2° partie : Probabilités

1. Introduction

C'est grâce aux jeux de hasard auxquels se sont intéressés le philosophe et écrivain Pascal (1623-1662) et le mathématicien Fermat (1601-1665) qu'est né le calcul des probabilités.
Originellement, ils se sont penchés sur le problème suivant :

Jeu de PILE OU FACE interrompu
Comment partager les gains ?
Comment répartir les enjeux, compte tenu des parties déjà jouées?

Ce problème a été posé par le chevalier Méré, et a été résolu par Pascal, expliqué dans une lettre qu'il envoie à Fermat en 1654.
Je vous invite à essayer vous-même de le résoudre, compte tenu des hypothèses suivantes :

Deux joueurs A et B jouent en trois manches. Ils misent chacun 32 pistoles (ou €, ou …). Le premier qui totalisera trois manches gagnantes reçoit les 64 pièces jouées.
Le jeu est équitable : Le jeu est symétrique. La probabilité de gain de l'un est égale à la probabilité de gain de l'autre. Pour ***A*** le gain peut-être direct: ***GGG*** ou indirect: ***G***PP***GG***. Même chose (réciproque) pour B, bien sûr.

Le jeu est interrompu alors que A mène 1 - 0 !

Bon amusement …

La théorie des probabilités est une théorie qui veut mathématiser des situations dont l'issue est **incertaine**, dont la réalisation est **due au hasard**.

2. Définitions

2.1. Expérience aléatoire

Une *expérience aléatoire* est une expérience dont le résultat ne peut pas être prévu (du moins avec certitude) et dont on connaît au préalable l'ensemble des résultats possibles.

L'ensemble E (ou Ω) de tous les résultats possibles est *l'espace des éventualités* (ou catégorie d'épreuve ou ensemble fondamental ou univers des possibles).

Remarque
On se limitera aux cas où l'ensemble des résultats possibles est fini (de type « discret » cf statistiques).

2.2 Événement

Un *événement* A d'une expérience aléatoire est un sous-ensemble de l'ensemble E des résultats possibles.
Donc $A \subset E$.
L'événement A est dit réalisé lors d'une expérience aléatoire si le résultat de l'expérience lui appartient.

Exemples

L'expérience consiste à jeter un dé et à noter le résultat obtenu.
Il y a 6 résultats possibles : l'apparition des points 1, 2, 3, 4, 5 et 6 donc l'ensemble des résultats possibles est : $E = \{ 1, 2, 3, 4, 5, 6 \}$

Un événement A est, par exemple : " obtenir un nombre pair "; $A = \{2, 4, 6\}$

En particulier :

- l'ensemble vide $(\varnothing)$ représente **l'événement impossible** (il ne se réalise jamais),
  Ex : obtenir le nombre 7

- l'ensemble E (ou Ω) représente **l'événement certain** *(ou sûr*, il se réalise toujours),
  Ex : obtenir un nombre pair ou impair

- tout singleton est appelé **événement élémentaire,**
  Ex : obtenir le point 2. Donc A=$\{2\}$ et A contient un seul élément

- $A \cup B$ est l'événement qui se produit si l'un des événements A **ou** B se produit (A se réalise ou B se réalise ou les deux se réalisent),

  Ex : Soit A : « obtenir un nombre pair » et B : « obtenir un point divisible par 5 »
  Donc $A \cup B$ est l'événement : « obtenir un point pair **OU** divisible par 5 »
  et $A \cup B = \{2,4,5,6\}$

- $A \cap B$ est l'événement qui se produit si A **et** B sont tous les deux réalisés,
  Ex : Soit A : « obtenir un point pair » et B : « obtenir un point divisible par 3 »
  Donc $A \cap B$ est l'événement : « obtenir un point pair **ET** divisible par 3 »
  et $A \cap B = \{6\}$

- Si $A \cap B = \varnothing$, A et B sont **disjoints** ou **incompatibles** (A et B ne peuvent se réaliser en même temps),

  Ex : Soit A : « obtenir un point pair » et B : « obtenir un point divisible par 5 » (dans le cadre du lancé du dé bien sûr)

- $\overline{A}$ est le **complémentaire** de A : c'est l'événement qui se produit si A n'est pas réalisé. A et $\overline{A}$ sont des événements **contraires** (l'un de l'autre).
  Donc $A \cup \overline{A} = E$ et $A \cap \overline{A} = \varnothing$.
  Ex : Soit A : « obtenir un point pair » et B : « obtenir un point impair »

3. Probabilité

3.1. Introduction

On lance plusieurs fois un dé non pipé (truqué) et on observe le résultat obtenu. On s'intéresse à la question suivante : "est-ce un 3 ou non ?"
On obtient les résultats suivants :

1re expérience

Nombres de lancers	Nombres de 3 obtenus	Fréquence	Nombre de "non 3"	Fréquence
10	3	0,3	7	0,7
20	4	0,2	16	0,8
30	5	0,166…	25	0,833…
50	8	0,16	42	0,84
100	15	0,15	85	0,85
500	83	0,166	417	0,834
800	130	0,1625	670	0,8375
2000	333	0,1665	1667	0,8335

2e expérience

Nombres de lancers	Nombres de 3 obtenus	Fréquence	Nombre de "non 3"	Fréquence
10	2	0,2	8	0,8
20	3	0,15	17	0,85
30	5	0,166…	25	0,833…
50	9	0,18	41	0,82
100	17	0,17	83	0,83
500	84	0,168	416	0,832
800	128	0,16	672	0,84
2000	331	0,1655	1669	0,8345

A priori, en lançant un dé non truqué, on a une chance sur 6 d'obtenir un "3".
C'est une manière de voir les choses. En fait, nous faisons l'hypothèse que chaque numéro présente *à priori* la même « éventualité » de sortir ; c'est-à-dire que nous faisons implicitement appel au préalable aux *propriétés de symétrie* de la situation considérée. C'était notamment le point de vue de Pascal.

Comme le dé n'est pas truqué, il n'est pas surprenant d'obtenir des fréquences relatives d'apparition du "3" voisines de 0,1666... et de 0,8333… pour l'obtention d'un autre résultat que "3".

On constate que plus le nombre de lancers est grand, plus la fréquence est voisine de la fréquence idéale qui est 0,1666… = 1/6 pour obtenir un "3".

La **probabilité de l'événement** A : "obtenir un 3" est :

$$P(A) = \lim_{n \to \infty} f_n = \lim_{n \to \infty} \frac{p_i}{n} = 0{,}1666... = \frac{1}{6}$$

où f_n représente la fréquence avec laquelle l'événement A s'est produit lors de n lancers et p_i le nombre de cas favorables lors de la i$^{\text{ième}}$ expérience lors de n lancers.
C'est par contre cette approche fréquentielle *à posteriori* qu'a adoptée Bernouilli pour définir la notion de probabilité.

On observe également que :

- Toutes les fréquences calculées sont comprises entre 0 et 1.
- La somme des fréquences pour un même nombre de lancers est 1.

