Chapitre 1

Variables aléatoires

Dans l'édition 2004 du petit Robert, ou trouve sous «le hasard» la définition suivante.

> «Cause fictive de ce qui arrive sans raison apparente ou explicable, souvent personnifiée au même titre que le sort, la fortune, etc. [...] Se dit à l'occasion d'un heureux concours de circonstances.»

Dans ce domaine que sont les statistiques, on étudie la structure qui se cache derrière quelques processus aléatoires. On y verra que lorsqu'on répète un tel processus un grand nombre de fois, «le hasard» s'efface pour laisser place à des lois statistiques.

1.1 Variables aléatoires discrètes

1.1.1 Jets d'une pièce bien équilibrée

Considérons une pièce de monnaie parfaitement bien équilibrée. Théoriquement, une telle pièce tombe une fois sur deux sur pile, et une fois sur deux sur face. En effet, par les propriétés de symétrie et d'homogénéité de la pièce, chacun de ses côtés à la même probabilité d'apparaître.

Il y a des pièces truquées dont les deux côtés sont les mêmes (deux côtés pile par exemple). On peut aussi truquer une pièce à l'aide d'une presse[1].

Effectuons 100 lancers de pièces en notant F *pour face et* P *pour pile*. On va étudier l'évolution de la proportion des piles sur le nombre de lancers.

Les 100 lancers sont	P	P	F	F	P	P	P	F	P	P	P	$\boldsymbol{F}$	$\boldsymbol{F}$	$\boldsymbol{F}$	$\boldsymbol{F}$	$\boldsymbol{F}$	$\boldsymbol{F}$	$\boldsymbol{F}$	P	P
	P	F	P	F	F	F	P	P	F	F	P	F	F	F	P	F	P	P	P	F
	F	P	F	F	P	P	P	F	F	F	$\boldsymbol{P}$	$\boldsymbol{P}$	$\boldsymbol{P}$	$\boldsymbol{P}$	$\boldsymbol{P}$	$\boldsymbol{P}$	$\boldsymbol{P}$	F	P	P
	P	P	F	F	P	F	P	P	P	P	P	F	P	F	F	F	P	P	F	F
	F	P	P	P	F	F	F	P	P	P	F	F	P	P	P	F	F	P	F	P

Comme le côté pile est sorti 54 fois, sa fréquence d'apparition est de 0.54.

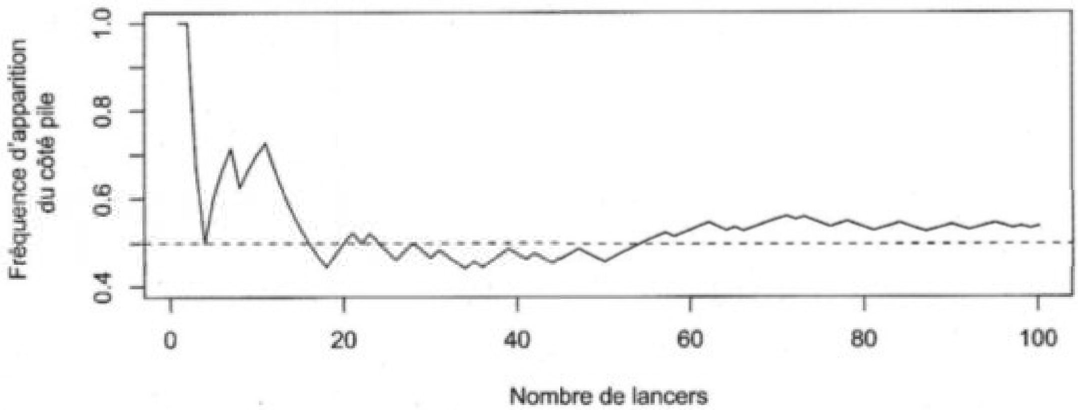
On remarque aussi qu'il y a une série de 7 piles de suite, et une série de 7 faces de suite. Si on devait demander à un humain d'écrire 100 piles ou faces de suite, il hésiterait fortement à en mettre 7 de suite et pourtant cela arrive fréquemment[2].

---

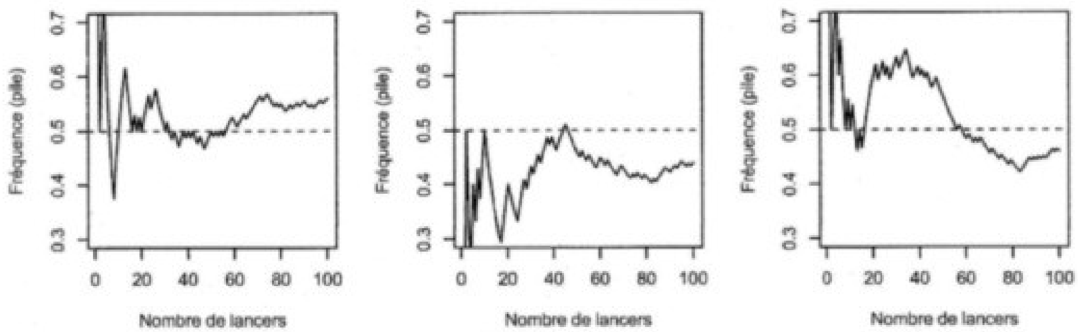
1. `http://www.stat.columbia.edu/~cook/movabletype/archives/2008/01/a_sighting_of_t.html`
2. Nicole Vogel montre, dans l'article «Peut-on imiter le hasard ?» de la brochure APMEP 451, que sur une série de 100 lancers, il y a plus d'une chance sur deux d'avoir une série d'au moins 7 piles ou 7 faces. Disponible sur `http://www.apmep.asso.fr/IMG/pdf/vogel-451.pdf`.

Voici la représentation graphique de l'évolution de la fréquence d'apparition du côté pile.

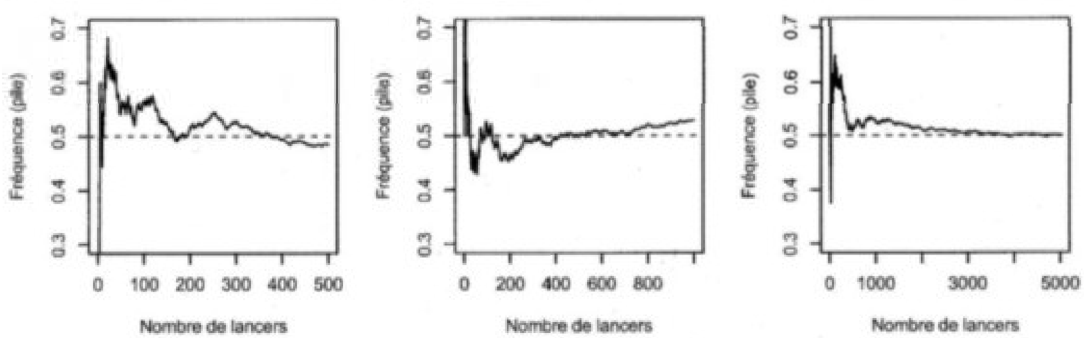
On commence à 1, puisque le premier lancer a donné pile. Le deuxième lancer donne aussi pile, la fréquence reste à $\frac{2}{2} = 1$. Le troisième lancer donne F, ainsi la fréquence donne $\frac{2}{3} \cong 0.67$. Le quatrième lancer donne F, ainsi la fréquence donne $\frac{2}{4} \cong 0.5$. Là le graphe touche l'horizontale à hauteur $\frac{1}{2} = 0.5$. Rappelons qu'au 100-ième lancer, la fréquence valait 0.54.



Effectuons trois autres séries de 100 lancers pour sentir ce qui se passe.



Pour les trois graphes ci-dessous, on a d'abord lancé 500 pièces, puis 1000, puis 5000.



On voit que plus le nombre de lancers augmente, plus la fréquence s'approche de 0.5. Ce phénomène illustre le résultat fondamental suivant de la théorie des probabilités.

La loi des grands nombres

Lorsqu'on répète une expérience aléatoire de manière indépendante un grand nombre de fois, alors les proportions expérimentales se rapprochent des proportions théoriques, appelée *probabilités*.

1.1.2 Jets d'un dé à six faces bien équilibré

Considérons un dé à six faces bien équilibré. Théoriquement, la fréquence d'apparition de chaque face est de $\frac{1}{6}$. En effet, par les propriétés de symétrie et d'homogénéité du cube, chacune de ses faces à la même probabilité d'apparaître.

On trouve sur le marché des dés pipés qui ont un numéro qui sort dans le 80% des cas[3].

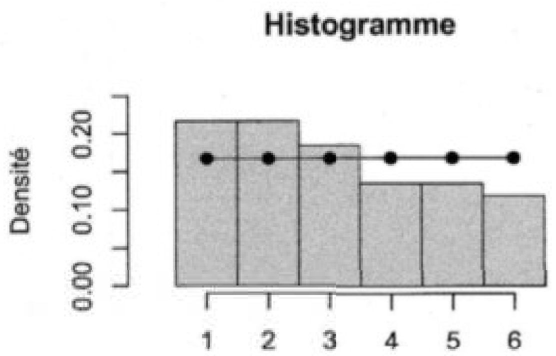
En effectuant 60 lancers, on obtient les données suivantes.

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

On construit, avec ces données, un *histogramme*. Pour chacune des issues possibles (ici les nombres de 1 à 6), on compte le nombre de fois qu'elle apparaît, c'est le nombre d'*occurrences*, puis on calcule la *densité* qui est le nombre d'occurrences divisé par le nombre total de jets.

face	1	2	3	4	5	6
nombre d'occurrences	13	13	11	8	8	7
densité	0.216	0.216	0.183	0.133	0.133	0.116

On fait ensuite un histogramme (voir le complément sur les histogrammes à la fin de cette sous-section en page 7).



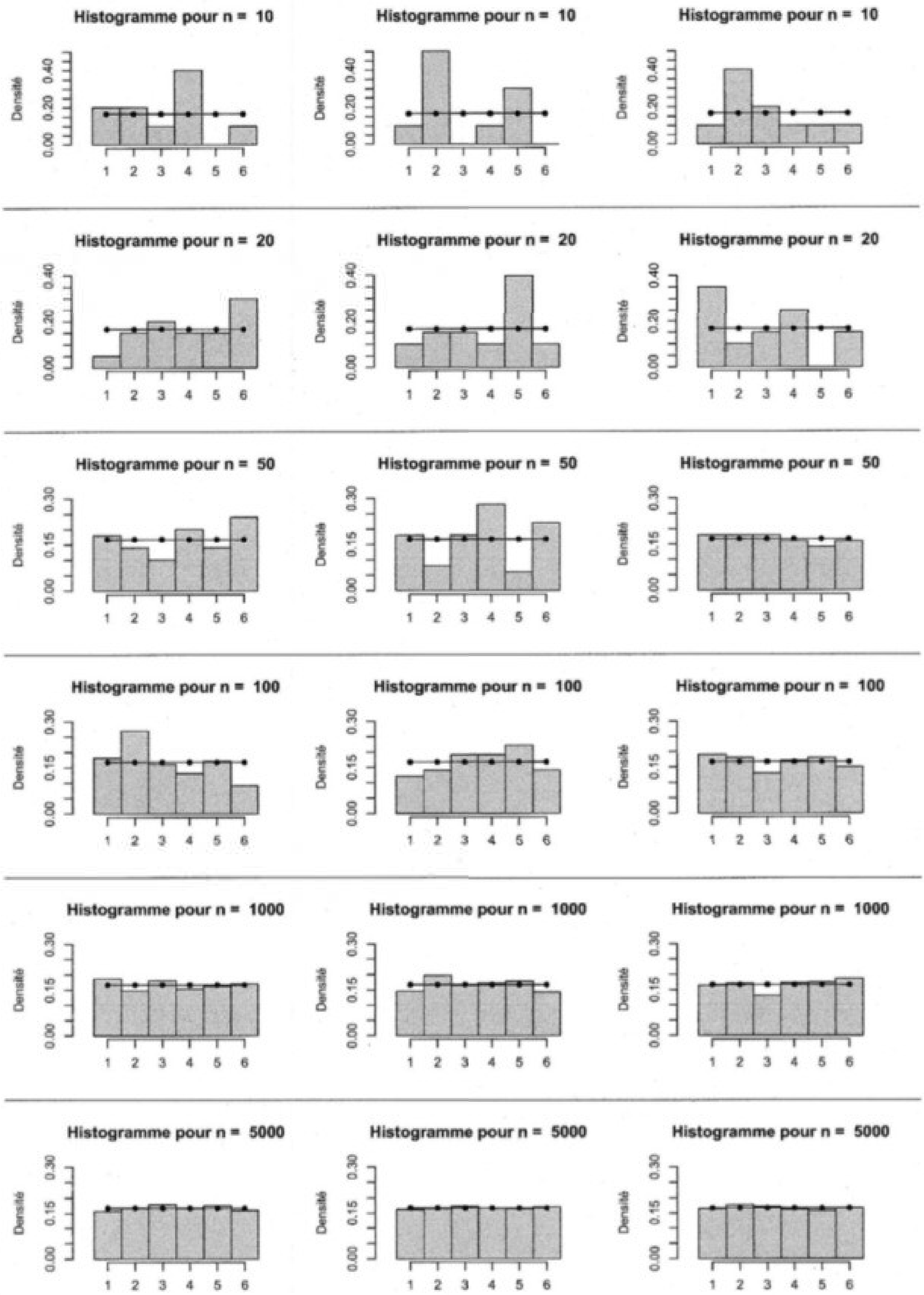
L'histogramme permet de mieux visualiser les données. Ici, on voit que les faces 1, 2 et 3 sont sorties bien plus souvent. Il est possible que le dé soit pipé (avec une petite bille de plomb vers le sommet adjacent aux faces 4, 5 et 6).

Néanmoins, ces données ne permettent pas de déduire que le dé est truqué. D'ailleurs un test du chi-carré indique que ce dé n'est probablement pas truqué (voir en page 52). Le test ne se trompe pas puisque ces données sont issues de simulations à partir d'un dé (informatique) non pipé.

La loi des grands nombres affirme que plus le nombre de lancers est grand plus les densités se rapprochent des probabilités (ou des densités théoriques) qui valent ici $\frac{1}{6} \cong 0.167$. Ces probabilités sont indiquées sur l'histogramme ci-dessus par les points reliés.

