Loi normale

Les poids indiqués sur les emballages correspondent rarement à la masse réelle du contenu. Intuitivement on peut penser que la variable aléatoire égale à cette masse prend des valeurs symétriques, de part et d'autre de la masse théorique. Ces valeurs étant de moins en moins nombreuses au fur et à mesure que l'on s'éloigne de la masse théorique.

Les lois exponentielles et uniformes ne sont donc pas les plus appropriées.
Au xix^e siècle, pour modéliser de telles fluctuations, les mathématiciens Laplace et Gauss ont défini la célèbre courbe en cloche, ou courbe de Gauss.
La fonction de densité permettant de construire cette courbe dépend de deux paramètres, la moyenne et l'écart-type. L'écriture de la fonction n'est pas à mémoriser.
Par contre, on retiendra le changement de variable permettant d'accéder à la table de lecture de la loi normale centrée réduite de paramètres 0 et 1.

1. Qu'est-ce qu'une loi normale ?

Loi normale : *N*(1 ; 2)

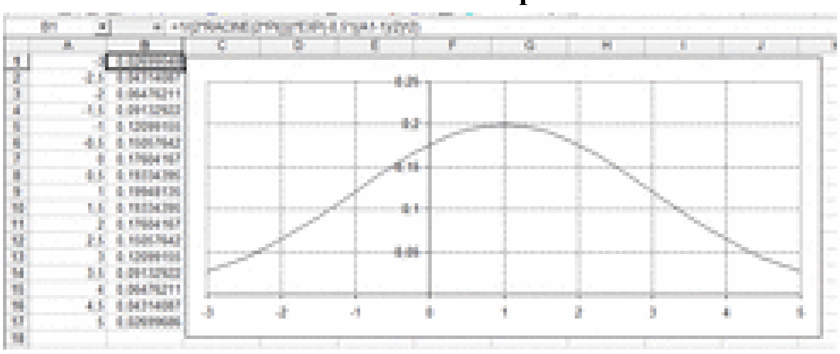
Les mathématiciens Laplace et Gauss ont mis au point cette fonction à densité *f*, définie sur $\mathbb{R}$, qui se représente par une courbe en cloche.

Elle s'écrit $f(x) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\frac{t-m}{\sigma}\right)^2}$ **, *m* et** σ **sont ses deux paramètres.**

On admet que si une variable aléatoire *X* a sa fonction de densité égale à *f*, alors $E(X) = m$ **,** $V(X) = \sigma^2$ **et** $\sigma(X) = \sigma$ **.**

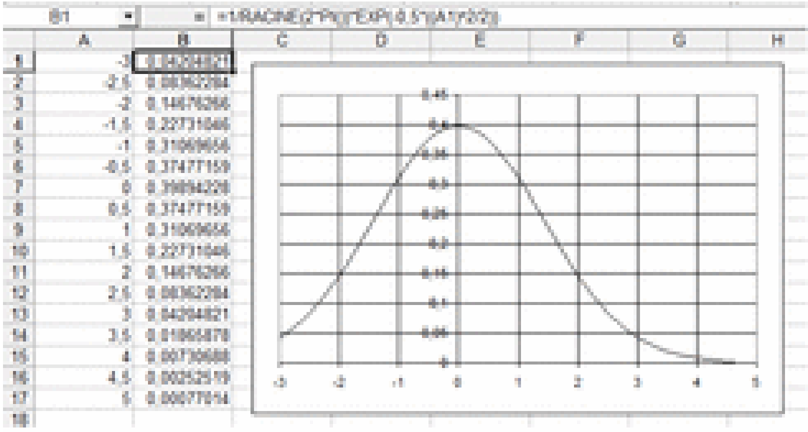
Tracé de la courbe en cloche pour la loi normale *N*(1 ; 2)

Son maximum est atteint pour $t = 1$**.**



Loi normale centrée réduite : *N*(0 ; 1)

La fonction à densité s'écrit à présent $f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}t^2}$ **.**
C'est une fonction paire, sa représentation graphique est symétrique par rapport à l'axe des ordonnées.