3.2. Définition (axiomatique)

Une probabilité P est une application de l'ensemble des événements dans l'ensemble des réels telle que, pour tout événement A, on ait :

- $P(A) \geq 0$;
- $P(E) = 1$:probabilité de l'événement certain et $P(\varnothing)$ =0 : probabilité d'un événement impossible;
- si $A \cap B = \varnothing$ alors $P(A \cup B) = P(A) + P(B)$: la probabilité de deux événements disjoints A et B est la somme des probabilités de ces événements :

Ces trois axiomes sont appelés : "**axiomes de Kolmogorov**". (Mathématicien russe 1903-1987)

3.3. Propriétés

a) $\boxed{\forall A, B \subset E : \quad A \subset B \quad \Rightarrow \quad P(A) \leq P(B)}$

En effet : $A \subset B$, on en déduit que : $B = A \cup (B \cap \overline{A})$ et $A \cap (B \cap \overline{A}) = \varnothing$

$$P(B) = P(A) + \underbrace{P(B \cap \overline{A})}_{\text{la probabilité est un nombre positif}} \geq P(A)$$

b) $\boxed{0 \leq P(A) \leq 1}$ En effet : $A \subset E \quad \Rightarrow \quad P(A) \leq P(E)$

$$P(A) \leq 1$$

c) $\boxed{P(\varnothing) = 0}$

Considérons un événement quelconque A. On a

$$A \cap \varnothing = \varnothing \quad \text{et} \quad A \cup \varnothing = A$$

Donc

$$P(A \cup \varnothing) = P(A) + P(\varnothing)$$
$$P(A) = P(A) + P(\varnothing)$$
$$P(\varnothing) = 0$$

d) $\boxed{\forall A \subset E: \quad P(\overline{A}) = 1 - P(A)}$

En effet :

$$A \cap \overline{A} = \varnothing$$
$$A \cup \overline{A} = E$$

ce qui entraîne

$$P(A \cup \overline{A}) = P(A) + P(\overline{A})$$
$$P(E) = P(A) + P(\overline{A})$$
$$1 = P(A) + P(\overline{A})$$
$$P(\overline{A}) = 1 - P(A)$$

Exemples

- Dans l'expérience aléatoire d'un jet de dé : "obtenir un point pair" et "obtenir un point impair" sont deux événements contraires.

- Dans l'expérience du jet d'un dé, l'événement A "obtenir un 3" et l'événement B "obtenir un autre résultat que 3" sont contraires.

$$P(A)=\frac{1}{6} \qquad \text{et} \qquad P(B)=\frac{5}{6}$$

ce qui donne, comme prévu,

$$P(A)+P(B)=1$$

Application importante : probabilité de l'événement "au moins un"

L'événement A "obtenir au moins un …" est l'événement contraire de l'événement $\overline{A}$ "n'obtenir aucun …". Donc $P(A)=1-P(\overline{A})$.

En pratique, il est souvent plus facile de calculer $P(\overline{A})$ et d'en tirer $P(A)$.

Exemple

Reprenons l'exemple du début : l'expérience consiste à lancer 3 fois de suite une pièce de monnaie et à observer la face supérieure des jets consécutifs.

$$P(\text{obtenir au moins 1 fois "face"})=1-P(\text{ne pas obtenir "face"})=1-\frac{1}{8}=\frac{7}{8}$$

c) $\boxed{\forall A,B\subset E:\quad P(A\cup B)=P(A)+P(B)-P(A\cap B)}$ (relation de Boole)

En effet :

En effet : $A\cup B=A\cup(B\cap\overline{A})$ donc $P(A\cup B)=P(A)+P(B\cap\overline{A})$ (1)

$B=(B\cap\overline{A})\cup(A\cap B)$ donc $P\left(B\right)=P(B\cap\overline{A})+P(A\cap B)$ (2)

(1) $P(A\cup B)=P(A)+P(B\cap\overline{A})$

(2) $P(B)=P(B\cap\overline{A})+P(A\cap B)$ donc $P(B\cap\overline{A})=P(B)-P(A\cap B)$

Remplaçons $P(B\cap\overline{A})$ par sa valeur dans la relation (1) ; il vient :

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

f) Généralisation du troisième axiome de Kolmogorov

Rappel : on a (3è axiome) : $\forall\ A,B\subset E:\qquad A\cap B=\varnothing\ \Rightarrow\ P(A\cup B)=P(A)+P(B)$

On peut le généraliser :

$$\forall A_1,A_2,\dots,A_n\subset E;avecA_i\cap A_j=0,\forall i\neq j:i,j=1,\dots,n$$
$$P(A_1\cup A_2\cup\dots\cup A_n)=P(A_1)+P(A_2)+\dots+P(A_n)$$

3.4. Equiprobabilité

Deux événements sont <u>équiprobables</u> lorsqu'ils ont la même probabilité de se produire.

<u>Exemple</u>
Lors du jet d'une pièce de monnaie, "obtenir pile" et "obtenir face" sont des événements équiprobables (du moins en considérant une pièce bien équilibrée).

Conséquence : Dans beaucoup de situations (**<u>mais pas dans toutes</u>**) les événements élémentaires de l'expérience aléatoire sont équiprobables, dès lors :
Dans le cas d'un univers des possibles $E = e_1, e_2, ..., e_n$, où $\{e_1\}, \{e_2\}, ..., \{e_n\}$ sont les événements élémentaires : $\{e_i\} \cap \{e_j\} = \varnothing$, $\forall i \neq j$: $i, j = 1, ..., n$.
En appliquant la généralisation du troisième axiome de Kolmogorov, on obtient :

$$P(E) = P(e_1) + P(e_2) + ... + P(e_n) = 1$$

Si les événements <u>élémentaires</u> sont <u>équiprobables</u>, on a : $\boxed{P(e_1) = P(e_2) = ... = P(e_n) = \frac{1}{n}}$

Considérons l'événement A formé de p événements élémentaires équiprobables $A = \{e_1, e_2, ..., e_p\}$.
On a $P(A) = P(\{e_1\}) + P(\{e_2\}) + ... + P(\{e_p\})$ et donc

$$P(A) = \frac{1}{n} + \frac{1}{n} + ... + \frac{1}{n} = \frac{p}{n}$$

p est le nombre de cas favorables à A, n est le nombre de cas possibles.

Dans le cas **<u>d'événements élémentaires équiprobables</u>**, la probabilité d'un événement se traduit par

$$\boxed{\text{probabilité d'un événement} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}}$$

Cette règle porte le nom de "Règle de Laplace".

<u>Exemples</u>

1. La probabilité d'obtenir un point pair en jetant un dé non truqué est $\frac{3}{6}$.
2. Lançons simultanément deux dés non truqués, l'un noir et l'autre blanc. On observe les points indiqués d'abord par le dé noir, puis par le dé blanc.

$$E = \{ \begin{array}{l} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\} \end{array}$$

Il y a donc 36 cas possibles (36 événements élémentaires, tous équiprobables).

- Quelle est la probabilité que la somme des points obtenus lors du lancement de deux dés soit 8 (événement A) ?

  On a : $A = \{(2,6) , (3,5) , (4,4) , (5,3) , (6,2)\}$

  D'où p(A) = $\frac{5}{36}$

- Quelle est la probabilité d'obtenir des points identiques lors du lancement de deux dés (événement B) ?