3. `http://www.toursdemagie.com/produit/magie/6-Des-Pipes---8-Des-truques.html`

En augmentant le nombre n de lancers effectués, on vérifie empiriquement la loi des grands nombres.



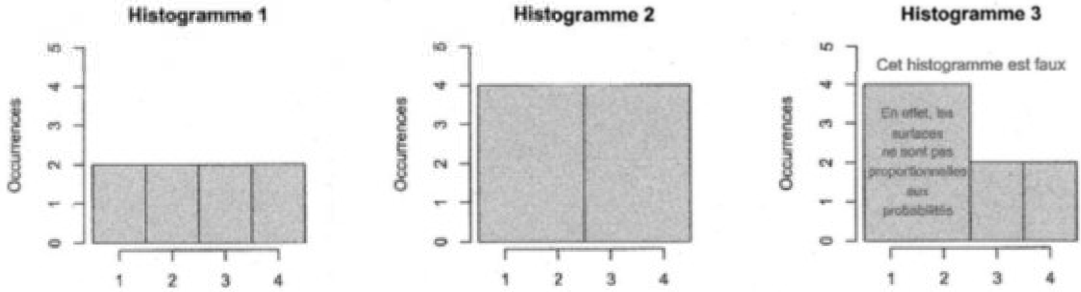
Compléments sur les histogrammes

Il y a deux sortes d'histogrammes : ceux qui ont en ordonnée les occurrences (c'est-à-dire le nombre de fois qu'une valeur mesurée apparaît ; dans l'exemple, il s'agissait du nombre montré par le dé) et ceux qui ont en ordonnée les densités.

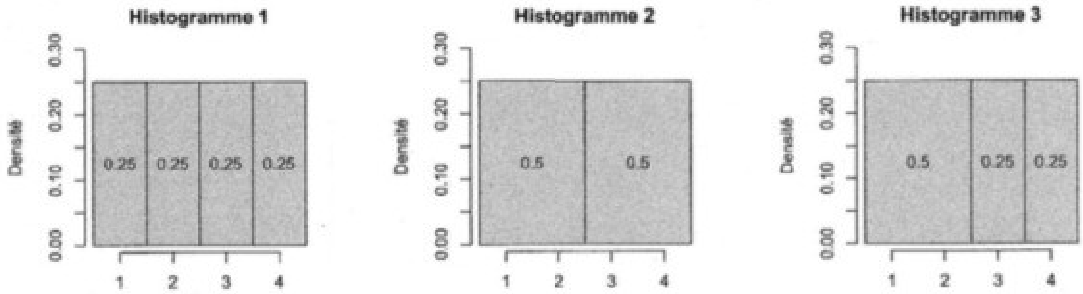
Imaginons que l'on lance un dé à quatre faces et que l'on obtienne les résultats suivants.

face	1	2	3	4
nombre d'occurrences	2	2	2	2
densité	0.25	0.25	0.25	0.25

Voici trois histogrammes des occurrences avec différentes façons de grouper les résultats.



Et ceux pour les densités (avec la même façon de grouper les résultats).



On voit le lien entre ces trois histogrammes et la densité de probabilité : les aires doivent correspondre aux probabilités (voir page 15).

Ces histogrammes s'obtiennent à l'aide du logiciel «R» (`http://www.r-project.org`) qui est la version gratuite du célèbre logiciel «S plus» (`http://stat.bell-labs.com/S`).

Voici ce qu'il faut taper dans «R» pour obtenir ces histogrammes (pour raccourcir le code, certains réglages ont été enlevés).

```
data <- c(1,1,2,2,3,3,4,4); layout(matrix(1:3, 1, 3, byrow=TRUE))

hist(data,breaks=c(0.5,1.5,2.5,3.5,4.5),prob=FALSE,col="gray",
     main=paste("Histogramme 1"),ylim=c(0,5),xlab="",ylab="Occurrences")

hist(data,breaks=c(0.5,2.5,4.5),prob=FALSE,col="gray",
     main=paste("Histogramme 2"),ylim=c(0,5),xlab="",ylab="Occurrences")

hist(data,breaks=c(0.5,2.5,3.5,4.5),prob=FALSE,col="gray",
     main=paste("Histogramme 3"),ylim=c(0,5),xlab="",ylab="Occurrences")
```

En mettant `prob=TRUE` au lieu de `prob=FALSE`, «R» passe directement des histogrammes des occurrences aux histogrammes des densités.

1.1.3 Variables aléatoires discrètes, espérance et variance

Lorsque les valeurs prises par la variable aléatoire X sont en nombre fini ou dénombrable, on dit que la variable aléatoire est *discrète*.

Définitions

L'ensemble des issues d'une expérience aléatoire est appelé *Univers* et noté Ω.

Une *variable aléatoire*[4] est une fonction $X : \Omega \to \mathbb{R}$ qui associe à chaque issue de l'expérience en question un unique nombre réel.

On note $\mathbf{P}(X = x)$ pour la probabilité[5] que la variable X prenne la valeur x (attention à bien distinguer majuscule (variable aléatoire) et minuscule (valeur prise par cette variable aléatoire)).

On définit l'*espérance* de X par $\mu = E(X)$ et la variance de X par $\sigma^2 = V(X)$ où

$$\boxed{\mu = E(X) = \sum_{\omega\in\Omega} X(\omega)\cdot\mathbf{P}(\omega) = \sum_{x\in\mathbb{R}} x\cdot\mathbf{P}(X = x)}$$

On définit la *variance* de X par $\sigma^2 = V(X)$ où

$$\boxed{\sigma^2 = V(X) = \sum_{x\in\mathbb{R}} (x-\mu)^2\cdot\mathbf{P}(X = x)}$$

On définit aussi l'*écart type* de X par $\sigma = \sqrt{\sigma^2}$.

Ces formules ne sont valables que lorsque $\mathbf{P}(X = x) \neq 0$ pour un nombre fini ou dénombrable[6] de $x \in \mathbb{R}$ (ce qui est le cas pour une variable aléatoire discrète). C'est pour cette raison qu'on se permet de noter $\sum_{x\in\mathbb{R}}$.

Interprétations

1. L'espérance de X est la moyenne des issues pondérée par leurs probabilités.

2. La variance de X est la moyenne des carrés des écarts par rapport à l'espérance pondérée par les probabilités. Autrement dit, plus les valeurs prises par X s'éloignent de l'espérance, plus la variance augmente fortement. Quant aux effets des petits écarts sur la variance, ils sont presque négligeables.

Ces notions sont liées à la variable aléatoire, donc à la théorie des probabilités. Elles ne sont pas liées aux mesures obtenues lors d'expérimentions. On verra qu'on pourra estimer ces paramètres (espérance, variance et écart type) à l'aide de mesures dans le chapitre 2.

Subtilité ! Contrairement à d'habitude, si X est une variable aléatoire, x est une valeur prise par cette variable aléatoire. Donc x est dans le domaine arrivée de X, qui est $\mathbb{R}$.

---

4. On comprendra mieux dans les chapitres qui suivent pourquoi on les appelle *variables aléatoires* alors que si on regarde la définition, il n'y a rien de variable, ni d'aléatoire.

5. Par rapport aux notions du cours DF, la probabilité que la variable aléatoire atteigne la valeur x est donnée par $\mathbf{P}(X = x) = \mathbf{P}(\{\omega \in \Omega : X(\omega) = x\})$.

6. dans ce cas, les séries sont infinies.

Retour à l'exemple du jet d'un dé

En reprenant l'exemple précédent. On peut considérer la variable aléatoire X qui représente le nombre montré par le jet d'un dé.

On peut ainsi avoir une notation efficace pour les probabilités : la probabilité que le nombre montré par le dé soit entre 4 et 6 se note $\mathbf{P}(4 \leqslant X \leqslant 6)$. Puisque la probabilité d'avoir chaque nombre vaut $\frac{1}{6}$, on a

$$\mathbf{P}(4 \leqslant X \leqslant 6) = \mathbf{P}(X=4) + \mathbf{P}(X=5) + \mathbf{P}(X=6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = 50\%$$

Le calcul de l'espérance est le suivant.

$$\begin{aligned} E(X) &= \mathbf{P}(X=1) + 2\mathbf{P}(X=2) + 3\mathbf{P}(X=3) + 4\mathbf{P}(X=4) + 5\mathbf{P}(X=5) + 6\mathbf{P}(X=6) \\ &= \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = \frac{7}{2} = 3.5 \end{aligned}$$

On vérifie que l'espérance n'est rien d'autre que la moyenne des résultats possibles.

$$E(X) = \frac{1+2+3+4+5+6}{6}$$

Le calcul de la variance est le suivant.

$$\begin{aligned} V(X) &= (1-\mu)^2 \cdot \mathbf{P}(X=1) + (2-\mu)^2 \cdot \mathbf{P}(X=2) + (3-\mu)^2 \cdot \mathbf{P}(X=3) \\ &\quad + (4-\mu)^2 \cdot \mathbf{P}(X=4) + (5-\mu)^2 \cdot \mathbf{P}(X=5) + (6-\mu)^2 \cdot \mathbf{P}(X=6) \\ &= \frac{(1-\frac{7}{2})^2}{6} + \frac{(2-\frac{7}{2})^2}{6} + \frac{(3-\frac{7}{2})^2}{6} + \frac{(4-\frac{7}{2})^2}{6} + \frac{(5-\frac{7}{2})^2}{6} + \frac{(6-\frac{7}{2})^2}{6} \\ &= \frac{\frac{25}{4}}{6} + \frac{\frac{9}{4}}{6} + \frac{\frac{1}{4}}{6} + \frac{\frac{1}{4}}{6} + \frac{\frac{9}{4}}{6} + \frac{\frac{25}{4}}{6} = \frac{70}{24} = \frac{35}{12} = 2.91\overline{6} \end{aligned}$$

L'écart type vaut $\sqrt{\frac{35}{12}} \cong 1.71$.

L'espérance de gain

Quand la variable aléatoire donne des valeurs en CHF, on parle d'*espérance de gain* au lieu d'espérance (mais cela ne change rien à sa définition).

Dans ce cas, l'unité de l'espérance est le CHF. Quant à l'unité de la variance, il s'agit de CHF^2 ! Ainsi, l'unité de l'écart type est le CHF. Il est donc plus naturel de parler de l'écart type lorsqu'il y a une unité.

Par exemple, imaginons que deux personnes Anka et Brad jouent à pile ou face. Si la pièce tombe sur pile, alors Anka donne 2 CHF à Brad ; si la pièce tombe sur face, alors c'est Brad qui donne 4 CHF à Anka. On peut traduire cette règle par une variable aléatoire $X : \{\text{pile}, \text{face}\} \to \mathbb{R}$ qui gère la fortune de Anka : $X(\text{pile}) = -2$ CHF et $X(\text{face}) = 4$ CHF.

Pour cette variable X, l'espérance vaut $E(X) = -2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{2} = 1$ CHF. La variance vaut $V(X) = (-2-1)^2 \cdot \frac{1}{2} + (4-1)^2 \cdot \frac{1}{2} = 9\ \text{CHF}^2$. L'écart type vaut $\sigma = 3$ CHF.

L'espérance de gain mesure le gain moyen : en jouant 1000 fois à un tel jeu, Anka est en droit d'espérer gagner 1000 CHF (si elle n'est ni trop chanceuse, ni trop malchanceuse). L'écart type est en relation avec le risque.

Généralisation de l'espérance

Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction réelle et X une variable aléatoire, alors $g \circ X$ est une variable aléatoire notée[7] $g(X)$ dont l'espérance est :

$$E(g(X)) = \sum_{\omega \in \Omega} g(X(\omega)) \cdot \mathbf{P}(\omega) = \sum_{x \in \mathbb{R}} g(x) \cdot \mathbf{P}(X = x)$$

Cela permet notamment d'écrire la variance comme une espérance.

$$V(X) = \sum_{x \in \mathbb{R}} (x - \mu)^2 \cdot \mathbf{P}(X = x) = E((X - \mu)^2)$$

1.1.4 Le théorème de Tchebychev

Soit X une variable aléatoire d'espérance μ et d'écart type σ. Alors, on a

$$\mathbf{P}(\mu - k\sigma \leqslant X \leqslant \mu + k\sigma) \geqslant 1 - \frac{1}{k^2} \qquad \text{pour tout } k > 0$$

Cas particuliers

1. Pour $k = 1$, on a $\mathbf{P}(\mu - \sigma \leqslant X \leqslant \mu + \sigma) \geqslant 0\%$.

   Cette affirmation n'est pas particulièrement utile.

2. Pour $k = 2$, on a $\mathbf{P}(\mu - 2\sigma \leqslant X \leqslant \mu + 2\sigma) \geqslant 75\%$.

   Ainsi il y a au moins 75% des valeurs de X qui sont dans l'intervalle $[\mu - 2\sigma, \mu + 2\sigma]$.

3. Pour $k = 3$, on a $\mathbf{P}(\mu - 3\sigma \leqslant X \leqslant \mu + 3\sigma) \geqslant 88.\overline{8}\%$.

   Ainsi il y a au moins 88% des valeurs de X qui sont dans l'intervalle $[\mu - 3\sigma, \mu + 3\sigma]$.

Illustration dans l'exemple du jeu d'argent de la page précédente

Dans cet exemple, on a $\mu = 1$ et $\sigma = 3$. On examine les gains d'Anka pour une partie.

1. La probabilité que la somme gagnée soit dans l'intervalle $[\mu - \sigma, \mu + \sigma] = [-2, 4]$ doit être plus grande ou égale à 0%.

   C'est vrai, car cette probabilité vaut 100%.

2. La probabilité que la somme gagnée soit dans l'intervalle $[\mu - 2\sigma, \mu + 2\sigma] = [-5, 7]$ doit être plus grande ou égale à 75%.

   C'est vrai, car cette probabilité vaut 100%.

3. La probabilité que la somme gagnée soit dans l'intervalle $[\mu - 3\sigma, \mu + 3\sigma] = [-8, 10]$ doit être plus grande ou égale à $88.\overline{8}\%$.

   C'est vrai, car cette probabilité vaut 100%.

Conclusion Le théorème ne donne pas des valeurs optimales, mais il est vrai quelque soit la variable aléatoire. Ici, on a pris un exemple un peu trop simple, ce qui explique la faible performance du théorème.