Lecture de la table de la loi normale centrée réduite

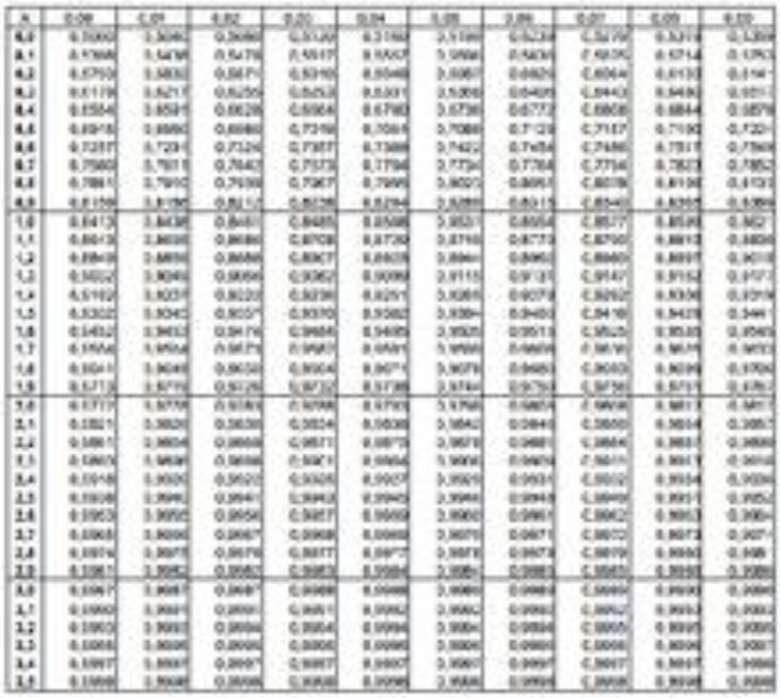
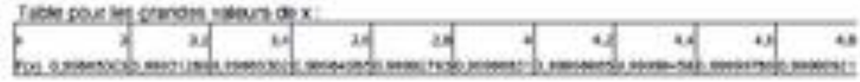


Table pour les grandes valeurs de x :



Si la variable aléatoire *X* suit la loi normale centrée réduite alors $P(X \leqslant a) = \Pi(a) = \int_{-\infty}^{a} f(t)\mathrm{d}t$ **.**

Exemple

$P(X \leqslant 1,32) = 0,906\,6$ **. (On lit le chiffre des unités et la première décimale dans la colonne de gauche, la seconde décimale sur la première ligne.)**

Cas d'un nombre comportant plus de deux décimales : on procède par interpolation linéaire.

On lit $P(X \leqslant 1,33) = 0,908\,2$ **et on veut calculer** $P(X \leqslant 1,324)$ **.**

Pour un écart de 0,01 la différence tabulaire est de $0,908\,2 - 0,906\,6 = 0,001\,6$ **, soit** $dt = 0,001\,6$ **; pour un écart de 0,004 il est alors de** $0,001\,6 \times \frac{0,004}{0,01} = 0,000\,64$ **.**

Donc $P(X \leqslant 1,324) = 0,906\,6 + 0,000\,64 = 0,907\,24$ **.**

Lecture sur la calculatrice

La variable aléatoire *X* suit la loi normale $N(m\,;\,\sigma)$. On veut calculer $P(a \leqslant X \leqslant b)$.
Sur *Texas* : dans le menu DISTRIBUTION on calcule « Normalcdf (*a*, *b*,*m*, σ) ».
Sur *Casio* : on sélectionne « Normal CD » et on note « $\text{lower} = a$ » et « $\text{upper} = b$ ».
(Pour calculer $P(X \leqslant b)$ on choisit une très petite valeur de *a*, par exemple −10 000.)
Inversement, on peut déterminer *t* tel que $P(X \leqslant t) \leqslant c$.
Sur *Texas* : on calcule à présent « invN(*c*, *m*, σ) » ; sur *Casio* : on affiche « invN » avec $\text{area} = c$ ».

Exercice n°1Exercice n°2Exercice n°3

2. Comment passer de la loi normale *N*(*m* ; σ) à la loi normale centrée réduite *N*(0 ; 1) ?

Propriété fondamentale

Si une variable aléatoire *X* suit la loi normale $N(m\,;\,\sigma)$ alors la variable aléatoire $T = \frac{X-m}{\sigma}$ suit la loi normale centrée réduite $N(0\,;\,1)$.

Justification

On sait que $E(aX+b) = aE(X)+b$.
Alors $E(T) = E(\frac{X}{\sigma} - \frac{m}{\sigma}) = \frac{1}{\sigma}(E(X) - m) = 0$.
De même $V(aX+b) = a^2V(X)$.
Alors $V(T) = V(\frac{X}{\sigma} - \frac{m}{\sigma}) = \frac{1}{\sigma^2}V(X) = \frac{1}{\sigma^2} \times \sigma^2 = 1$.