On a : $B = \{(1,1)\ ,\ (2,2)\ ,\ (3,3)\ ,\ (4,4)\ ,\ (5,5)\ ,\ (6,6)\}$

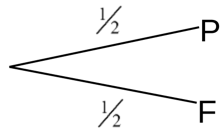
D'où p(B) = $\frac{6}{36} = \frac{1}{6}$

3. L'exemple suivant montre le danger qu'il y a d'utiliser la formule des événements équiprobables sans réfléchir.

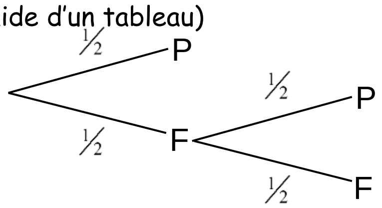
Un joueur lance une pièce de monnaie parfaitement symétrique.
Si le côté "pile" sort, le jeu s'arrête et le joueur a gagné.
Si le côté "face" sort, le joueur relance la pièce et le jeu s'arrête définitivement.
Si lors du second lancer le côté "pile" sort, le joueur a gagné, sinon il a perdu.
Quelle est la probabilité de gain du joueur ?

D'Alembert (1717-1783) étudia ce problème et en publia une solution fausse dans l'Encyclopédie :
"il y a trois cas possibles dont deux sont favorables, la probabilité demandée vaut donc $\frac{2}{3}$ ".
D'Alembert appliquait la formule des événements équiprobables alors qu'ils ne l'étaient pas !

Au premier jet, les probabilités se répartissent de la façon suivante :



Considérons un deuxième jet ; les probabilités se répartissent de la façon suivante :
(Rem. : on peut aussi représenter la situation à l'aide d'un tableau)



La probabilité de gain pour le joueur est donc de $\frac{1}{2} + \frac{1}{4} = \frac{3}{4}$ et non de $\frac{2}{3}$.

4. Probabilité conditionnelle.

4.1 Définition

La probabilité d'un événement est souvent modifiée quand on apprend qu'un autre événement s'est produit (par exemple, la probabilité d'entendre un tonnerre pendant un orage change soudainement si on voit un éclair!).
La nouvelle valeur de la probabilité d'un événement *A* quand on sait que l'événement *B* s'est produit, est la *probabilité conditionnelle de A étant donné B*, et est représentée par $P(A|B)$(et se lit probabilité de l'événement *A* **à la condition** B).
En d'autres termes,

> **La probabilité conditionnelle d'un événement A par rapport à un événement B est la probabilité que A se réalise lorsque B est réalisé.** On la note $P(A/B)$

En fait, l'opération de conditionnement est équivalente à une redéfinition de l'espace des réalisations, (ensemble des possibles) qui est contraint au sous-ensemble de l'espace initial où l'événement conditionnant est vérifié.

Exemple 1:

Un groupe de cadres est classé selon son poids et l'incidence d'hypertension. Les proportions dans les différentes catégories sont données par le tableau suivant:

	obèse	normal	maigre	total
hypertension	.10	.08	.02	.20
pas d'hypertension.	.15	.45	.20	.80
Total	.25	.53	.22	1.00

Si l'on sait qu'une personne est obèse, alors les catégories correspondantes aux deuxième et troisième colonnes n'ont pas d'intérêt. On peut conclure de ce tableau, par exemple, qu'étant donné qu'une certaine personne est obèse, la probabilité pour qu'elle souffre d'hypertension passe de la valeur (non-conditionnée) de 0.2 à 0.1/0.25=0.4, c'est-à-dire deux fois supérieures.

L'exemple précédent illustre l'application de la formule qui définit les probabilités conditionnelles:

4.2. Formule de calcul de $P(A/B)$:

La probabilité conditionnelle de l'événement *A* étant donné l'événement *B* est, par définition:

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

Exemple 2 :

On lance deux dés, un blanc et un noir. Quelle est la probabilité que le dé blanc donne "1" si la somme des points obtenus sur les deux dés est strictement inférieure à 6 ?

Appelons A l'événement "le 1 du dé blanc apparaît" et B l'événement "la somme des points des deux dés est strictement inférieure à 6"

$$P(A) = \frac{6}{36} \quad , \quad P(B) = \frac{10}{36} \quad , \quad P(A \cap B) = \frac{4}{36}$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{4}{36}}{\frac{10}{36}} = \frac{4}{10} = \frac{2}{5}$$

4.3. Théorème de la multiplication[1]

La probabilité de la réalisation simultanée de deux événements vaut le produit de la probabilité de l'un par la probabilité conditionnelle de l'autre par rapport au premier.

$$\boxed{\begin{aligned} P(A \cap B) &= P(A)\,.\,P(B/A) \\ &= P(B)\,.\,P(A/B) \end{aligned}}$$

Exemple 1

On tire successivement et <u>sans remise</u> deux cartes dans un jeu de 52 cartes. Quelle est la probabilité de tirer

1°) un as au premier tirage ?

2°) un deuxième as au deuxième tirage sachant qu'un as est déjà sorti ?

3°) deux as ?

Soient A l'événement « tirer un as au premier tirage », B l'événement « tirer un as au deuxième tirage » et C l'événement « tirer deux as ».

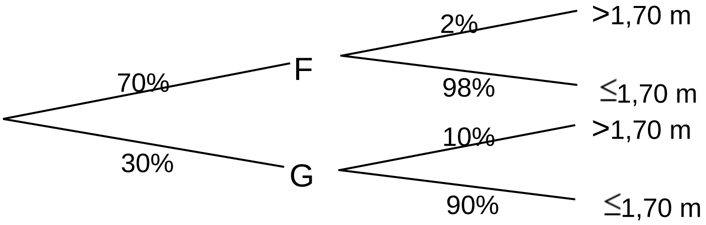
1°) $P(A) = \dfrac{4}{52} = \dfrac{1}{13}$.

2°) $P(B/A) = \dfrac{3}{51} = \dfrac{1}{17}$.

3°) $P(C) = P(A \cap B) = P(A).P(B/A) = \dfrac{4}{52}.\dfrac{3}{51} = \dfrac{1}{221}$

Exemple 2

Dans une classe d'âge d' élèves, 10% des garçons et 2% des filles mesurent plus de 1,70 m. On sait que 70% des élèves sont des filles. Si l'on prend un élève au hasard et que celui-ci mesure plus de 1,70 m , quelle est la probabilité que ce soit une fille ?



Soient A l'événement « mesurer plus de 1,70 m », F l'événement « être une fille » et G l'événement « être un garçon ».

$$P(F/A) = \frac{P(F \cap A)}{P(A)} = \frac{P(F).P(A/F)}{P(G).P(A/G) + P(F).P(A/F)} = \frac{70\%.2\%}{30\%.10\% + 70\%.2\%} = \frac{140}{440} = 0{,}318$$

[1] Ce théorème peut être généralisé sans difficulté.

4.4 Événements indépendants

Deux événements sont *indépendants* si leurs probabilités ne changent pas quand on conditionne l'un à l'autre. Dans ces conditions, l'information que *B* s'est produit ne change pas la vraisemblance de l'occurrence de *A* (et inversément) :

Deux événements sont <u>indépendants</u> si la réalisation de l'un n'a aucune influence sur la réalisation de l'autre et réciproquement.