---

7. Cette notation est abusive : $g(X)(\omega)$ n'a pas de sens ! La justification est celle-ci : si $X : \Omega \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$, alors $(g \circ X)(\omega) = g(X(\omega)) = g(x)$ avec $x = X(\omega) \in \mathrm{Im}(X)$. D'où la notation $g \circ X = g(X)$.

1.1.5 Jets de plusieurs dés à six faces bien équilibrés

Considérons plusieurs dés à six faces bien équilibrés. On examine la somme obtenue des valeurs montrées par un jet de plusieurs dés.

Jets de deux dés

Voici un tableau représentant la somme, notée $\sum$, de chacune des $6^2 = 36$ issues.

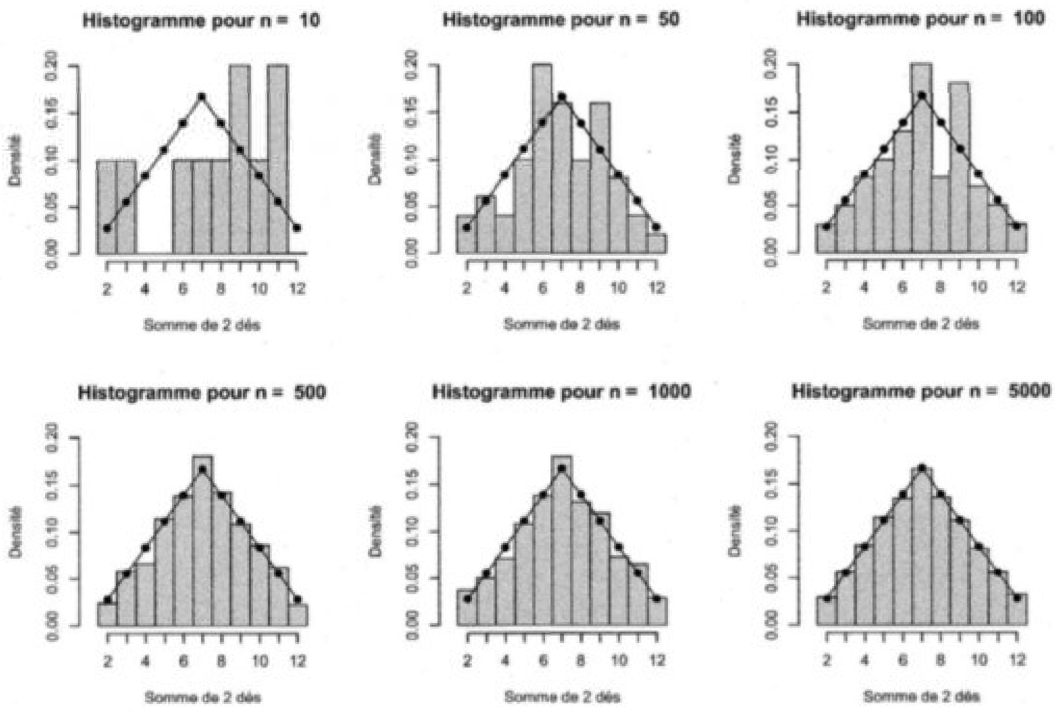
Issue	$\sum$	Issue	$\sum$	Issue	$\sum$	Issue	$\sum$	Issue	$\sum$	Issue	$\sum$
⚀⚀	2	⚀⚁	3	⚀⚂	4	⚀⚃	5	⚀⚄	6	⚀⚅	7
⚁⚀	3	⚁⚁	4	⚁⚂	5	⚁⚃	6	⚁⚄	7	⚁⚅	8
⚂⚀	4	⚂⚁	5	⚂⚂	6	⚂⚃	7	⚂⚄	8	⚂⚅	9
⚃⚀	5	⚃⚁	6	⚃⚂	7	⚃⚃	8	⚃⚄	9	⚃⚅	10
⚄⚀	6	⚄⚁	7	⚄⚂	8	⚄⚃	9	⚄⚄	10	⚄⚅	11
⚅⚀	7	⚅⚁	8	⚅⚂	9	⚅⚃	10	⚅⚄	11	⚅⚅	12

Si on suppose que les dés sont bien équilibrés, alors chaque issue a autant de chances de se produire (c'est pour cette raison qu'on a tenu compte de l'ordre). On trouve ainsi le tableau des probabilités suivant.

somme	2	3	4	5	6	7	8	9	10	11	12
probabilité	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

La loi des grands nombres affirme que plus le nombre n de lancers est grand plus les densités se rapprochent des probabilités ci-dessus. Ces probabilités sont indiquées sur les histogrammes ci-dessous par les points reliés.

Voici une vérification empirique de la loi des grands nombres.



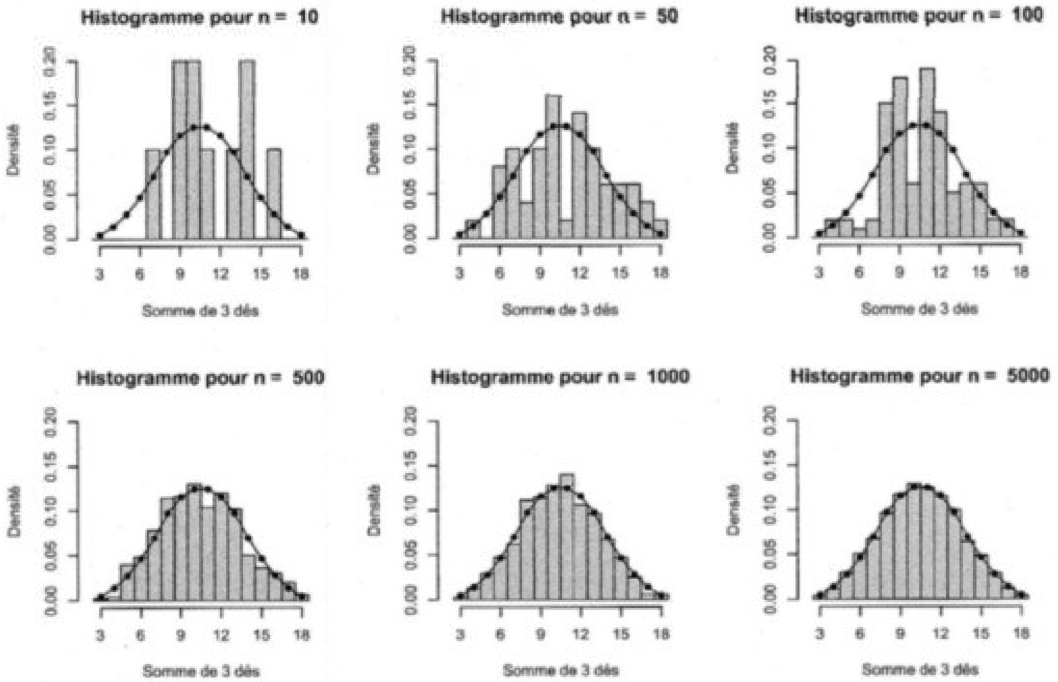
Jets de trois dés

Lorsqu'on jette 3 dés, il y a $6^3 = 216$ issues. On a le tableau de probabilités suivant.

somme	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
proba	$\frac{1}{216}$	$\frac{3}{216}$	$\frac{6}{216}$	$\frac{10}{216}$	$\frac{15}{216}$	$\frac{21}{216}$	$\frac{25}{216}$	$\frac{27}{216}$	$\frac{27}{216}$	$\frac{25}{216}$	$\frac{21}{216}$	$\frac{15}{216}$	$\frac{10}{216}$	$\frac{6}{216}$	$\frac{3}{216}$	$\frac{1}{216}$

La loi des grands nombres affirme que plus le nombre n de lancers est grand plus les densités se rapprochent des probabilités ci-dessus. Ces probabilités sont indiquées sur les histogrammes ci-dessous par les points reliés.

Voici une vérification empirique de la loi des grands nombres.



Propriétés de l'espérance et de la variance

En reprenant l'exemple précédent. On peut considérer les variables aléatoires X_1, X_2 et X_3 qui représentent le nombre montré par les dés numéro 1, 2 et 3 respectivement.

On note dans ce cas, la variable aléatoire $X = \sum_{i=1}^{3} X_i = X_1 + X_2 + X_3$ qui représente la somme des nombres montrés par les trois dés.

Grâce aux probabilités qui se trouvent dans le tableau de l'exemple précédent, on a

$$\mathbb{P}(12 \leqslant X \leqslant 18) = \frac{25}{216} + \frac{21}{216} + \frac{15}{216} + \frac{10}{216} + \frac{6}{216} + \frac{3}{216} + \frac{1}{216} = \frac{81}{216} = 0.375$$

On utilise les propriétés suivantes (voir section 1.4).

$$\boxed{E\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} E(X_i)} \quad \text{et} \quad \boxed{V\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} V(X_i) \text{ si les } X_i \text{ sont indépendantes}}$$

Ainsi

$$\begin{aligned} E(X) &= E(X_1) + E(X_2) + E(X_3) = 3 \cdot \tfrac{7}{2} = 10.5 \Longrightarrow \mu = 10.5 \\ V(X) &= V(X_1) + V(X_2) + V(X_3) = 3 \cdot \tfrac{35}{12} = 8.75 \Longrightarrow \sigma \cong 2.96 \end{aligned}$$

1.1.6 La loi de Bernoulli

On considère une expérience aléatoire à deux issues : le succès et l'échec. À cette expérience, on associe la variable aléatoire X qui donne 1 en cas de succès et 0 en cas d'échec.

$$X : \{\text{succès}, \text{échec}\} \to \mathbb{R};\quad \text{succès} \mapsto 1,\ \text{échec} \mapsto 0$$

Si $p \in]0,1[$ est la probabilité de succès, alors on a

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

On dit que la variable aléatoire X suit une loi de Bernoulli de paramètre p, notée $B(p)$.

Il est facile de montrer que $E(X) = p$ et $V(X) = p(1-p)$.

Application

On tente de lancer une pièce de monnaie bien équilibrée sur pile. L'espérance vaut $\frac{1}{2}$ (on va réussir en moyenne une fois sur 2), la variance vaut $\frac{1}{4}$, l'écart type vaut $\frac{1}{2}$.

1.1.7 La loi binomiale

On considère une expérience aléatoire constituée de n épreuves successives identiques, indépendantes, à deux issues chacune : le succès et l'échec. Autrement dit, à la la i-ième épreuve, on associe une variable aléatoire X_i qui suit une loi de Bernoulli $B(p)$ où $p \in]0,1[$ est la probabilité d'un succès à chaque épreuve successive. À cette expérience, on associe la variable aléatoire X qui compte le nombre de succès, ainsi $X = \sum_{i=1}^{n} X_i$.

La probabilité d'avoir k succès (et donc $n-k$ échecs) est donnée grâce à la technique des anagrammes[8] par la formule suivante (la i-ième case correspond au i-jet ; il y a n cases).

$$\begin{aligned}
\mathbf{P}(X=k) &= \mathbf{P}\Bigg(\overbrace{\underbrace{\boxed{\text{succès}}\boxed{\text{succès}}\boxed{\ldots}\boxed{\text{succès}}}_{k \text{ succès}}\underbrace{\boxed{\text{échec}}\boxed{\ldots}\boxed{\text{échec}}}_{n-k \text{ échecs}}}^{\text{permutations}}\Bigg) \\
&= \binom{n}{k} p^k (1-p)^{n-k} \text{où} \quad \binom{n}{k} = \left\{ \begin{array}{cl} \frac{n!}{k!(n-k)!} & \text{si } k \in \{0,1,2,\ldots,n\} \\ 0 & \text{sinon} \end{array} \right.
\end{aligned}$$

En fait[9], $\binom{n}{k}$ est le coefficient de x^k dans le polynôme $(1+x)^n$.

On dit que la variable aléatoire X suit une binomiale de paramètres n et p, notée $B(n,p)$.

Remarque importante

Une variable qui suit une loi binomiale $B(n,p)$ est une somme de n variables qui suivent une loi de Bernoulli $B(p)$. Grâce aux propriétés de l'espérance et de la variance (voir section 1.4), on a $E(X) = np$ et $V(X) = np(1-p)$.

Application

On compte le nombre de piles obtenus en lançant n fois une pièce de monnaie bien équilibrée. L'espérance vaut $\frac{n}{2}$ (le nombre de piles obtenus est en moyenne égal à la moitié des lancers), l'écart type vaut $\frac{n}{4}$.

---

8. Se référer à la technique des anagrammes du cours de probabilité

9. Par rapport au résultat de la page suivante, on a $(1+x)^n = (x^0 + x^1)^n$ car notre pièce de monnaie est comme un dé à deux faces dont les numéros sont 0 pour face et 1 pour pile.

1.1.8 La loi des jets de k dés

Cas particulier où $k = 5$

La variable aléatoire X compte la somme des résultats des jets de chacun des 5 dés.

On a évidemment $P(X = 5) = (\frac{1}{6})^5 = \frac{1}{7776}$, puisque chacun des 5 dés doit tomber sur 1. De même, on a $P(X = 30) = \frac{1}{7776}$. Il est plus dur de calculer $P(X = k)$ avec $5 < m < 30$.

Regardons en détails comment calculer $P(X = 8)$ avec la méthode des anagrammes.

$$\begin{aligned} P(X=8) &= \mathbf{P}\Big(\overbrace{\boxed{1\,|\,1\,|\,2\,|\,2\,|\,2}}^{\text{perm.}}\Big) + \mathbf{P}\Big(\overbrace{\boxed{1\,|\,1\,|\,1\,|\,2\,|\,3}}^{\text{perm.}}\Big) + \mathbf{P}\Big(\overbrace{\boxed{1\,|\,1\,|\,1\,|\,1\,|\,4}}^{\text{perm.}}\Big) \\ &= \frac{5!}{2!\,3!}\cdot\left(\frac{1}{6}\right)^5 + \frac{5!}{3!\,1!\,1!}\cdot\left(\frac{1}{6}\right)^5 + \frac{5!}{4!\,1!}\cdot\left(\frac{1}{6}\right)^5 \\ &= \left(\frac{1}{6}\right)^5\left(\frac{5!}{2!\,3!} + \frac{5!}{3!\,1!\,1!} + \frac{5!}{4!\,1!}\right) \end{aligned}$$

On voit que ce qui n'est pas facile à trouver dans le cas général, c'est le contenu de la parenthèse qui contient les fractions avec des factorielles.