Exemple

Le délai de livraison d'un matériel, compté en jours, est une variable aléatoire *X* qui suit la loi normale $N(30\,;\,5)$.
Quelle est la probabilité que le délai soit compris entre 22 et 38 jours ?
On pose $T = \frac{X-30}{5}$. Alors *T* suit la loi normale centrée réduite $N(0\,;\,1)$.
$22 \leqslant X \leqslant 38$ **équivaut à** $-8 \leqslant X - 30 \leqslant 8$ **ou** $-1,6 \leqslant \frac{X-30}{5} \leqslant 1,6$**, soit**
$P(22 \leqslant X \leqslant 38) = P(-1,6 \leqslant T \leqslant 1,6) = 2\Pi(1,6) - 1$.
Sur la table, on lit $\Pi(1,6) = 0,945\,2$**, d'où** $P(22 \leqslant X \leqslant 38) = 2 \times 0,945\,2 - 1 = 0,890\,4$.
Il y a 89,04 % de chances que le délai de livraison soit compris entre 22 et 38 jours.

Exercice n°4

3. Quelles sont les probabilités de $X \in [m - \sigma\,;\, m + \sigma]$ **et** $X \in [m - 2\sigma\,;\, m + 2\sigma]$ **?**

Probabilité de $X \in [m - \sigma\,;\, m + \sigma]$

La variable aléatoire *X* suit la loi normale $N(m\,;\,\sigma)$**.**

On pose $T = \frac{X - m}{\sigma}$ **. Alors *T* suit la loi normale centrée réduite.**

$P(m - \sigma \leqslant X \leqslant m + \sigma)$ **équivaut à** $P(-\sigma \leqslant X - m \leqslant +\sigma)$ **ou** $P(-1 \leqslant \frac{X-m}{\sigma} \leqslant 1)$**.**

Soit $P(-1 \leqslant T \leqslant 1)$**.**

Donc $P(m - \sigma \leqslant X \leqslant m + \sigma) = P(-1 \leqslant T \leqslant 1) = 2\Pi(1) - 1$**.**

Sur la table on lit $\Pi(1) = 0,841\,3$**.**

D'où $P(m - \sigma \leqslant X \leqslant m + \sigma) = 2 \times 0,841\,3 - 1 = 0,682\,6$**.**

On peut estimer que l'intervalle $[m - \sigma\,;\, m + \sigma]$ **regroupe environ 68 % des valeurs de *X*.**

Probabilité de $X \in [m - 2\sigma\,;\, m + 2\sigma]$

$P(m - 2\sigma \leqslant X \leqslant m + 2\sigma)$ **équivaut à** $P(-2\sigma \leqslant X - m \leqslant +2\sigma)$ **ou** $P(-2 \leqslant \frac{X-m}{\sigma} \leqslant 2)$

.

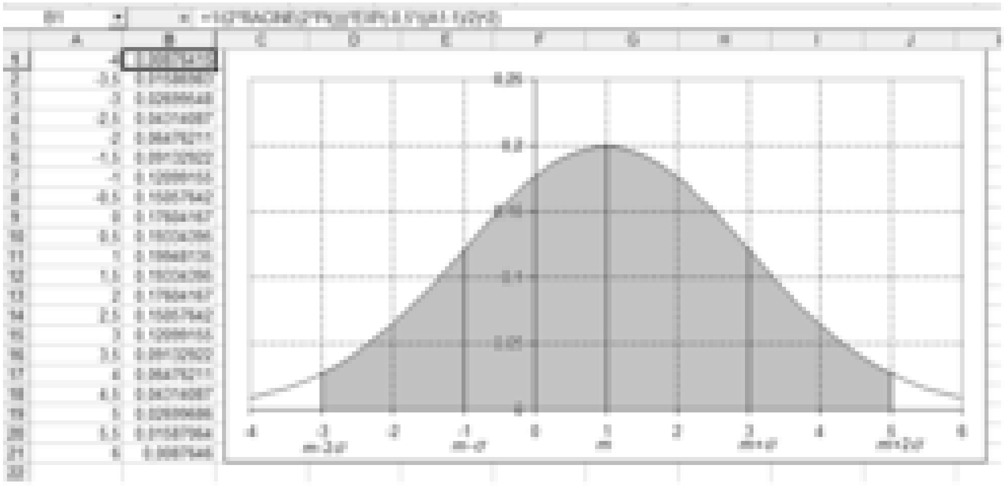
Soit $P(-2 \leqslant T \leqslant 2)$**.**

Sur la table on lit $\Pi(2) = 0,977\,2$**.**

D'où $P(m - 2\sigma \leqslant X \leqslant m + 2\sigma) = 2 \times 0,977\,2 - 1 = 0,954\,4$**.**

On peut estimer que l'intervalle $[m - 2\sigma\,;\, m + 2\sigma]$ **regroupe environ 95 % des valeurs de *X*.**

Ces deux intervalles sont appelés intervalles de normalité.