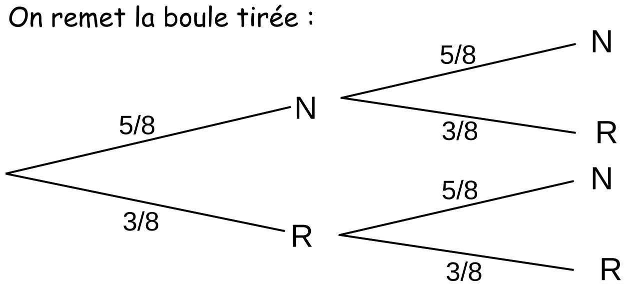
A **et** B **sont indépendants si et seulement si** $P(A/B)=P(A)$ **ou** $P(B/A)=P(B)$

Le théorème de la multiplication devient alors: $\boxed{P(A\cap B)=P(A).P(B)}$

Exemples

1. Considérons une urne contenant 5 boules noires et 3 boules rouges.
Retirons successivement 2 boules de l'urne.
Considérons les événements A et B suivants :
A = "tirer une boule rouge au premier tirage"
B = "tirer une boule rouge au deuxième tirage"
Ces événements sont-ils indépendants?

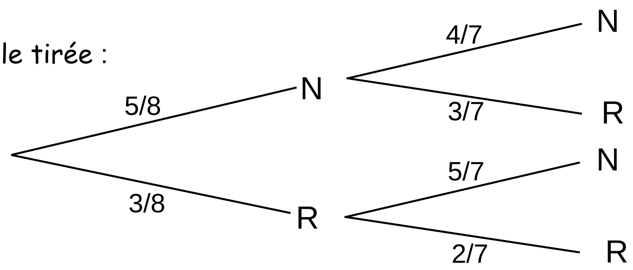
- On remet la boule tirée :



$$P(A)=\frac{3}{8}\ ;\ P(B)=\frac{5}{8}.\frac{3}{8}+\frac{3}{8}.\frac{3}{8}=\frac{3}{8}\ ;\ P(B/A)=\frac{3}{8}$$

Puisque $P(B/A)=P(B)$, les événements sont indépendants : la probabilité du deuxième tirage ne dépend pas du résultat du premier tirage.

- On ne remet pas la boule tirée :



$$P(A)=\frac{3}{8}\ ;\ P(B)=\frac{5}{8}.\frac{3}{7}+\frac{3}{8}.\frac{2}{7}=\frac{15+6}{56}=\frac{21}{56}=\frac{3}{8};\ P(B/A)=\frac{2}{7}$$

Les deux tirages sont donc dépendants : la probabilité du deuxième tirage dépend du résultat du premier tirage.

3° partie : Lois de probabilités

1. Variable aléatoire

1.1 Définitions

Considérons une expérience aléatoire dont l'ensemble fini des résultats est E.
A chacun des éléments e_i de E associons un nombre réel x_i .
On définit ainsi une fonction de E dans $\boldsymbol{R}$ appelée **variable aléatoire et notée X.**

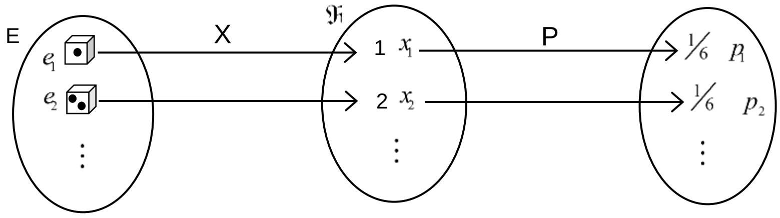
Une variable aléatoire est discrète si l'ensemble de ses valeurs possibles est fini ou dénombrable.
Dans le cas contraire elle est continue.

N.B. : La distinction entre variables discrètes et continues est identique à celle rencontrée en Statistique en 4°.

Exemples

1. Considérons le jet d'un dé.

$X : E \to \boldsymbol{R}$
$e_i \to x_i$



La variable aléatoire X prend les valeurs 1,2,3,4,5 et 6. Elle est discrète.

2. On jette simultanément 4 pièces de monnaie indiscernables. On associe à chaque résultat le nombre de "face" obtenu. La variable aléatoire X ainsi définie peut prendre les valeurs 0,1,2,3,4
Elle est discrète.

3. On jette simultanément 2 dés marqués chacun 1 sur trois faces, 2 sur deux faces et 5 sur une face. On associe à chaque résultat la somme des points obtenus. La variable aléatoire X ainsi définie peut prendre les valeurs 2, 3, 4, 6, 7et 10. Elle est discrète.

4. On mesure les tailles d'un groupe d'individus se situant entre 145 cm et 200 cm. On associe à chaque résultat la taille obtenue. La variable aléatoire peut prendre des valeurs comprises entre 145 et 200. Elle est continue.

2. Variable aléatoire discrète

2.1. Distribution discrète de probabilité

La distribution de probabilité d'une variable aléatoire discrète est une **fonction** qui à x_i associe la probabilité p_i d'obtenir cette valeur :$p_i = p(X = x_i)$ avec *i = 1,...,k* où *k* est le nombre de résultats possibles.

La représentation graphique de la distribution de probabilité est un diagramme en bâtonnets (voir page suivante).

Remarque : $\sum_{i=1}^{k} p_i = 1$

Exemple :
Un magasin solde 15 radios, dont 5 sont défectueuses. Si on essaye trois radios différentes, prises au hasard parmi les 15, quelle est la distribution de probabilité de la variable aléatoire X si
X désigne le nombre de radios défectueuses détectées ?

Réponse: P(x=0)=24/91; P(X=1)=45/91; P(x=2)=20/91; P(x=3)=2/91. (A vérifier sur feuille séparée en utilisant l'analyse combinatoire)

2.2 Fonction de répartition

En considérant les probabilités cumulées, on obtient la fonction de répartition définie par :

$$\boxed{F : \mathbf{R} \to [0,1] \quad , \quad F(x) = P(X \le x)}.$$

c.à-d.la probabilité pour que la variable aléatoire prenne des valeurs inférieures à une certaine valeur *x*.

Des propriétés des probabilités, résultent certaines propriétés de la fonction de répartition:

- $F_X(x)$ est une fonction croissante de *x*.
- Les valeurs de la fonction passeront de 0 à 1 quand x varie de moins l'infini à plus l'infini:
  (de façon « imagée » : $F_X(-\infty) = 0,\ F_X(\infty) = 1$)

La représentation graphique de la fonction de répartition est une fonction en escalier.

Reprenons les **exemples**[2] **2 et 3 :**
Exemple 2 : La variable aléatoire X prend les valeurs 0, 1, 2, 3, 4.

x_i	0	1	2	3	4
p_i	1/16	4/16	6/16	4/16	1/16
$F(x_i)$	1/16	5/16	11/16	15/16	1

Par exemple : $F(2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} = \frac{11}{16}$

C'est la probabilité d'obtenir au plus deux fois face.

Notons que p.ex. : $F(1{,}7) = F(1) = \frac{5}{16}$!!! … nous aurons bien une fonction en escalier.