Mais en fait, ces fractions peuvent être expliquées[10] ainsi.

$\frac{5!}{2!\,3!}$ est le coefficient de $x_1^2x_2^3$ dans $(x_1+x_2+x_3+x_4+x_5+x_6)^5$

$\frac{5!}{3!\,1!\,1!}$ est le coefficient de $x_1^3x_2x_3$ dans $(x_1+x_2+x_3+x_4+x_5+x_6)^5$

$\frac{5!}{4!\,1!}$ est le coefficient de $x_1^4x_4$ dans $(x_1+x_2+x_3+x_4+x_5+x_6)^5$

On utilise l'astuce suivante : en remplaçant x_1 par x, x_2 par x^2, x_3 par x^3, …, x_6 par x^6, on transforme les monômes à six variables ci-dessus en un polynôme à une seule variable. Ainsi, les trois monômes ci-dessus ($x_1^2x_2^3$, $x_1^3x_2x_3$ et $x_1^4x_4$) sont fusionnés sur un nouveau monôme à une seule variable dont la puissance correspond à la somme des dés, ici x^8.

En effet, $x_1^2x_2^3$ devient $x^2x^6 = x^8$; $x_1^3x_2x_3$ devient $x^3x^2x^3 = x^8$; $x_1^4x_4$ devient $x^4x^4 = x^8$. Ainsi

$\frac{5!}{2!\,3!} + \frac{5!}{3!\,1!\,1!} + \frac{5!}{4!\,1!}$ est le coefficient de x^8 dans $(x+x^2+x^3+x^4+x^5+x^6)^5$

Règle générale pour k dés

La probabilité de lancer k dés et d'obtenir un total de m est donnée par le coefficient x^m du polynôme $(x+x^2+x^3+x^4+x^5+x^6)^k$ multiplié par $(\frac{1}{6})^k$. Ce coefficient est facilement déterminable à l'aide de `Maxima` : la commande qui établit la liste des coefficients est `makelist(coeff(ratsimp((x+x^2+x^3+x^4+x^5+x^6)^k),x,m),m,k,6*k)`.

Par exemple, la probabilité de faire un total de 120 avec 30 dés vaut

$$2\,633\,742\,030\,711\,604\,837\,831 \cdot \frac{1}{221\,073\,919\,720\,733\,357\,899\,776} \cong 0.0119$$

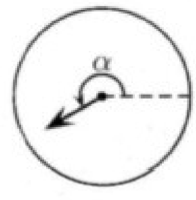
10. On utilise les combinaisons vues dans le chapitre sur le dénombrement

1.2 Variables aléatoires continues

Lorsque les valeurs prises par la variable aléatoire X varient dans un intervalle (ou une réunion d'intervalles), on dit que la variable aléatoire est *continue.*

1.2.1 La loi uniforme

Imaginons l'expérience aléatoire suivante : on fait tourner une aiguille fixée au milieu d'un disque. L'aiguille s'arrêtera de manière aléatoire.



Regardons la variable aléatoire X qui associe à chaque position de l'aiguille l'angle correspondant dans l'intervalle $[0, 2\pi[$.

La probabilité que l'angle soit exactement $\frac{\pi}{2}$ est nulle. On voit tout de même que la probabilité que l'angle soit entre $\frac{\pi}{4}$ et $\frac{3\pi}{4}$ devrait valoir $\frac{1}{4}$.

L'astuce consiste à se ramener à regarder des aires sous une courbe appelée *densité (continue) de probabilité* ou encore *distribution* et notée $f(x)$. On a

$$\underbrace{\mathbf{P}(a \leqslant X \leqslant b)}_{\substack{\text{Probabilité que } X \text{ prenne} \\ \text{une valeur entre } a \text{ et } b}} \quad = \quad \underbrace{\int_a^b f(x)\,dx}_{\substack{\text{Notation mathématique pour} \\ \text{décrire l'aire entre } f \text{ et l'axe des } x, \\ \text{de } a \text{ jusqu'à } b}}$$

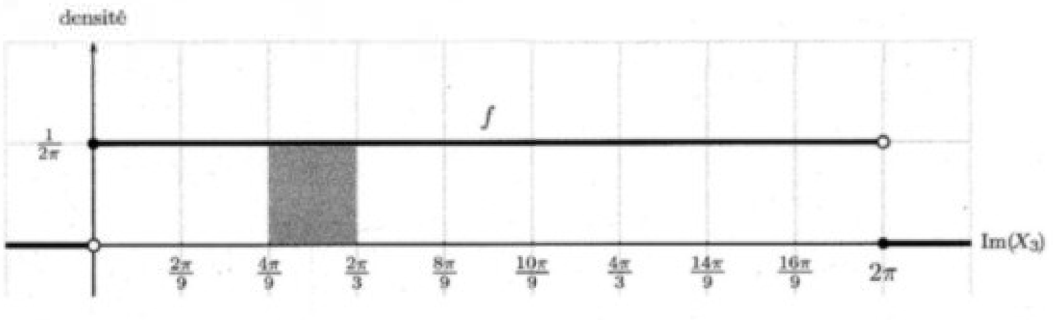
Comme une probabilité est un nombre entre 0 et 1 (ou entre 0% et 100%), on doit avoir

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1 \qquad \text{et} \qquad f(x) \geqslant 0 \text{ pour tout } x \in \mathbb{R}$$

La densité de la variable aléatoire X associée à cette expérience aléatoire est donnée par

$$f(x) = \begin{cases} \frac{1}{2\pi} & \text{si } x \in [0, 2\pi[\\ 0 & \text{si } x \notin [0, 2\pi[\end{cases}$$

Voici son graphe.



Cette densité est dite *uniforme.*

On a

$$\mathbf{P}(\tfrac{4\pi}{9} \leqslant X \leqslant \tfrac{2\pi}{3}) \; = \; \int_{\frac{4\pi}{9}}^{\frac{2\pi}{3}} f(x)dx \; = \; \frac{1}{9}$$

En effet, en observant le quadrillage, on constate que la zone grise définie sur l'intervalle $\left[\frac{4\pi}{9}, \frac{2\pi}{3}\right]$ a une proportion de $\frac{1}{9}$.

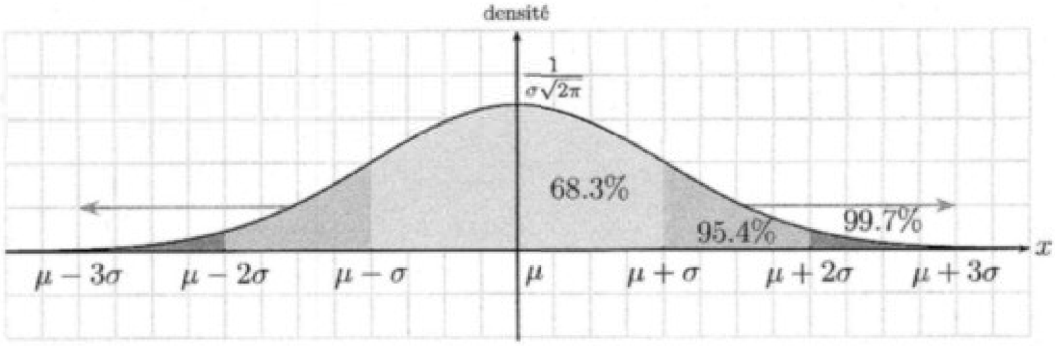
1.2.2 La loi normale

Il existe une loi, notée $\mathcal{N}(\mu, \sigma^2)$ et appelée *loi normale* (ou *loi de Laplace-Gauss*) qui, comme on le verra, permet d'approximer certaines probabilités.

Cette loi dépend de deux paramètres : l'espérance μ et la variance σ^2 (qui est le carré de l'écart type σ). Elle est donnée par la densité

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

où e est le nombre d'Euler qui vaut environ 2.71828. Voici le graphe de f.

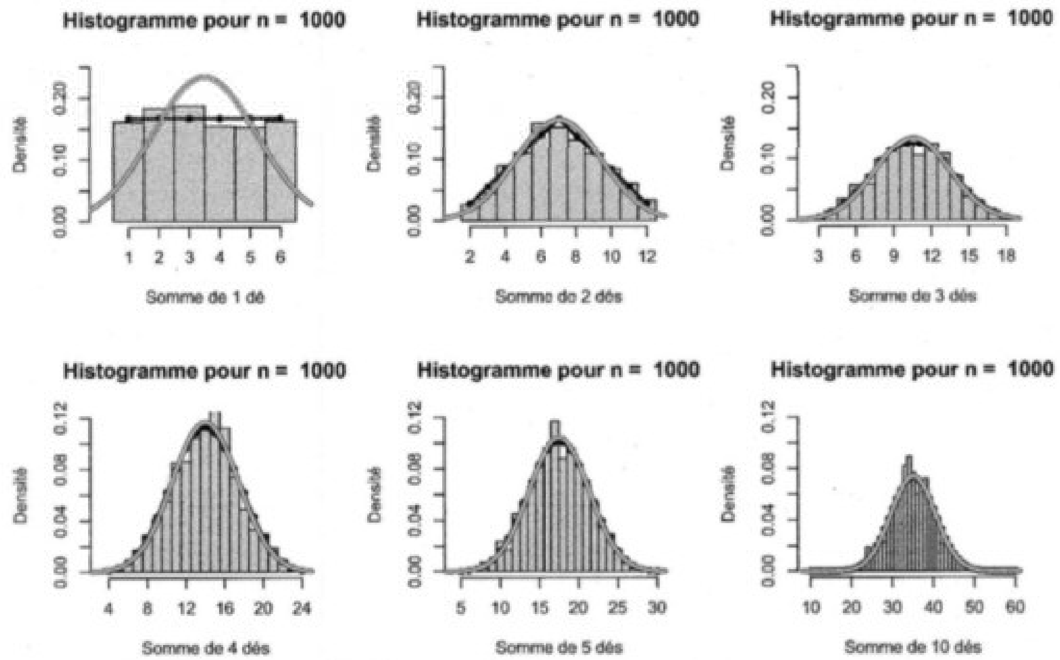


Le lecteur pourra comparer les pourcentages indiqués avec ceux donnés par le théorème de Tchebychev.

Même si cette loi est très importante, tous les phénomènes ne peuvent pas forcément être estimés par la loi normale.

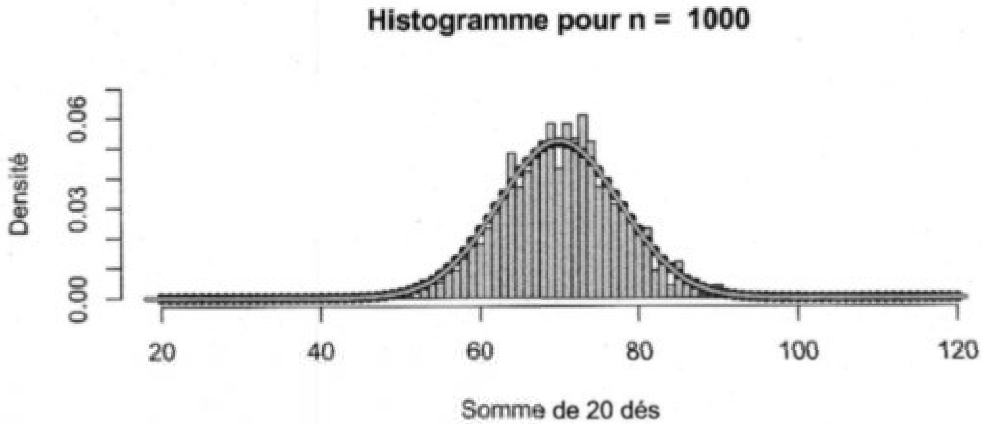
Approximation de la somme d'un jet de dés

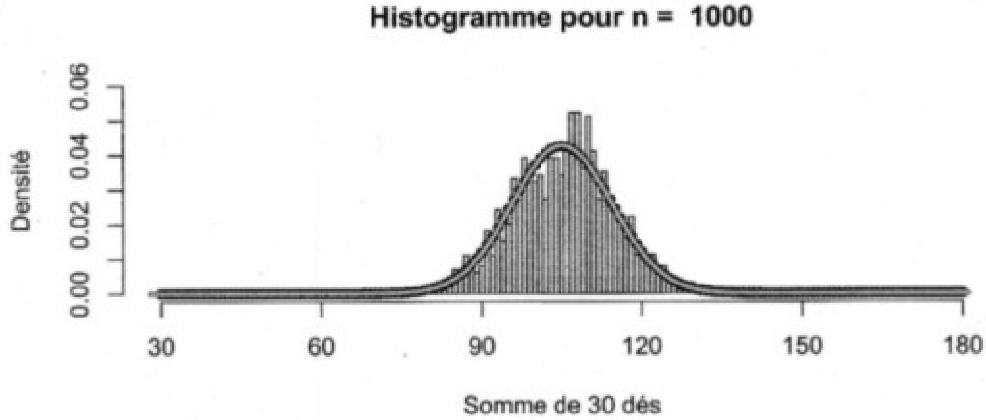
Superposons en rouge/blanc la loi normale $\mathcal{N}(k\mu, k\sigma^2)$ correspondant aux lancers de k dés pour $k \in \{1, 2, 3, 4, 5, 10\}$. Rappelons que $\mu = 3.5$ et $\sigma = \frac{35}{12}$.



On voit que lorsque n grandit, l'approximation par la loi normale devient meilleure.

Encore quelques comparaisons.





Comparons la probabilité exacte d'avoir, en lançant 30 dés, une somme comprise entre 120 et 180 avec son approximation par la loi normale $\mathcal{N}(105, 87.5)$.

méthodes de calcul	réponse
vraies probabilités (voir section 1.1.8)	$\mathbf{P}(120 \leqslant X \leqslant 180) \cong 0.06067$
loi normale $\mathcal{N}(105, 87.5)$	$\int_{120}^{180} f(x)\,dx \cong 0.05440$
loi normale $\mathcal{N}(105, 87.5)$ avec correction de continuité (voir aussi la page 28)	$\int_{119.5}^{180.5} f(x)\,dx \cong 0.06056$

La correction de continuité consiste à enlever $\frac{1}{2}$ à la première borne d'intégration et à ajouter $\frac{1}{2}$ à la deuxième borne. En effet, si on veut calculer la probabilité $\mathbf{P}(X = 100)$ et qu'on calcule $\int_{100}^{100} f(x)\,dx$, on va trouver 0, parce qu'il n'y a pas d'aire. Si on veut approximer $\mathbf{P}(X = 100)$ à l'aide de la loi normale, il faut calculer l'intégrale $\int_{99.5}^{100.5} f(x)\,dx$, car les barres des histogrammes sont centrées en 100 mais sont de largeur 1.