Exercice n°5

4. Comment approcher la loi binomiale par une loi normale ?

Conditions de correction

En classe de Première, on a établi qu'une variable aléatoire *X* qui suit la loi binomiale $B(n\,;\,p)$ a pour espérance $E(X) = np$ et pour variance $V(X) = \sqrt{np(1-p)}$.

Si *n* est grand et *p* voisin de 0,5 alors la loi $B(n\,;\,p)$ est très proche de la loi normale $N(np\,;\,\sqrt{np(1-p)})$.

Dans la pratique on retient $n \geqslant 30$, $np \geqslant 5$ et $\sqrt{np(1-p)} \geqslant 5$.
Le but est la simplification des calculs.

Exemple

On lance 40 fois une pièce de monnaie. La variable aléatoire *X*, égale au nombre de « pile », suit la loi binomiale $B(40\,;\,0,5)$.

$P(25 \leqslant X \leqslant 30) = P(X=25) + P(X=26) + P(X=27) + P(X=28) + P(X=29) + P(X=30)$.

$E(X) = 40 \times 0,5 = 20$ **et** $V(X) = \sqrt{40 \times 0,5 \times 0,5} = \sqrt{10}$.

En substituant à la variable aléatoire *X* la variable aléatoire *Z* suivant la loi normale $N(20\,;\,\sqrt{10})$, $P(25 \leqslant Z \leqslant 30)$ ne prend pas en compte les valeurs 25 et 30, car

$P(X=25)=P(X=30)=0$.

Pour prendre en compte ces deux valeurs, on calcule $P(24,5 \leqslant Z \leqslant 30,5)$.

En posant $T=\frac{Z-20}{\sqrt{10}}$ **,** ***T*** **suit la loi normale centrée réduite** $N(0\,;\,1)$.

Alors $P(24,5 \leqslant Z \leqslant 30,5)=P(4,5 \leqslant Z-20 \leqslant 10,5)=P(1,423 \leqslant T \leqslant 3,320)$.

Corrections de continuité à apporter

Si ***X*** **suit la loi binomiale** $B(n\,;\,p)$ **et** ***Z*** **la loi normale** $N(np\,;\,\sqrt{np(1-p)})$ **alors :**

$P(X=a) \simeq P(a-0,5 \leqslant Z \leqslant a+0,5)$ **;** $P(X \leqslant a) \simeq P(Z \leqslant a+0,5)$ **;**

$P(a \leqslant X \leqslant b) \simeq P(a-0,5 \leqslant Z \leqslant b+0,5)$.

Exercice n°6Exercice n°7

5. Comment déterminer un intervalle de confiance à ***t*** **% ?**

Définition

La variable aléatoire ***X*** **suit la loi normale** $N(m\,;\,\sigma)$.

Déterminer un intervalle de confiance à ***t*** **%, c'est déterminer le nombre réel positif** ***a*** **tel que** $P(m-a \leqslant X \leqslant m+a)=\frac{t}{100}$.

On a vu précédemment que :

- $P(m-\sigma \leqslant X \leqslant m+\sigma)=0,68$ **;**
- $P(m-2\sigma \leqslant X \leqslant m+2\sigma)=0,95$ **;**
- $P(m-3\sigma \leqslant X \leqslant m+3\sigma)=0,99$.

Exemple

La variable aléatoire ***X*****, égale à la durée de vie en années d'un composant électronique, suit la loi normale** $N(4,4\,;\,1,63)$ **. On veut déterminer** ***a*** **tel que** $P(4,4-a \leqslant X \leqslant 4,4+a)=0,9$.

En posant $T=\frac{X-4,4}{1,63}$ **on obtient** $P(-\frac{a}{1,63} \leqslant T \leqslant \frac{a}{1,63})=2\Pi(\frac{a}{1,63})-1=0,9$.

$2\Pi(\frac{a}{1,63})-1=0,9$ **équivaut à** $\Pi(\frac{a}{1,63})=\frac{1+0,9}{2}=0,95$.

Par lecture inverse de la table on obtient $\frac{a}{1,63}=1,645$ **, soit** $a=1,645 \times 1,63=2,681\,35$.

$[4,4-2,7\,;\,4,4+2,7]=[1,7\,;\,7,1]$, $0,7\times 12=8,4$ et $0,1\times 12=1,2$.