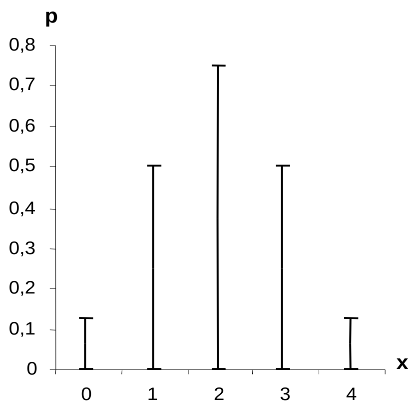
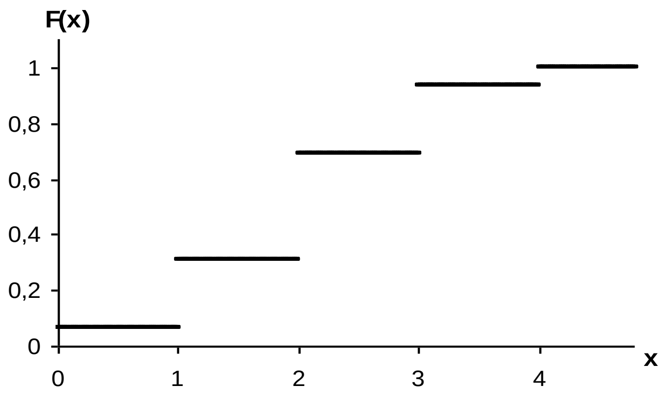


Diagramme en bâtons



fonction de répartition

[2] Le diagramme représentant la fonction de répartition est incomplet : il y a lieu de préciser quelles extrémités des segments horizontaux (« marches ») appartiennent aux graphiques.

Exemple 3 : La variable aléatoire X prend les valeurs 2, 3, 4, 6, 7, 10.

x_i	2	3	4	6	7	10
p_i	9/36	12/36	4/36	6/36	4/36	1/36
$F(x_i)$	9/36	21/36	25/36	31/36	35/36	36/36

Ainsi p.ex. : $F(4) = P(X=2) + P(X=3) + P(X=4) = \dfrac{25}{36} ... = F(4{,}2) = F(4{,}9) = F(5{,}6) = ...$

C'est la probabilité que la somme des points obtenus sur les deux dés soit inférieure ou égale à 4, mais aussi inférieure ou égale à tout nombre supérieur ou égal à 4 et strictement inférieur à 6.

2.3. Paramètres d'une distribution discrète

Considérons une variable aléatoire discrète X prenant les valeurs x_i et les probabilités p_i correspondantes. $p_i = P(X = x_i)$ avec *i = 1,…,k* où k est le nombre de résultats possibles.

- La <u>*moyenne m*</u> ou <u>*espérance mathématique E(X)*</u> de la variable aléatoire est le nombre réel défini par :

$$\boxed{E(X) = \sum_{i=1}^{k} p_i\, x_i}$$

Remarque

Dans les jeux de hasard, si x_i représente un gain, <u>l'espérance mathématique représente le gain moyen</u> d'un joueur.

- Le <u>*mode*</u> est la (ou les) valeur(s) de X donnant la probabilité maximum.

- La <u>*variance*</u> $V(X)$ traduit l'écart par rapport à la moyenne, elle est définie par :

$$V(X) = \sum_{i=1}^{k} p_i (x_i - E(X))^2$$

En pratique on utilisera plutôt l'expression suivante :

$$\boxed{V(X) = \sum_{i=1}^{k} p_i x_i^{\,2} - \left[E(X)\right]^2}$$

La démonstration est laissée à titre d'exercice.

L'<u>écart type</u> est la racine carrée de la variance : $\boxed{\sigma(X) = \sqrt{V(X)}}$

Il est bien évident que ces notions sont tout à fait à comparer à celles rencontrées en statistique.

Exemple :

Un joueur lance un dé : s'il fait 1, il perd 10 euros ; s'il fait 2, il perd 6 euros ; s'il fait 3 ou 4, il ne perd ni ne gagne ; s'il fait 5, il gagne 2 euros ; et s'il fait 6, il gagne 20 euros .

On associe à chaque résultat du dé le gain correspondant. La variable aléatoire X ainsi définie peut prendre les valeurs : -10, -6, 0, 2, 20.

x_i	-10	-6	0	2	20
p_i	1/6	1/6	2/6	1/6	1/6
$p_i \, . x_i$	-10/6	-6/6	0	2/6	20/6

- $E(X) = \dfrac{-10}{6} + \dfrac{-6}{6} + \dfrac{2}{6} + \dfrac{20}{6} = \dfrac{6}{6} = 1$ Signification : le joueur gagnera en moyenne 1 € par partie.
- Mode : 0

- $V(X)=\frac{1}{6}.\left((-10-1)^2+(-6-1)^2+2(0-1)^2+(2-1)^2+(20-1)^2\right)=89$

$\sigma(X)=9{,}43$ On constate donc que l'écart type représente une estimation de l'écart moyen entre le gain moyen et le gain réel.

2.4. Cas particulier 1: La variable binomiale

2.4.1. Définition

On considère une expérience aléatoire. Une ***épreuve de Bernoulli*** est une épreuve dans laquelle on ne s'intéresse qu'à la réalisation (succès) ou la non réalisation (échec) d'un événement A lié à cette épreuve : $P(A)+P(\overline{A})=1$.

A se produit avec une probabilité p et donc $\overline{A}$ se produit avec une probabilité $q=1-p$.

Un schéma de Bernoulli est une suite de n épreuves de Bernoulli identiques, indépendantes entre elles avec la même probabilité de succès pour chacune d'elles.

Exemple

On lance un dé équilibré. On s'intéresse à l'événement A : "obtenir un 6".

$P(A)=\frac{1}{6}$ et $P(\overline{A})=\frac{5}{6}$ est une épreuve de Bernoulli.

Le lancement de ce dé 10 fois de suite donne un schéma de Bernoulli.

2.4.2. Distribution binomiale (ou loi binomiale)

On considère n répétitions indépendantes d'une épreuve de Bernoulli. A chaque répétition, l'événement A peut se produire avec une probabilité p. La variable aléatoire X est le nombre de réalisations de A au cours des n expériences. Elle peut prendre les valeurs 0, 1, 2, …, n. Dans l'exemple précédent, on s'intéresse au nombre de fois que l'on obtient le "6" comme résultat au cours des 10 lancers du dé.

La probabilité que A se réalise k fois lors des n répétitions est

$$\boxed{P(X=k)=C_n^k\ p^k\,q^{n-k}}$$

En effet

- Il existe plusieurs manières de réaliser k fois A au cours des n répétitions de l'expérience : ce sont les combinaisons de k éléments choisis parmi n : C_n^k .
- Chacune de ces manières a une probabilité $p^k\,q^{n-k}$ de se produire : A se produit k fois et $\overline{A}$ se produit donc n-k fois (les événements sont indépendants).

La variable aléatoire X ainsi définie est appelée *variable binomiale* et est notée $B(n,p)$, où n est le nombre d'expériences et p la probabilité de succès lors de chaque expérience.

Exemple

Une boite contient 10 boules : 2 rouges et 8 noires.

On va tirer successivement 5 boules, avec remise de la boule tirée à chaque fois.

On a ainsi un schéma de Bernoulli. On s'intéresse à la variable aléatoire X "nombre de boules noires obtenues".

Quand on tire une boule, la probabilité qu'elle soit noire est de 0,8 et la probabilité qu'elle soit rouge est de 0,2.

La variable X est une variable binomiale : $B(5\,;\,0{,}8)$.
Les valeurs possibles de X sont : 0, 1, 2, 3, 4, 5.