1.2.3 La fonction gamma

La fonction *gamma* est définie par l'intégrale suivante.

$$\Gamma(n) = \int_{0}^{+\infty} x^{n-1} e^{-x}\, dx \qquad n \in]0, +\infty[$$

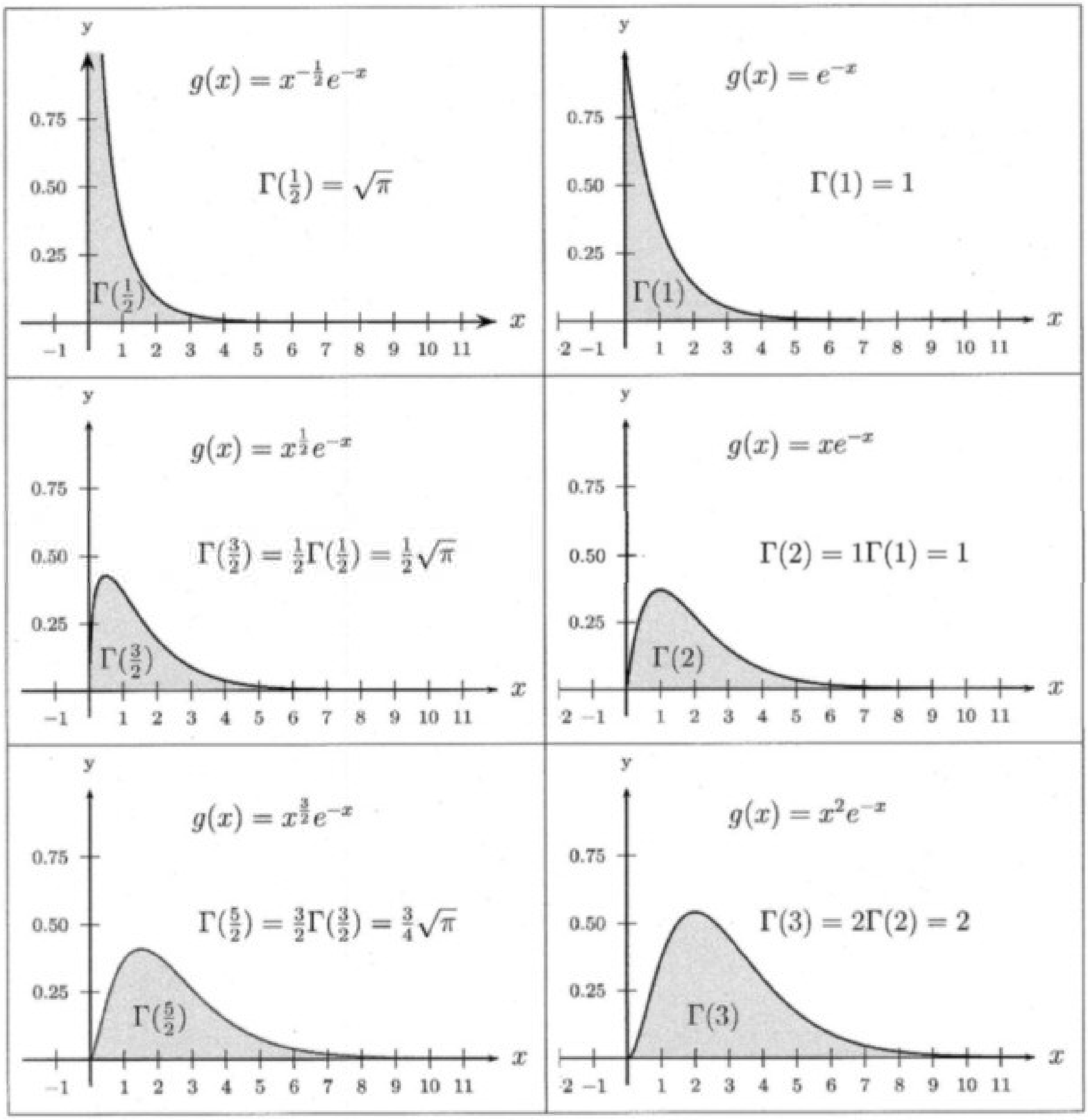
Propriétés de la fonction gamma.

$$\Gamma(n+1) = n \cdot \Gamma(n) \text{ pour tout } n > 0 \qquad \text{avec} \qquad \Gamma(\tfrac{1}{2}) = \sqrt{\pi} \;\text{ et }\; \Gamma(1) = 1$$

Ces propriétés permettent de calculer $\Gamma(\frac{n}{2})$ avec $n \geqslant 1$.

On voit aussi que lorsque n est un nombre naturel, $n \geqslant 1$, la fonction *gamma* est une généralisation de la factorielle, autrement dit qu'on a $\Gamma(n) = (n-1)!$.

En effet, la factorielle satisfait les propriétés $n! = n \cdot (n-1)!$ et $0! = 1$.



1.2.4 Les lois de Student

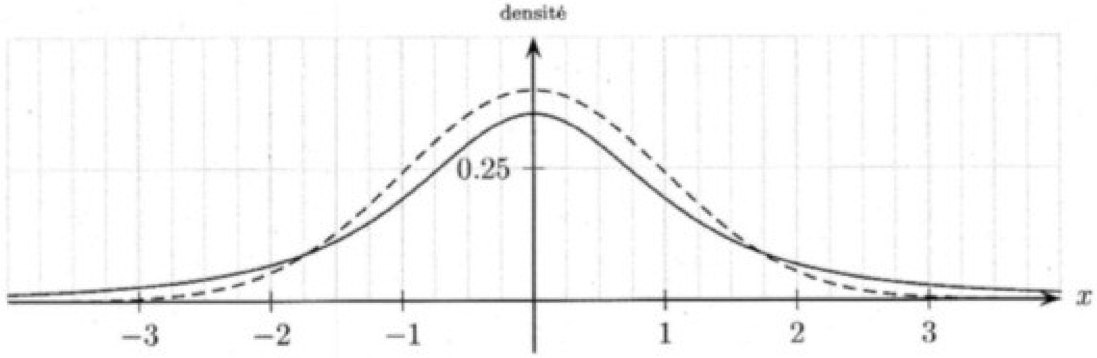
La densité de probabilité de la loi de Student T_ν à ν degrés de liberté est donnée par la fonction suivante.

$$f_{T_\nu}(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{1}{\sqrt{\nu\pi}} \cdot \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad , \; x \in \mathbb{R}$$

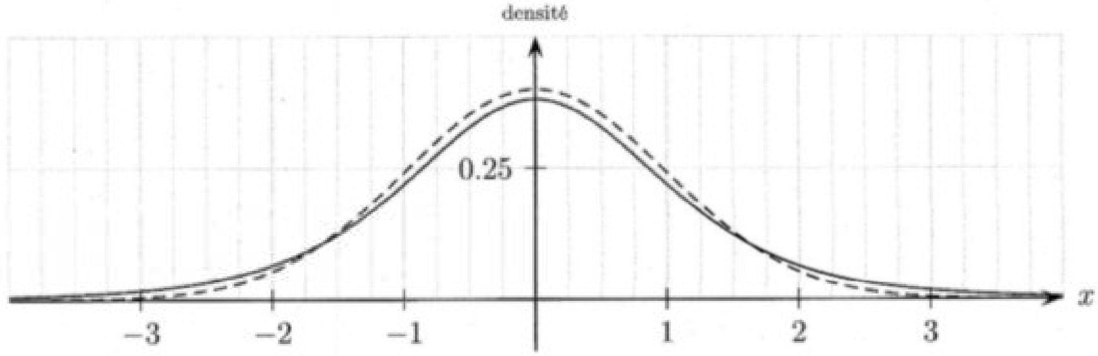
Il s'agit d'une distribution qui ressemble à celle de la loi normale centrée réduite, mais elle est plus haute sur les côtés.

Sur les graphiques suivants, on voit la loi normale centrée réduite $\mathcal{N}(0,1)$ en traitillés, les lois de Student pour $\nu = 2$, $\nu = 5$ et $\nu = 30$.

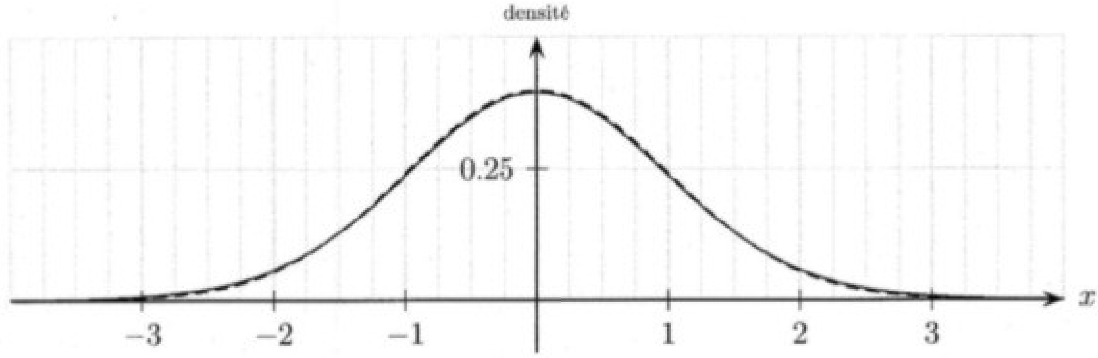
Comparaison en la loi normale $\mathcal{N}(0,1)$ et T_2.



Comparaison en la loi normale $\mathcal{N}(0,1)$ et T_5.

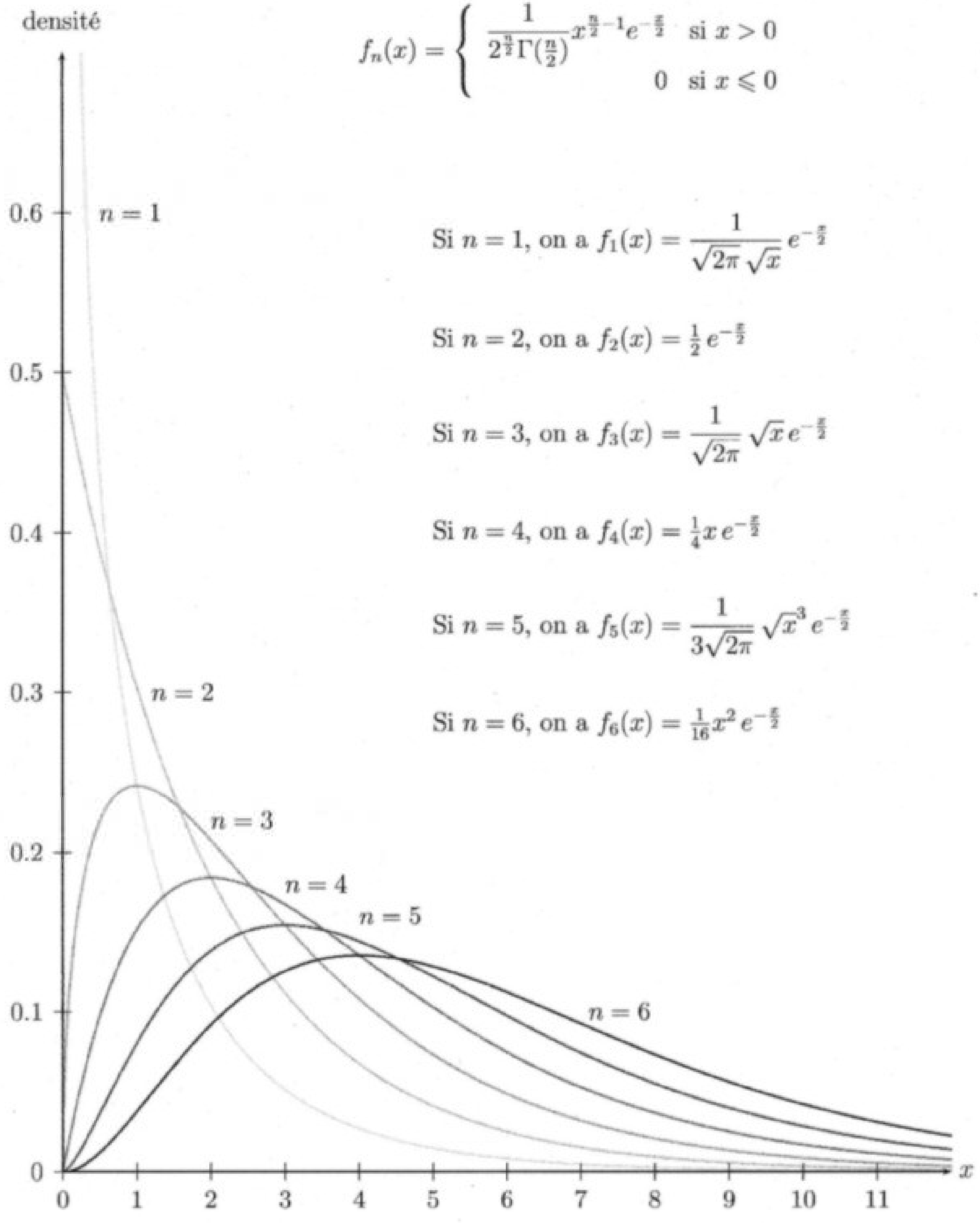


Comparaison en la loi normale $\mathcal{N}(0,1)$ et T_{30}.



