	0.0844	0.2600	0.0663	0.1848	0.0496	0.1270
12	0.1144	0.5760	0.1099	0.4631	0.0984	0.3585	0.0829	0.2676	0.0661	0.1931
13	0.1056	0.6815	0.1099	0.5730	0.1060	0.4644	0.0956	0.3632	0.0814	0.2745
14	0.0905	0.7720	0.1021	0.6751	0.1060	0.5704	0.1024	0.4657	0.0930	0.3675
15	0.0724	0.8444	0.0885	0.7636	0.0989	0.6694	0.1024	0.5681	0.0992	0.4667
16	0.0543	0.8987	0.0719	0.8355	0.0866	0.7559	0.0960	0.6641	0.0992	0.5660
17	0.0383	0.9370	0.0550	0.8905	0.0713	0.8272	0.0847	0.7489	0.0934	0.6593
18	0.0255	0.9626	0.0397	0.9302	0.0554	0.8826	0.0706	0.8195	0.0830	0.7423
19	0.0161	0.9787	0.0272	0.9573	0.0409	0.9235	0.0557	0.8752	0.0699	0.8122
20	0.0097	0.9884	0.0177	0.9750	0.0286	0.9521	0.0418	0.9170	0.0559	0.8682
21	0.0055	0.9939	0.0109	0.9859	0.0191	0.9712	0.0299	0.9469	0.0426	0.9108
22	0.0030	0.9970	0.0065	0.9924	0.0121	0.9833	0.0204	0.9673	0.0310	0.9418
23	0.0016	0.9985	0.0037	0.9960	0.0074	0.9907	0.0133	0.9805	0.0216	0.9633
24	0.0008	0.9993	0.0020	0.9980	0.0043	0.9950	0.0083	0.9888	0.0144	0.9777
25	0.0004	0.9997	0.0010	0.9990	0.0024	0.9974	0.0050	0.9938	0.0092	0.9869
26	0.0002	0.9999	0.0005	0.9995	0.0013	0.9987	0.0029	0.9967	0.0057	0.9925
27	0.0001	0.9999	0.0002	0.9998	0.0007	0.9994	0.0016	0.9983	0.0034	0.9959
28	0	1	0.0001	0.9999	0.0003	0.9997	0.0009	0.9991	0.0019	0.9978
29	0	1	0.0001	1	0.0002	0.9999	0.0004	0.9996	0.0011	0.9989
30	0	1	0	1	0.0001	0.9999	0.0002	0.9998	0.0006	0.9994
31	0	1	0	1	0	1	0.0001	0.9999	0.0003	0.9997
32	0	1	0	1	0	1	0	1	0.0001	0.9999
33	0	1	0	1	0	1	0	1	0.0001	0.9999

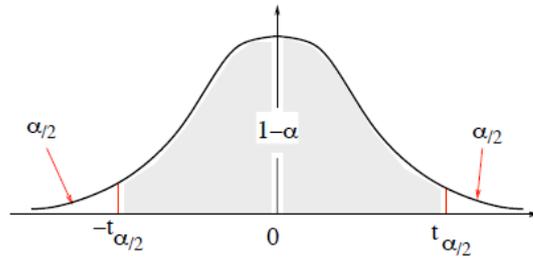
# LOI DE LAPLACE-GAUSS



La table donne  $\Phi(t) = P(0 < X < t)$

t	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
4.0	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

## LOI DE STUDENT



La table donne  $t_{\alpha/2}$  tel que  $P(|T_\nu| > t_{\alpha/2}) = \alpha$ .

$\nu \backslash \alpha$	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	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	31.82	63.65	636.6
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	31.59
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	12.92
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	3.768
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
60	0.126	0.254	0.387	0.527	0.679	0.848	1.045	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

# Bibliographie

- [1] N. Bouleau, Probabilités de l'ingénieur, Hermann, 1986.
- [2] S. M. Ross, Initiation aux probabilités", PPUR, 1987.
- [3] D. Revus, Probabilités, Hermann 1994.
- [4] Philippe Barbe et Michel Ledoux, Probabilité, 2007, EDP Sciences.
- [5] Hervé Carrieu, Probabilité : Exercices corrigés, 2008, EDP Sciences
- [6] Cours de statistique de J. Gergaud donné à l'école supérieur d'agronomie de Toulouse.
- [7] Introduction à la méthode statistique. Manuel et exercices corrigés. Bernard Goldfarb Catherine Pardoux. Dunod.
- [8] Série Schaum Algèbre linéaire 1 Seymour Lipshutz Mc Graw Hill.
- [9] Biomathématiques Analyse Algèbre probabilités statistiques, S Bénazeth & all Masson L 510-821.
- [10] Méthodes statistiques en biologie et en agronomie A.Vessereau L/519.041
- [11] Mini manuel de mathématiques pour les sciences de la vie et de l'environnement : cours et exos corrigés Driss Boularas, Fredon Daniel, Petit Daniel L/510.1003.