- Probabilité d'obtenir 0 boule noire : il faut tirer successivement 5 boules rouges.
$$P(X=0)=(0{,}2)^5=0{,}00032$$
- Probabilité d'obtenir 1 boule noire (et 4 rouges) : on a $C_5^1=5$ manières d'obtenir une boule noire parmi les 5 tirages. $$P(X=1)=C_5^1\,(0{,}8).(0{,}2)^4=0{,}0064$$
- De même : Probabilité d'obtenir 2 boules noires (et 3 rouges) : on a $C_5^2=10$ manières d'obtenir 2 boules noires parmi les 5 tirages.
$$P(X=2)=C_5^2\,(0{,}8)^2.(0{,}2)^3=0{,}0512$$
… et ainsi de suite

2.4.3. Paramètres de la distribution binomiale

- Moyenne ou espérance mathématique : $\boxed{E(X)=np}$

Elle représente le nombre moyen de succès au cours de n répétitions de l'expérience aléatoire.

Démonstration

$$\begin{aligned}E(X)&=\sum_{k=1}^{n}k\,.P(X=k)\\&=\sum_{k=1}^{n}k\,.C_n^k\;p^k\,q^{n-k}\\&=\sum_{k=1}^{n}k\,.\frac{n!}{k!.(n-k)!}\,p^k q^{n-k}\\&=\sum_{k=1}^{n}\frac{n!}{(k-1)!(n-k)!}\;p^k\;q^{n-k}\\&=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!(n-k)!}\,p^{k-1}\,q^{n-k}\end{aligned}$$

En posant $i=k-1$, on obtient :

$$\begin{aligned}E(X)&=np\sum_{i=0}^{n-1}\frac{(n-1)!}{i!\,(n-1-i)!}\,p^i\,q^{n-1-i}\\&=np\sum_{i=0}^{n-1}C_{n-1}^i\;p^i\,(1-p)^{n-1-i}\end{aligned}$$

Grâce au binôme de Newton, cette expression devient

$$\begin{aligned}E(X)&=np\;\left(p+(1-p)\right)^{n-1}\\&=np\end{aligned}$$

- Variance : $\boxed{V(X)=n\,p\,q}$

Elle représente la moyenne des carrés des écarts par rapport à la moyenne (même notion qu'en statistique)

- Ecart-type : $\boxed{\sigma(X)=\sqrt{n\,p\,q}}$

Exemple
Pour l'exemple précédent, on a
$$E(X)=n\,p=5.0{,}8=4$$
En moyenne sur les 5 tirages on obtiendra 4 boules noires.
$$\sigma(X)=\sqrt{5.0{,}8.0{,}2}=0{,}894$$

La moyenne des écarts à la moyenne est proche de 0,894.

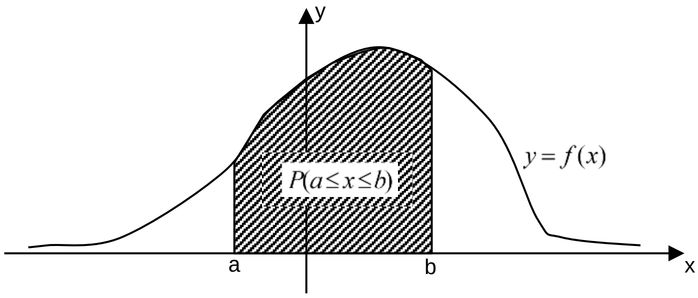
3. Variable aléatoire continue

3.1. Densité de probabilité

Reprenons l'exemple de la taille des individus. La variable aléatoire X ainsi obtenue est continue.

La représentation graphique de cette variable est une courbe continue appelée courbe de distribution de fréquence dont l'équation est donnée par : $y = f(x)$.

La fonction f est appelée densité de probabilité.



Cette courbe vérifie les propriétés suivantes

- $P(a \le x \le b) = \int_a^b f(x)dx$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

Remarque
La probabilité pour que x prenne exactement une valeur donnée est nulle.

3.2. Fonction de répartition

La fonction de répartition d'une variable continue est définie par :

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

3.3.Paramètres d'une distribution continue

- moyenne : $E(X) = \int_{-\infty}^{+\infty} x.f(x)dx$
- variance : $V(X) = \int_{-\infty}^{+\infty} x^2.f(x)dx - (E(X))^2$
- écart type : $\sigma(X) = \sqrt{V(X)}$

On pourrait démontrer (mais nous ne le ferons pas ici) que les formules données ci-dessus ne sont rien d'autres que des généralisations « naturelles » de celles rencontrées dans le cas discret, ce que l'on peut montrer en travaillant de la même manière que dans le chapitre du calcul intégral (p.ex. calcul d'aire sous la courbe comme limite d'une somme d'aires de rectangles ; où l'on était de même ainsi passé du « discret » au « continu » en considérant des sommes d'un nombre infini de termes).

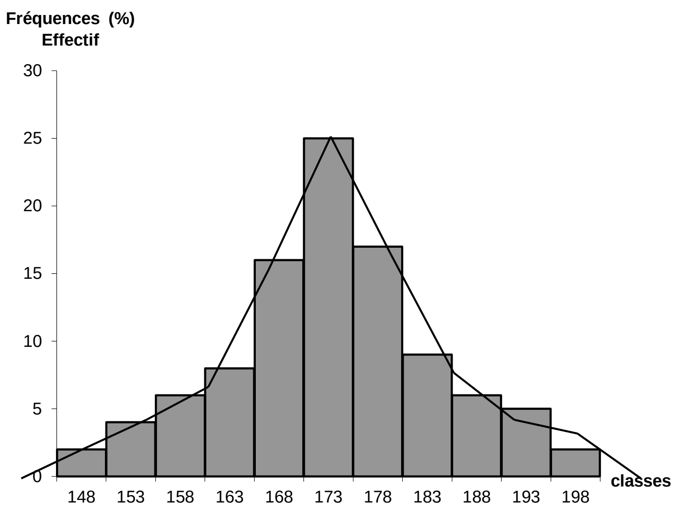
3.4 Cas particulier 2 : La variable normale (continue)

3.4.1. Exemple introductif

Une étude statistique sur les tailles des garçons de 15 à 18 ans (représentés par un échantillon de 100 garçons) a donné les résultats suivants :

Moyenne : $\mu = 173{,}35\,cm$
Écart-type : $\sigma = 10{,}64\,cm$

L'histogramme des fréquences et le polygone des fréquences se présentent comme suit :
Le polygone des fréquences ressemble à une courbe en cloche, presque symétrique par rapport à la droite

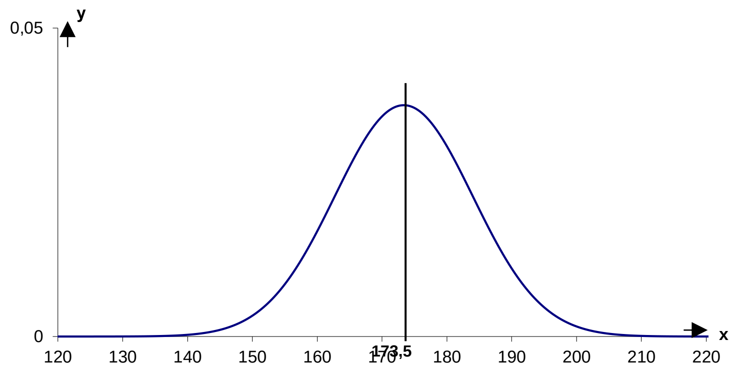


d'équation : $x = 173$, c'est à dire presque la valeur de la moyenne de l'échantillon observé.