1.2.5 Les lois du chi-carré

Il existe des lois, notées χ^2_n, et appelées lois du chi-carré (prononcer «ki») avec n degrés de liberté. Voici les expressions fonctionnelles de leur densité et leur représentation graphique.



$$f_n(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$$

Si $n=1$, on a $f_1(x) = \dfrac{1}{\sqrt{2\pi}\sqrt{x}} e^{-\frac{x}{2}}$

Si $n=2$, on a $f_2(x) = \frac{1}{2} e^{-\frac{x}{2}}$

Si $n=3$, on a $f_3(x) = \dfrac{1}{\sqrt{2\pi}} \sqrt{x}\, e^{-\frac{x}{2}}$

Si $n=4$, on a $f_4(x) = \frac{1}{4} x\, e^{-\frac{x}{2}}$

Si $n=5$, on a $f_5(x) = \dfrac{1}{3\sqrt{2\pi}} \sqrt{x}^3 e^{-\frac{x}{2}}$

Si $n=6$, on a $f_6(x) = \frac{1}{16} x^2 e^{-\frac{x}{2}}$

1.2.6 Espérance et variance dans le cas continu

Espérance d'une variable aléatoire continue

L'*espérance* d'une variable aléatoire X, notée $E(X)$ ou μ, est la moyenne pondérée des valeurs images de X par leur probabilité :

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\, dx \qquad \text{(en modifiant la formule du cas discret)}$$

Variance et écart type d'une variable aléatoire continue

La *variance* d'une variable aléatoire X, notée $V(X)$, est la moyenne pondérée des carrés des écarts par rapport à l'espérance μ par leur probabilité.

$$V(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x)\, dx$$

L'*écart type* de X est défini par $\sigma(X) = \sqrt{V(X)}$.

Généralisation de l'espérance

Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction réelle et X une variable aléatoire, alors $g \circ X$ est une variable aléatoire notée[10] $g(X)$ dont l'espérance est :

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

Cela permet notamment d'écrire la variance comme une espérance.

$$V(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x)\, dx = E((X-\mu)^2)$$

Exemple de calcul

Pour la variable aléatoire X définie en page 15, l'espérance est le résultat que l'on peut le plus souvent espérer (en moyenne) et vaut :

$$E(X) = \int_0^{2\pi} \frac{x}{2\pi}\, dx \overset{\text{TFCI}}{=} \left. \frac{x^2}{4\pi} \right|_0^{2\pi} = \frac{4\pi^2}{4\pi} = \pi$$

La variance est :

$$V(X) = \int_0^{2\pi} \frac{(x-\pi)^2}{2\pi}\, dx \overset{\text{TFCI}}{=} \left. \frac{(x-\pi)^3}{6\pi} \right|_0^{2\pi} = \frac{\pi^3}{6\pi} - \frac{(-\pi)^3}{6\pi} = \frac{\pi^2}{6} + \frac{\pi^2}{6} = \frac{\pi^2}{3}$$

Par conséquent, l'écart type vaut $\sigma(X) = \frac{\pi}{\sqrt{3}}$.

10. Cette notation est abusive : $g(X)(\omega)$ n'a pas de sens ! La justification est celle-ci : si $X : \Omega \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$, alors $(g \circ X)(\omega) = g(X(\omega)) = g(x)$ avec $x = X(\omega) \in \mathrm{Im}(X)$. D'où la notation $g \circ X = g(X)$.

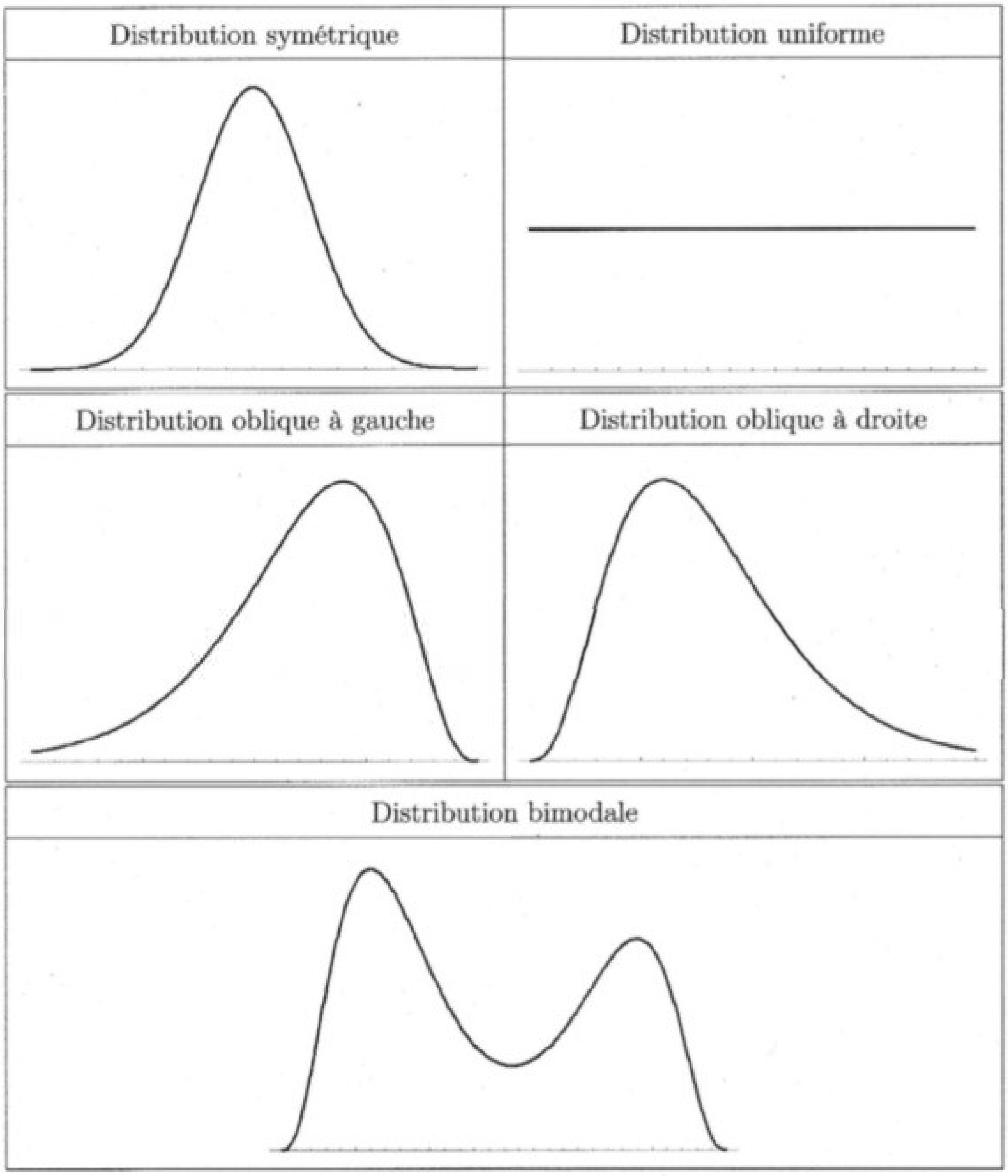
1.3 Les différentes formes de densité de distribution

Il s'agit d'une notion issue de la théorie des probabilités : une *densité* (ou *distribution*) est une description de la courbe que *devrait* suivre un histogramme lorsque les mesures utilisées pour faire l'histogramme sont en très grands nombre (grâce à la loi des grands nombres). Il s'agit d'une *courbe théorique* !

Tandis que l'histogramme est basé sur les observations, la distribution est théorique. Il faut donc faire preuve de prudence lorsqu'on dit que les données mesurées suivent une certaine distribution !

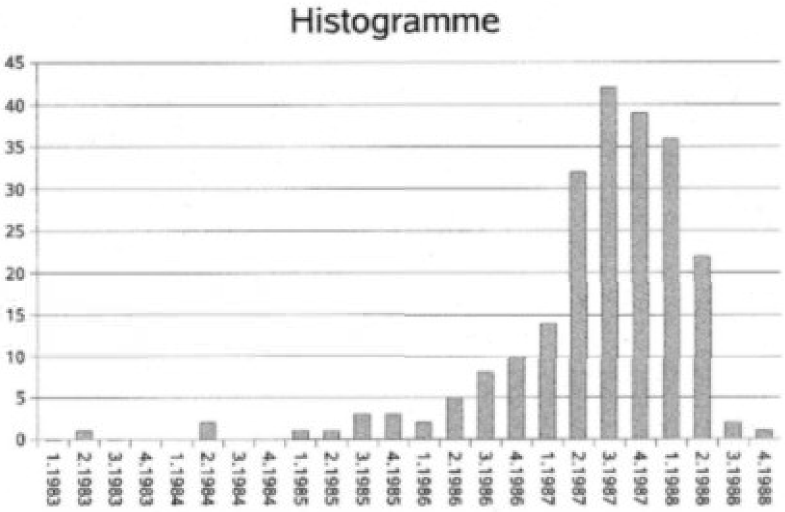
Dans les exemples précédents, les densités ont toujours été déterminées de manière exacte (loi de Bernoulli, loi binomiale ou loi multinomiale) ou approximative (loi normale, lois de Student, lois du chi-carré).

Voici les noms de certaines formes de distributions que l'on peut trouver dans la nature.

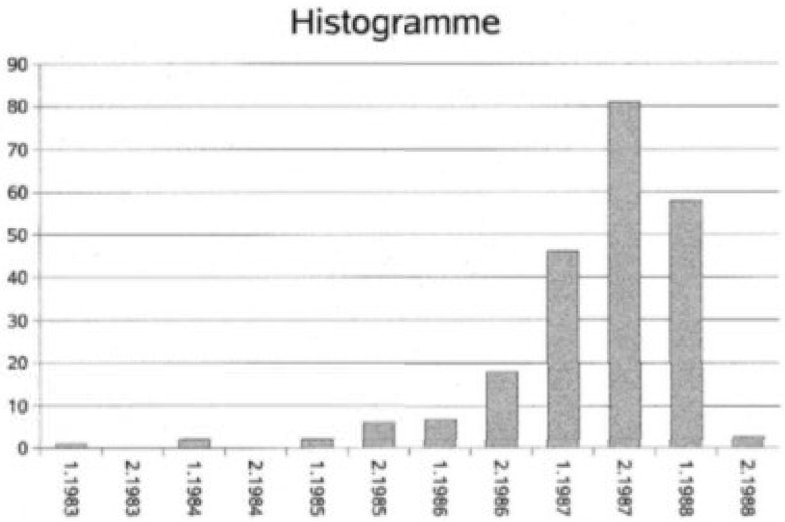


Date de naissance de élèves de première année (2003-2004)

On peut regarder la date de naissance des élèves de première année. Il y a 224 élèves inscrits en août 2003. Voici l'histogramme où les classes sont des trimestres (trois mois).



On voit beaucoup mieux la nécessité de classer les données dans des classes, afin d'avoir un graphe lisible. Voici l'histogramme où les classes sont des semestres (six mois).



Il faut faire attention à l'interprétation que l'on donne à l'histogramme, puisque le choix des classes influence légèrement l'histogramme. Pour cette raison, il est important de ne pas faire de calculs à partir des nombres indiqués sur l'histogramme, mais plutôt à partir des données initiales.

Même si le choix des classes modifie l'histogramme, on s'aperçoit que la forme générale reste la même. Ici, on voit qu'en première année, il y a beaucoup d'élèves qui sont nés en 1987 et dans la première moitié de 1988.

La distribution est oblique à gauche, car il y a aura certes quelques élèves surdoués qui auront sauté une année (donc qui entreront au lycée très jeune), mais il y a aura bien plus d'élèves qui entreront au lycée un peu plus vieux (redoublement, réorientation professionnelle, ...).

1.4 Propriétés de l'espérance et de la variance

Propriétés de l'espérance

L'espérance est linéaire (cela provient des propriétés du symbole somme et de l'intégrale). Cela signifie que l'espérance satisfait :

$$\begin{array}{lrcl} \text{additivité :} & E(X+Y) & = & E(X)+E(Y) \\ \text{homogénéïté :} & E(\lambda X) & = & \lambda\, E(X) \end{array}$$

De plus, si C est une variable aléatoire constante, on a $E(C) = C$.

Propriétés de la variance

La variance n'est pas linéaire. On a tout de même :

$$\begin{array}{lrcl} \text{additivité si } X \text{ et } Y \text{ sont indépendants :} & V(X+Y) & = & V(X)+V(Y) \\ \text{homogénéïté de degré 2 :} & V(\lambda X) & = & \lambda^2\, V(X) \\ \text{lorsque } C \text{ est une variable aléatoire constante :} & V(C) & = & 0 \end{array}$$

Pour démontrer la formule de l'additivité, il faut définir la notion de variables aléatoires indépendantes (et aussi la notion de covariance) : cela dépasse le cadre de ce cours. Par contre, les deux autres propriétés sont évidentes.

Autre formule pour la variance

En utilisant les propriétés de l'espérance, on montre que $\boxed{V(X) = E(X^2) - \big(E(X)\big)^2}$. En effet, en notant $\mu = E(X)$, on a :

$$V(X) = E\Big(\big(X-\mu\big)^2\Big) = E\big(X^2-2\mu X+\mu^2\big) = E(X^2)-2\mu E(X)+\mu^2 = E(X^2)-\big(E(X)\big)^2$$

Théorème

On considère n variables aléatoires indépendantes $X_1, X_2, \ldots, X_n$ de même espérance μ et de même variance σ^2. La moyenne de ces n variables aléatoires est aussi une variable aléatoire, notée $\overline{X}$ et définie par $\overline{X} = \frac{X_1+X_2+\cdots+X_n}{n}$. Sous ces hypothèses, on a :

$$E(\overline{X}) = \mu \qquad V(\overline{X}) = \frac{\sigma^2}{n}$$

Preuve

On a : $$E(\overline{X}) = E\left(\frac{\sum_{i=1}^n X_i}{n}\right) \overset{E \text{ lin.}}{=} \frac{\sum_{i=1}^n E(X_i)}{n} = \frac{\sum_{i=1}^n \mu}{n} = \frac{n\mu}{n} = \mu$$

Pour la variance :

$$V(\overline{X}) = V\left(\frac{\sum_{i=1}^n X_i}{n}\right) = V\Big(\frac{1}{n}\sum_{i=1}^n X_i\Big) = \frac{1}{n^2}\, V\Big(\sum_{i=1}^n X_i\Big)$$

Point délicat : on a $V(\sum_{i=1}^n X_i) = \sum_{i=1}^n V(X_i)$ parce que les variables aléatoires X_i sont indépendantes. Par conséquent :

$$V(\overline{X}) = \frac{1}{n^2}\, V\Big(\sum_{i=1}^n X_i\Big) = \frac{1}{n^2}\, \sum_{i=1}^n V(X_i) = \frac{1}{n^2}\, \sum_{i=1}^n \sigma^2 = \frac{1}{n^2}\, n\sigma^2 = \frac{\sigma^2}{n}$$

□

1.5 Centrage et réduction

Définitions

- Une variable aléatoire (discrète ou continue) est dite *centrée* si son espérance est nulle.
- Une variable aléatoire (discrète ou continue) est dite *réduite* si sa variance vaut 1.