La durée de vie d'un composant est comprise entre 1 an 8 mois et demi et 7 ans 1 mois, avec un indice de confiance de 90 %.

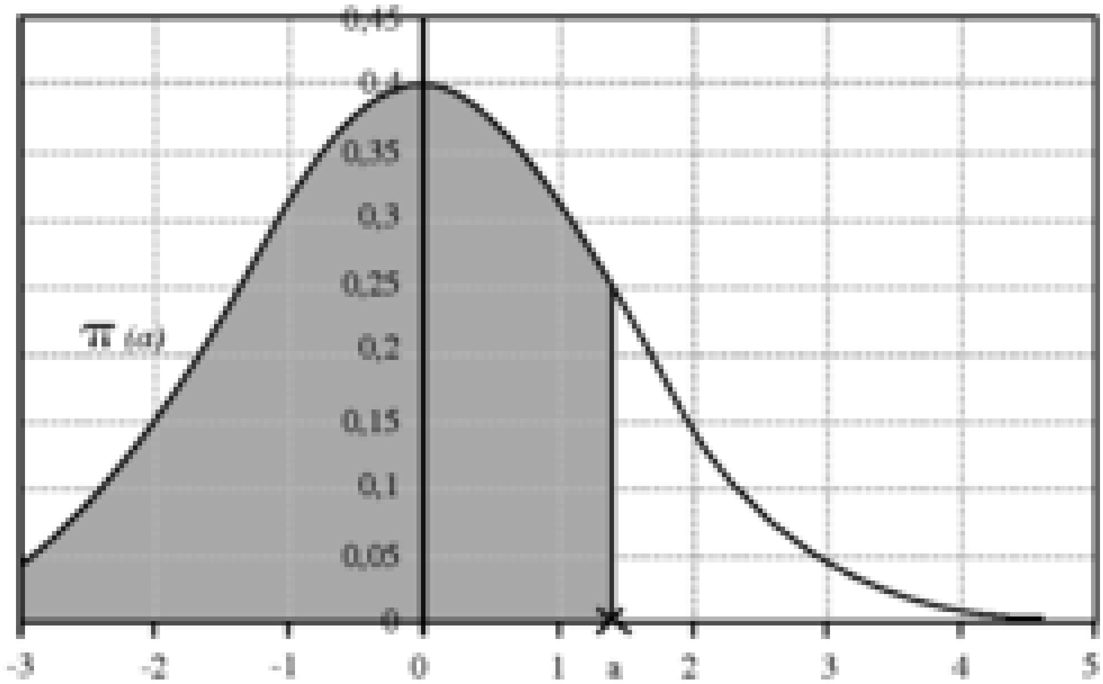
Exercice n°8

À retenir

Loi normale et loi normale centrée réduite

Pour passer d'une variable aléatoire *X* suivant la loi normale $N(m\,;\,\sigma)$ à la variable aléatoire *T* suivant la loi normale centrée réduite $N(0\,;\,1)$, on pose $T=\frac{X-m}{\sigma}$. Les valeurs de la fonction de répartition $\Pi(a)=P(T\leqslant a)$ se lisent alors sur une table.

Zoom



Écritures à savoir retrouver graphiquement

- $P(T \geqslant a) = P(T \leqslant -a) = 1 - P(T \leqslant a)$, d' où $\Pi(-a) = 1 - \Pi(a)$;
- $P(a \leqslant X \leqslant b) = P(T \leqslant b) - P(T \leqslant a) = \Pi(b) - \Pi(a)$;
- $P(-a \leqslant X \leqslant a) = 2P(T \leqslant a) - 1 = 2\Pi(a) - 1$.

Approche d'une loi binomiale par une loi normale

Si *X* suit la loi binomiale $B(n\,;\,p)$, alors *Z* suit la loi normale $N(np\,;\,\sqrt{np(1-p)})$.

- $P(X = a) \simeq P(a - 0,5 \leqslant Z \leqslant a + 0,5)$;
- $P(X \leqslant a) \simeq P(Z \leqslant a + 0,5)$;
- $P(a \leqslant X \leqslant b) \simeq P(a - 0,5 \leqslant Z \leqslant b + 0,5)$.

Intervalles de confiance

- $P(m - \sigma \leqslant X \leqslant m + \sigma) = 0,68$;
- $P(m - 2\sigma \leqslant X \leqslant m + 2\sigma) = 0,95$;
- $P(m - 3\sigma \leqslant X \leqslant m + 3\sigma) = 0,99$.