Si l'on augmente de plus en plus l'effectif total de l'échantillon, alors il devient proche de celui de la population entière. Dans ce cas là, le nombre de tailles observées étant élevé, on peut les répartir dans des classes dont l'amplitude est de plus en plus petite, le nombre de classes devenant de plus en plus grand.
A la limite, l'amplitude de chaque classe se réduit à un point de l'axe.
Le polygone des fréquences devient une courbe en cloche, appelée <u>courbe normale</u> ou <u>courbe de Gauss</u>, parfaitement symétrique par rapport à la moyenne.

Cette courbe sert de modèle mathématique pour la plupart des distributions de variables continues.

En augmentant la taille de l'échantillon dans l'exemple précédent, le polygone des fréquences donne la courbe de Gauss suivante :



3.4.2. Distribution normale (ou loi normale)

Une variable aléatoire continue est appelée **variable normale de moyenne μ et d'écart type** σ si sa fonction de densité est

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

où e = 2,71828… et π = 3,14159…

La loi ainsi définie est appelée **loi normale** et est notée $\underline{N(\mu,\,\sigma)}$.

Dans cette expression de la densité de probabilité, x représente une valeur de la variable observée, y la fréquence de la valeur observée x, μ la moyenne de la distribution de probabilité et σ l'écart type.

La représentation graphique de cette fonction, quels que soient μ et σ, a toujours la forme caractéristique d'une cloche (voir paragraphe précédent). Rappelons cependant que :

- La courbe est symétrique par rapport à $x = \mu$;
- Elle admet un maximum en μ ;
- Elle admet Ox comme asymptote horizontale ;
- Elle a des points d'inflexion en $\mu - \sigma$ et en $\mu + \sigma$;
- Mode, moyenne et médiane coïncident.

3.4.3. Distribution normale centrée réduite

Parmi les lois normales, il en est une qui joue un rôle particulier : la loi normale centrée réduite pour laquelle $\mu = 0$ et $\sigma = 1$, notée $N(0,1)$.

Toute loi normale $N(\mu,\,\sigma)$ peut être ramenée à une loi normale centrée réduite par le changement suivant : $\frac{X-\mu}{\sigma} = Z$.

Pour obtenir la distribution de probabilité de la loi normale centrée réduite, on pose $t = \frac{x-\mu}{\sigma}$, $dt = \frac{dx}{\sigma}$.

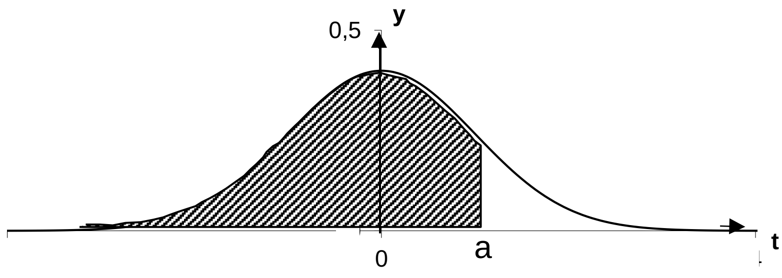
La distribution de probabilité devient : $y(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$.

Cette fonction est définie sur $]-\infty,\,+\infty[$, elle est positive, en forme de "cloche" et symétrique par rapport à l'axe Oy. On peut démontrer que $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}.e^{\frac{-1}{2}t^2}\,dt = 1$, c'est à dire que l'aire comprise entre la courbe et l'axe des abscisses vaut 1, ce qui doit être le cas pour s'accorder avec les définitions précédentes de la fonction de répartition.

$P(Z < a)$ représente l'aire de la surface comprise entre la courbe, l'axe des abscisses et la droite d'équation $t = a$.

On a :

$$\int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}}.e^{\frac{-1}{2}t^2}\,dt = P(Z \le a)$$



Remarque :
En pratique, il existe des tables (cf ci-dessous) qui donnent l'aire de la surface comprise entre la courbe normale centrée réduite, l'axe des abscisses et la droite d'équation $t = a$, c'est à dire la probabilité de trouver une observation inférieure à la valeur *a* ; il s'agit graphiquement de la valeur de l'ordonnée du point d'abscisse *a* de la courbe en S qui correspond au graphe de la fonction <u>primitive</u> de la fonction exponentielle gaussienne.

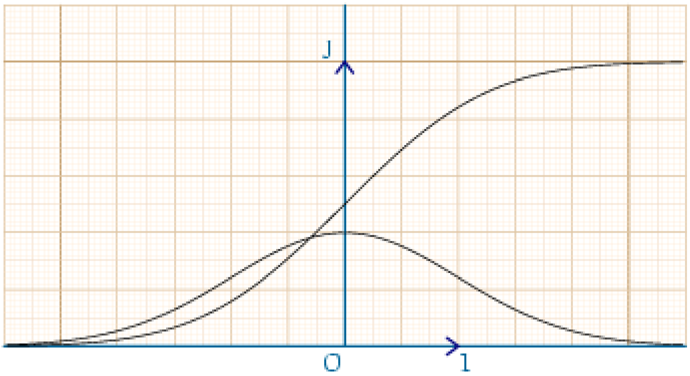


TABLE DE LA FONCTION DE REPARTITION DE LA LOI NORMALE CENTREE REDUITE

(Aire sous la courbe de la Loi Normale Centrée Réduite)

Dans ce tableau figurent les valeurs de $P(X \leq t)$ pour t ≥ 0 .

N.B. : ! Tenir compte de l'approximation; avec une précision de 5 décimales au lieu de 4, on aurait
0,99995 pour t = 3,9 et 0,99997 pour t = 4.

Intégrale $\Pi(t)$ de la Loi Normale Centrée Réduite $\mathcal{N}(0;\,1)$.

$$\Pi(t) = P(X \leq t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx \quad \text{et} \quad \Pi(-t) = 1 - \Pi(t).$$

t	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Exemple

Considérons que la taille des individus est une variable aléatoire X qui suit une loi normale de moyenne $\mu = 173{,}35$ et d'écart-type $\sigma = 10{,}64$: $X \approx N(173{,}35\ ;\ 10{,}64)$.

Quelle est la probabilité qu'un individu mesure moins de 180 cm, moins de 170 cm, plus de 180 cm, entre 170 et 180 cm ?

- $P(X < 180) = P(Z < \frac{180 - 173{,}35}{10{,}64})$ = *P(Z < 0,63)* = 0,7357 = 73, 57 %

73,57 % des individus mesurent moins de 180 cm.

- $P(X < 170) = P(Z < \frac{170 - 173{,}35}{10{,}64}) = P(Z < -0{,}31)$ = *P(Z > 0,31) = 1 – P(Z < 0,31) =*

  1 – 0,6217 = 0,3783 = 37, 83 %

37,83 % des individus mesurent moins de 170 cm.

- $P(X > 180) = P(Z > \frac{180 - 173{,}35}{10{,}64}) = P(Z > 0{,}63)$ = 1 – *P(Z <* 0,63) = 1 – 0,7357 = 0,2643 = 26,43%

26,43 % des individus mesurent plus de 180 cm.

- $$P(170 < X < 180) = P(\frac{170 - 173{,}35}{10{,}64} < Z < \frac{180 - 173{,}35}{10{,}64}) = P(-0{,}31 < Z < 0{,}63)$$
  $$= P(Z < 0{,}63) - P(Z < -0{,}31) = P(Z < 0{,}63) - P(Z > 0{,}31)$$
  $$= P(Z < 0{,}63) - [1 - P(Z < 0{,}31)] = P(Z < 0{,}63) - 1 + P(Z < 0{,}31)$$
  $$= 0{,}7356 - 1 + 0{,}6217 = 0{,}3574 = 35{,}74\%$$

35,74 % des individus mesurent entre 170 et 180 cm.