Théorème

Si X est une variable aléatoire (discrète ou continue) d'espérance μ et de variance σ^2, alors

$$\frac{X-\mu}{\sigma} \text{ est une variable centrée réduite}$$

Preuve

1. On a, par linéarité :

$$E\left(\frac{X-\mu}{\sigma}\right) = E\left(\frac{1}{\sigma}(X-\mu)\right) \overset{E \text{ lin.}}{=} \frac{1}{\sigma}\cdot E(X-\mu) \overset{E \text{ lin.}}{=} \frac{1}{\sigma}\cdot \overbrace{(E(X)-\mu)}^{=0} = 0$$

2. On commence par utiliser la formule $V(X) = E\big((X-E(X))^2\big)$ afin d'avoir :

$$V\left(\frac{X-\mu}{\sigma}\right) = E\left(\left(\frac{X-\mu}{\sigma} - \underbrace{E\left(\frac{X-\mu}{\sigma}\right)}_{=0 \text{ (voir ci-dessus)}}\right)^2\right)$$

On peut ainsi continuer le calcul :

$$\begin{aligned} V\left(\frac{X-\mu}{\sigma}\right) &= E\left(\left(\frac{X-\mu}{\sigma}\right)^2\right) = E\left(\frac{(X-\mu)^2}{\sigma^2}\right) = E\left(\frac{1}{\sigma^2}\cdot(X-\mu)^2\right) \\ &= \frac{1}{\sigma^2}\cdot E\big((X-\mu)^2\big) = \frac{1}{\sigma^2}\cdot E\big(X^2-2X\mu+\mu^2\big) \\ &\overset{E \text{ lin.}}{=} \frac{1}{\sigma^2}\cdot\Big(E(X^2)-2\mu E(X)+\mu^2\Big) \overset{\mu=E(X)}{=} \frac{1}{\sigma^2}\cdot\Big(E(X^2)-\big(E(X)\big)^2\Big) \end{aligned}$$

On termine en se souvenant de la formule $V(X) = E(X^2) - \big(E(X)\big)^2$. En effet :

$$V\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\cdot\Big(E(X^2)-\big(E(X)\big)^2\Big) = \frac{1}{\sigma^2}\cdot V(X) = \frac{1}{\sigma^2}\cdot\sigma^2 = 1$$

$\square$

1.6 Fonction de répartition d'une variable aléatoire

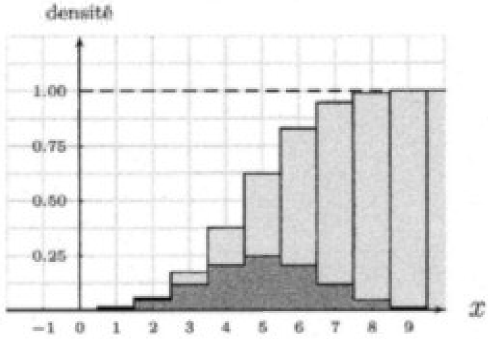
1.6.1 Fonction de répartition d'une variable discrète

On définit la fonction de répartition d'une variable discrète comme suit.

$$\phi(x) = \mathbf{P}(X \leqslant x) = \sum_{k=-\infty}^{x} \mathbf{P}(X = k) \quad \text{où } \mathbf{P}(X = k) \text{ est la densité discrète de probabilité}$$

Exemple

Ci-dessous, on trouve la loi binomiale $B(10, \frac{1}{2})$ (dont la densité est en gris foncé) et sa fonction de répartition (en gris clair).



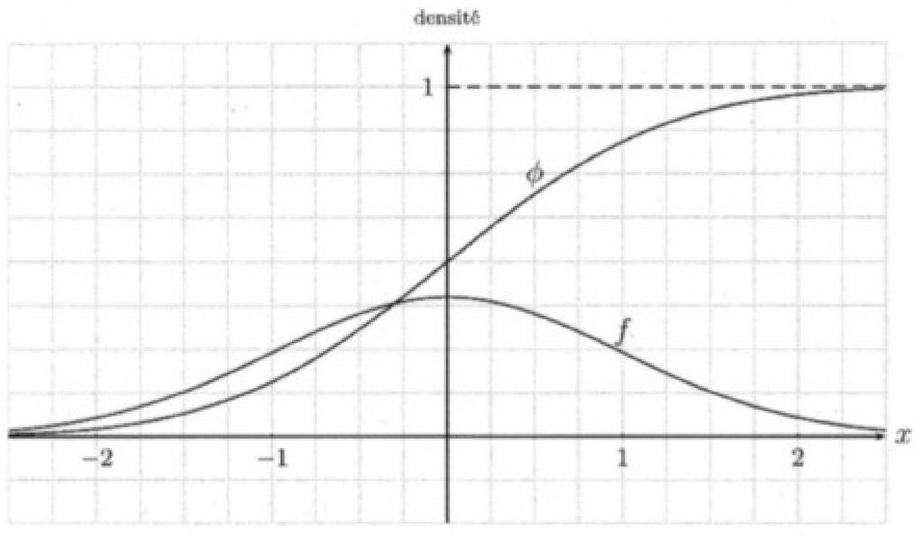
1.6.2 Fonction de répartition d'une variable continue

On définit la fonction de répartition d'une variable continue comme suit.

$$\phi(x) = \mathbf{P}(X \leqslant x) = \int_{-\infty}^{x} f(t)\, dt \quad \text{où} \quad f(x) \text{ est la densité de probabilité}$$

Exemple

Ci-dessous, on trouve la loi normale $\mathcal{N}(0, 1)$ et sa fonction de répartition.



1.7 Deux propriétés essentielles de la loi normale

1.7.1 La loi normale permet d'approximer la loi binomiale

Le théorème central-limite de Laplace

Soit a, $b \in \mathbb{R}$ et $p \in]0,1[$. Alors

La loi binomiale $B(n,p)$ suit approximativement une loi normale $\mathcal{N}(np, np(1-p))$

Autrement dit, pour n assez grand, on peut approximer la probabilité que la variable X, qui suit une loi binomiale $B(n,p)$ d'espérance np et de variance $np(1-p)$, soit entre a et b de la manière suivante.

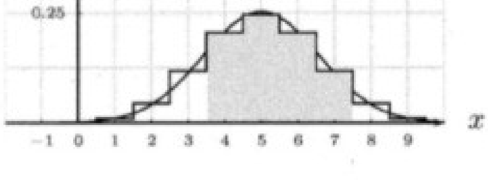
$$\begin{aligned}
\mathbf{P}\Big(a \leqslant X \leqslant b\Big) \underset{N \sim \mathcal{N}(np,np(1-p))}{\overset{\text{approx par}}{\cong}} & \ \mathbf{P}\Big(a \leqslant N \leqslant b\Big) \\
\underset{\text{réduction}}{\overset{\text{centrage}}{=}} & \ \mathbf{P}\Big(\frac{a-np}{\sqrt{np(1-p)}} \leqslant \overbrace{\frac{N-np}{\sqrt{np(1-p)}}}^{=Z\sim\mathcal{N}(0,1)} \leqslant \frac{b-np}{\sqrt{np(1-p)}}\Big) \\
\underset{\text{répartition}}{\overset{\text{fonction de}}{=}} & \ \mathbf{P}\Big(Z \leqslant \frac{b-np}{\sqrt{np(1-p)}}\Big) - \mathbf{P}\Big(Z \leqslant \frac{a-np}{\sqrt{np(1-p)}}\Big) \\
\underset{\text{page 30}}{\overset{\text{notation}}{=}} & \ \phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - \phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right)
\end{aligned}$$

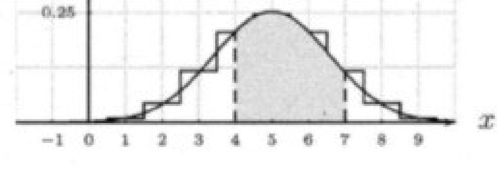
La plupart du temps, on considère que si $np \geqslant 5$ et si $n(1-p) \geqslant 5$ (autrement dit si n est suffisamment grand par rapport aux probabilités de succès p et d'échec $1-p$), alors l'approximation est de bonne qualité.

Exemple

Soit X une variable aléatoire qui suit la loi binomiale $B(10, \frac{1}{2})$ (on lance une pièce de monnaie bien équilibrée 10 fois de suite et on compte le nombre de piles obtenu). On a $np = n(1-p) = 5$. Ainsi, on estime $B(10, \frac{1}{2})$ avec la loi normale $\mathcal{N}(np, np(1-p)) = \mathcal{N}(5, 2.5)$.

En gris, on a $\mathbf{P}(4 \leqslant X \leqslant 7)$ que l'on voit sur le graphe de gauche et que l'on estime par la loi normale sur le graphe de droite.





On a :

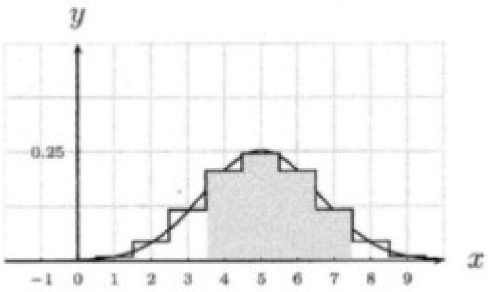
$$\begin{aligned}\mathbf{P}(4 \leqslant X \leqslant 7) &= \mathbf{P}(X=4)+\mathbf{P}(X=5)+\mathbf{P}(X=6)+\mathbf{P}(X=7) \\ &= \binom{10}{4}\left(\tfrac{1}{2}\right)^4\left(\tfrac{1}{2}\right)^6+\binom{10}{5}\left(\tfrac{1}{2}\right)^5\left(\tfrac{1}{2}\right)^5+\binom{10}{6}\left(\tfrac{1}{2}\right)^6\left(\tfrac{1}{2}\right)^4+\binom{10}{7}\left(\tfrac{1}{2}\right)^7\left(\tfrac{1}{2}\right)^3 \\ &= \tfrac{792}{1024}=\tfrac{99}{128}=0.7734375\end{aligned}$$

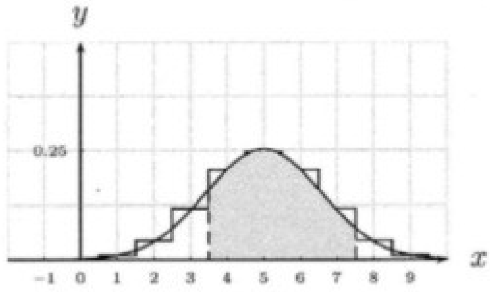
L'approximation est donnée par

$$\mathbf{P}(4 \leqslant X \leqslant 7) \cong \phi(\tfrac{7-5}{\sqrt{2.5}})-\phi(\tfrac{4-5}{\sqrt{2.5}}) \cong \phi(1.26)-\phi(-0.63) \cong 0.6319$$

Amélioration de l'estimation par correction de continuité

L'estimation de la page précédente n'est pas extraordinaire. Pour l'améliorer, on va légèrement modifier les bornes de l'intégrale. En effet, en allant de 4 à 7, l'intégrale précédente 'oublie' la moitié des deux rectangles aux extrémités. C'est pourquoi il est plus judicieux de partir à 3.5 et de finir à 7.5 comme le montre les schémas ci-dessous.





On a :

$$\mathbf{P}(4 \leqslant X \leqslant 7) = 0.7734375$$

L'approximation est maintenant bien meilleure

$$\mathbf{P}(4 \leqslant X \leqslant 7) \cong \phi(\tfrac{7.5-5}{\sqrt{2.5}})-\phi(\tfrac{3.5-5}{\sqrt{2.5}}) \cong \phi(1.58)-\phi(-0.95) \cong 0.7718$$

Formule d'approximation améliorée

La formule d'approximation améliorée est donc dans le cas général donnée par

$$\mathbf{P}\Big(a \leqslant X \leqslant b\Big) \cong \phi\left(\frac{b+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)-\phi\left(\frac{a-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)$$

Remarque

Si n est très grand, les deux approximations (l'améliorée et celle du théorème) seront les deux très proches.

1.7.2 Le théorème de la limite centrale

Le *théorème de la limite centrale*, aussi appelé *théorème central limite* est l'un des résultats les plus importants de la théorie des probabilités. De façon informelle, ce théorème[11] donne une estimation de l'erreur que l'on commet en approchant la moyenne théorique μ par la moyenne arithmétique $\overline{x}$.

Théorème de la limite centrale

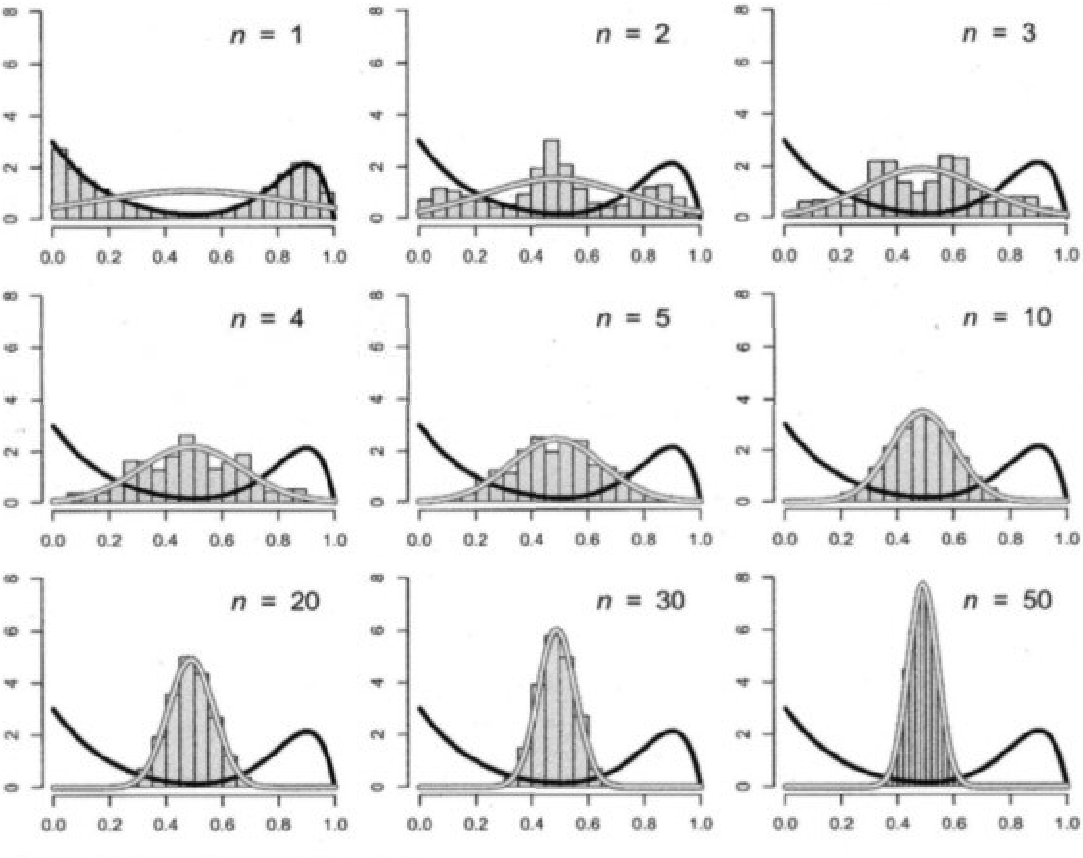
Supposons qu'on a n variables aléatoires $X_1, \ldots, X_n$ indépendantes et qui suivent une même loi de probabilité d'espérance μ et de variance σ^2.