Inversement, on peut envisager d'estimer la taille que dépasse 10% de la population, ce qui revient à déterminer la taille maximale atteinte par 90% de cette même population .

Dans ce cas, il faut regarder dans la table à quelle valeur de la variable réduite correspond 0,9, et au besoin, arrondir proportionnellement à la différence des valeurs données dans la table si la valeur 0,9 n'apparaît pas précisément (ce qui est le cas).

On a : 0,8997 correspond à t = 1,28 et 0,9015 correspond à t = 1,29 . La valeur de t correspondant à 0,9 devra être estimée proportionnellement , c'est-à-dire qu'il faut ajouter à 1,28 la quantité 0,01. 3/18 (plutôt 0,01. 0,0003/0,0018)

Nous obtenons : $t \cong 1{,}28166...$ or : $t = \frac{X - \mu}{\sigma} = \frac{X - 173{,}35}{10{,}64} = 1{,}28166...$

Donc on aura : $X \cong 1{,}28166.10{,}64 + 173{,}35 = 186{,}98 \cong$ ……… c'est la taille que dépasse 10% de la population concernée.

Note

C'est au dix-huitième siècle que les scientifiques remarquèrent une étonnante régularité dans les erreurs de mesure en laboratoire ; ils étudièrent cette régularité et appelèrent la distribution obtenue « Distribution normale ». Elle est aussi appelée « Distribution de Gauss » car c'est ce mathématicien qui l'a découverte.

f) <u>***Cas particulier 3***</u> : **La loi de Poisson**

La loi de Poisson décrit une variable aléatoire discrète qui prend des valeurs dans les entiers positifs:

$$P_X(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \qquad \textit{x=1,2,3,...}$$

Cette loi intervient dans des processus aléatoires dont les **éventualités sont faiblement probables** et survenant **indépendamment les unes des autres** : cas des phénomènes accidentels, d'anomalies diverses, des problèmes d'encombrement ("files d'attente"), des ruptures de stocks, etc.

Mais d'où vient cette exponentielle $e^{-\lambda}$ *et cette factorielle k ?*

En fait, la loi de Poisson est un <u>cas limite de la loi binomiale</u> (quand $n \to \infty$ et $p \to 0$) pour lequel intervient le facteur *temps* :

Considérons un phénomène aléatoire A dont il apparaît que la probabilité d'apparition sur un laps de temps infinitésimal (très petit) Δt est proportionnelle à ce Δt, non reproductible (A apparaît au plus une fois) et indépendante de la période considérée; le nombre d'apparitions de A dans l'unité de temps choisie est une variable aléatoire discrète X, dite *de Poisson.*

Posons n.Δt = 1 : par hypothèse, dans le laps de temps Δt = 1/n, A se produira au plus une fois. Le nombre k d'apparitions de A dans l'unité de temps (n sous intervalles de durée Δt) fournit la valeur de la variable aléatoire X qui apparaît comme variable binomiale, somme de n <u>variables de Bernoulli</u> identiques dont la probabilité de réalisation est λ/n où λ est le coefficient de proportionnalité issu de notre hypothèse. On a :

$$\mathrm{Prob}(X=k) = C_n^k p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} \times \frac{n(n-1)...(n-k+1)}{n^k} \times \left(1-\frac{\lambda}{n}\right)^n \times \left(1-\frac{\lambda}{n}\right)^{-k}$$

Le dernier membre est un produit de quatre facteurs : *lorsque n tend vers l'infini,* le premier est constant (indépendant de n), le second tend vers 1, et le quatrième tend vers 1.

En se référant au cours d'analyse, on peut voir que <u>le troisième facteur tend vers $e^{-\lambda}$</u> :

En effet, $\lim\limits_{n\to\infty} \ln\left(1-\frac{\lambda}{n}\right)^n = \lim\limits_{n\to\infty} n.\ln\left(1-\frac{\lambda}{n}\right) = \lim\limits_{n\to\infty} \frac{\ln\left(1-\frac{\lambda}{n}\right)}{\frac{1}{n}}$

Et par la règle de L'Hospital , en dérivant numérateur et dénominateur: $= \lim\limits_{n\to\infty} \frac{-\lambda}{1-\frac{\lambda}{n}} = -\lambda$

On en conclut que $\lim\limits_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$

d'où : <u>Prob(X = k) = $e^{-\lambda}\lambda^k/k!$</u>

Sachant (depuis Euler) que $e^{\lambda} = \sum \frac{\lambda^k}{k!}$; on montrera que l'espérance E(X) de la loi de Poisson de paramètre λ est λ lui-même, puis que $E(X^2) = \lambda + \lambda^2$, donc que la variance est aussi égale à λ puisque $V(X) = E(X^2) - [E(X)]^2$.

Le développement ci-dessus montre que la loi de Poisson peut être utilisée en tant qu'approximation d'une loi binomiale B(n,p) lorsque n est "grand" et p "petit" avec λ = np.

La constance de λ = np, contrainte posée pour l'approximation peut s'interpréter ainsi : plus n est grand, plus la probabilité d'apparition du phénomène est faible, sa moyenne restant la même ; elle s'explique par l'hypothèse de proportionnalité de la probabilité d'apparition du phénomène au laps de temps considéré.

Rem. : On peut estimer une bonne approximation avec des valeurs de l'ordre de : n > 50, p < 0,1 et np < 10 (un λ= np "trop grand" risque d'écraser les probabilités et donner des valeurs non significatives car $e^{-\lambda}$ sera "très petit". Cependant un n "très grand" (> 1000) et un p "très petit" (quelques millièmes ou centièmes) fourniront généralement une très bonne approximation.

Exemple : *Suite à une vaccination contre le paludisme, dans une population à risque, on estime à 2%, compte tenu du délai d'immunisation, la proportion de personnes qui seront pourtant atteintes de la maladie. Quelle est la probabilité de constater, lors d'un contrôle dans un petit village de 100 habitants tous récemment vaccinés, plus d'une personne malade ? (on supposera l'indépendance des éventualités).*

Compte tenu des hypothèses, le nombre de malades est ici régi par une loi binomiale de paramètres n = 100 et p = 0,02. On a np = 2 et les conditions d'approximation par une loi de Poisson sont réalisées. Soit m la probabilité cherchée; avec les notations ci-dessus, on a :

$Prob(X = k) = e^{-2} \times 2^k/k!$, donc 1 - m @ Prob(X = 0) + Prob(X = 1) = 0,406 , soit : m @ 0,6

L'application (peu pratique) de la loi binomiale aurait fourni $1 - m = (0,98)^{100} + 2 \times (0,98)^{99}$ @ 0,403. Soit m @ 0,597.

L'approximation est donc ici excellente.

Exercice :
Pour effectuer un contrôle de qualité, on procède au tirage au hasard de 20 objets dans une fabrication industrielle. Supposons savoir que 5% des objets de cette fabrication sont défectueux.
Le nombre **N** d'objets défectueux dans l'échantillon prélevé suit alors une loi binomiale de paramètres *n=20* et *p=0.05*.
Comparer pour des valeurs de *x* entre *0* et *4* la qualité de l'approximation de la loi binomiale par une loi de Poisson de paramètre $\lambda = np = 1$.