Le théorème central limite dit que, si n est suffisamment grand (la plupart du temps $n \geqslant 30$ suffit), alors la variable aléatoire $\overline{X} = \frac{X_1 + \cdots + X_n}{n}$ qui correspond à la moyenne des variables $X_1, \ldots, X_n$ suit approximativement une loi normale $\mathcal{N}(\mu, \frac{\sigma^2}{n})$.

Simulation numérique

En noir, la loi d'origine. En rouge/blanc, la loi normale du théorème. En gris, l'histogramme des moyennes de 10 000 échantillons de taille n.

Pour $n = 1$, l'histogramme suit la loi d'origine, puisque la moyenne d'une mesure est égale à la mesure elle-même. Puis plus n grandit, plus l'histogramme cesse de suivre la loi d'origine pour se rapprocher de la loi normale (qui devient de plus en plus étroite, car $\frac{\sigma^2}{n}$ devient de plus en plus petit, et haute, car son aire vaut toujours 1).



11. Ce phénomène a d'abord été observé par Gauss qui l'appelait loi des erreurs ; mais ce dernier n'en a pas donné de démonstration rigoureuse. La preuve du théorème a été apportée par deMoivre et Laplace ; le théorème porte donc parfois leurs noms.

1.8 Tables

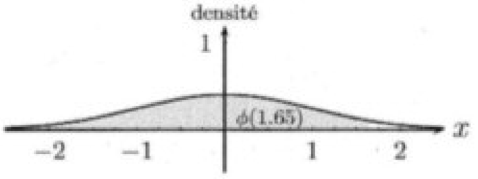
1.8.1 Fonction de répartition de la loi normale

La fonction de répartition $\phi(x)$ de la loi normale centrée réduite $\mathcal{N}(0,1)$ est une intégrale qui n'admet pas de primitive explicite. On recourt ainsi à une table pour donner des valeurs approximatives.

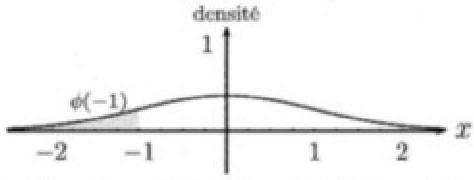
$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt$$

	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	0.7703	0.7734	0.7764	0.7793	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	0.8906	0.8925	0.8943	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	0.9986	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

On lit les décimales dans les lignes, et les centièmes en colonnes. Par exemple, on trouve l'approximation arrondie à 4 décimales de $\phi(1.65)$, qui vaut 0.9505, à l'intersection de la ligne 1.6 et de la colonne 0.05. Si x est négatif, on utilise la relation $\phi(-x) = 1 - \phi(x)$ qui provient du fait que la densité est paire et que l'aire sous la courbe vaut 1.



$\phi(1.65) \cong 0.9505$



$\phi(-1) = 1 - \phi(1) \cong 1 - 0.8413 = 0.1587$

1.8.2 Les quantiles des lois de Student

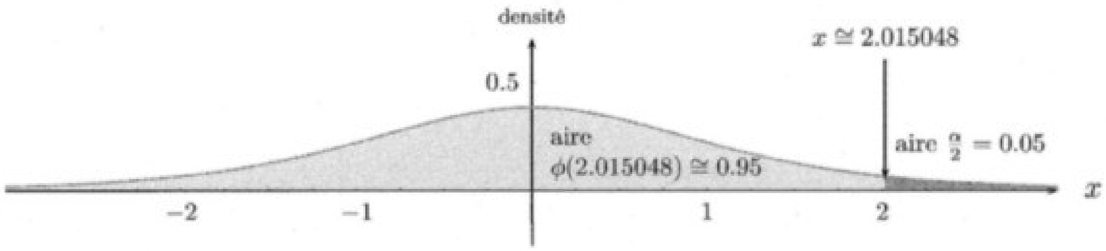
La fonction de répartition ϕ_ν de la loi de Student T_ν à ν degrés de liberté est une intégrale qui n'admet pas de primitive explicite (sauf si $\nu = 1$).

$$\phi_\nu(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{1}{\sqrt{\nu\pi}} \cdot \int_{-\infty}^{x} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} dt \quad , \; x \in \mathbb{R}$$

Contrairement à la table de la loi normale, il est coutume de préférer une table des quantiles (c'est-à-dire une table qui donne la valeur de x lorsqu'on connaît la valeur de l'aire $\phi_\nu(x)$).

ν \ $\phi_\nu(x)$	0.60	0.75	0.90	0.95	0.975	0.99	0.995	0.9995
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.302653	6.964557	9.924843	31.59906
3	0.276671	0.764892	1.637744	2.353363	3.182446	4.540703	5.840909	12.92398
4	0.270722	0.740697	1.533206	2.131847	2.776445	3.746947	4.604095	8.610302
5	0.267181	0.726687	1.475884	2.015048	2.570582	3.364930	4.032143	6.868827
6	0.264835	0.717558	1.439756	1.943180	2.446912	3.142668	3.707428	5.958816
7	0.263167	0.711142	1.414924	1.894579	2.364624	2.997952	3.499483	5.407883
8	0.261921	0.706387	1.396815	1.859548	2.306004	2.896459	3.355387	5.041305
9	0.260955	0.702722	1.383029	1.833113	2.262157	2.821438	3.249836	4.780913
10	0.260185	0.699812	1.372184	1.812461	2.228139	2.763769	3.169273	4.586894
11	0.259556	0.697445	1.363430	1.795885	2.200985	2.718079	3.105807	4.436979
12	0.259033	0.695483	1.356217	1.782288	2.178813	2.680998	3.054540	4.317791
13	0.258591	0.693829	1.350171	1.770933	2.160369	2.650309	3.012276	4.220832
14	0.258213	0.692417	1.345030	1.761310	2.144787	2.624494	2.976843	4.140454
15	0.257885	0.691197	1.340606	1.753050	2.131450	2.602480	2.946713	4.072765
16	0.257599	0.690132	1.336757	1.745884	2.119905	2.583487	2.920782	4.014996
17	0.257347	0.689195	1.333379	1.739607	2.109816	2.566934	2.898231	3.965126
18	0.257123	0.688364	1.330391	1.734064	2.100922	2.552380	2.878440	3.921646
19	0.256923	0.687621	1.327728	1.729133	2.093024	2.539483	2.860935	3.883406
20	0.256743	0.686954	1.325341	1.724718	2.085963	2.527977	2.845340	3.849516
21	0.256580	0.686352	1.323188	1.720743	2.079614	2.517648	2.831360	3.819277
22	0.256432	0.685805	1.321237	1.717144	2.073873	2.508325	2.818756	3.792131
23	0.256297	0.685306	1.319460	1.713872	2.068658	2.499867	2.807336	3.767627
24	0.256173	0.684850	1.317836	1.710882	2.063899	2.492159	2.796940	3.745399
25	0.256060	0.684430	1.316345	1.708141	2.059539	2.485107	2.787436	3.725144
26	0.255955	0.684043	1.314972	1.705618	2.055529	2.478630	2.778715	3.706612
27	0.255858	0.683685	1.313703	1.703288	2.051831	2.472660	2.770683	3.689592
28	0.255768	0.683353	1.312527	1.701131	2.048407	2.467140	2.763262	3.673906
29	0.255684	0.683044	1.311434	1.699127	2.045230	2.462021	2.756386	3.659405
30	0.255605	0.682756	1.310415	1.697261	2.042272	2.457262	2.749996	3.645959
31	0.255532	0.682486	1.309464	1.695519	2.039513	2.452824	2.744042	3.633456
32	0.255464	0.682234	1.308573	1.693889	2.036933	2.448678	2.738481	3.621802
33	0.255399	0.681997	1.307737	1.692360	2.034515	2.444794	2.733277	3.610913
34	0.255339	0.681774	1.306952	1.690924	2.032245	2.441150	2.728394	3.600716
35	0.255281	0.681564	1.306212	1.689572	2.030108	2.437723	2.723806	3.591147
inf	0.253347	0.674490	1.281552	1.644854	1.959964	2.326348	2.575829	3.290527

Si, par exemple, pour un test d'hypothèses symétriques, le seuil de signification α est égal à 10%, c'est-à-dire que $\alpha = 0.1$, alors on a $\phi_\nu(x) = 1 - \frac{0.1}{2} = 0.95$. On trouve $x = \phi_\nu^{-1}(0.95)$ en regardant la valeur se trouvant dans la colonne correspondant à 0.95 et dans la ligne correspondant aux degrés de liberté ν. Si, de plus, on choisit $\nu = 5$, on trouve $x \cong 2.015048$.



1.8.3 Les quantiles des lois du chi-carré

La fonction de répartition ϕ_ν de la loi du chi-carré χ^2_ν à ν degrés de liberté est une intégrale qui n'admet pas de primitive explicite (sauf si ν est pair).

$$\phi_\nu(x) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \int_0^x t^{\frac{n}{2}-1} e^{-\frac{t}{2}}\, dt \quad , \ x \in [0, +\infty[$$

Comme pour la table des lois de Student, il est coutume de préférer une table des quantiles (c'est-à-dire une table qui donne la valeur de x lorsqu'on connaît la valeur de l'aire $\phi_\nu(x)$).

ν \ $\phi_\nu(x)$	0.500	0.750	0.800	0.900	0.950	0.990	0.995	0.999
1	0.455	1.323	1.642	2.706	3.841	6.635	7.879	10.828
2	1.386	2.773	3.219	4.605	5.991	9.210	10.597	13.816
3	2.366	4.108	4.642	6.251	7.815	11.345	12.838	16.266
4	3.357	5.385	5.989	7.779	9.488	13.277	14.86	18.467
5	4.351	6.626	7.289	9.236	11.07	15.086	16.75	20.515
6	5.348	7.841	8.558	10.645	12.592	16.812	18.548	22.458
7	6.346	9.037	9.803	12.017	14.067	18.475	20.278	24.322
8	7.344	10.219	11.030	13.362	15.507	20.090	21.955	26.124
9	8.343	11.389	12.242	14.684	16.919	21.666	23.589	27.877
10	9.342	12.549	13.442	15.987	18.307	23.209	25.188	29.588
11	10.341	13.701	14.631	17.275	19.675	24.725	26.757	31.264
12	11.340	14.845	15.812	18.549	21.026	26.217	28.300	32.909
13	12.340	15.984	16.985	19.812	22.362	27.688	29.819	34.528
14	13.339	17.117	18.151	21.064	23.685	29.141	31.319	36.123
15	14.339	18.245	19.311	22.307	24.996	30.578	32.801	37.697
16	15.338	19.369	20.465	23.542	26.296	32.000	34.267	39.252
17	16.338	20.489	21.615	24.769	27.587	33.409	35.718	40.790
18	17.338	21.605	22.760	25.989	28.869	34.805	37.156	42.312
19	18.338	22.718	23.900	27.204	30.144	36.191	38.582	43.820
20	19.337	23.828	25.038	28.412	31.410	37.566	39.997	45.315
21	20.337	24.935	26.171	29.615	32.671	38.932	41.401	46.797
22	21.337	26.039	27.301	30.813	33.924	40.289	42.796	48.268
23	22.337	27.141	28.429	32.007	35.172	41.638	44.181	49.728
24	23.337	28.241	29.553	33.196	36.415	42.980	45.559	51.179
25	24.337	29.339	30.675	34.382	37.652	44.314	46.928	52.620
30	29.336	34.800	36.250	40.256	43.773	50.892	53.672	59.703
35	34.336	40.223	41.778	46.059	49.802	57.342	60.275	66.619
40	39.335	45.616	47.269	51.805	55.758	63.691	66.766	73.402
45	44.335	50.985	52.729	57.505	61.656	69.957	73.166	80.077
50	49.335	56.334	58.164	63.167	67.505	76.154	79.490	86.661
55	54.335	61.665	63.577	68.796	73.311	82.292	85.749	93.168
60	59.335	66.981	68.972	74.397	79.082	88.379	91.952	99.607
65	64.335	72.285	74.351	79.973	84.821	94.422	98.105	105.988
70	69.334	77.577	79.715	85.527	90.531	100.425	104.215	112.317
75	74.334	82.858	85.066	91.061	96.217	106.393	110.286	118.599
80	79.334	88.130	90.405	96.578	101.879	112.329	116.321	124.839
85	84.334	93.394	95.734	102.079	107.522	118.236	122.325	131.041
90	89.334	98.650	101.054	107.565	113.145	124.116	128.299	137.208
95	94.334	103.899	106.364	113.038	118.752	129.973	134.247	143.344

Si, par exemple, le seuil de signification α vaut 5%, c'est-à-dire que $\alpha = 0.05$, alors on a $\phi_\nu(x) = 1 - 0.05 = 0.95$. On trouve $x = \phi_\nu^{-1}(0.95)$ en regardant la valeur se trouvant dans la colonne correspondant à 0.95 et dans la ligne correspondant aux degrés de liberté ν. Si, de plus, on choisit $\nu = 4$, on trouve $x \cong 9.488$.