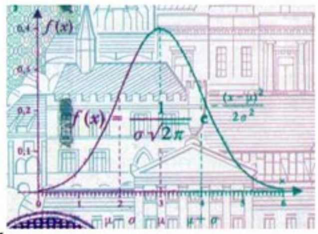
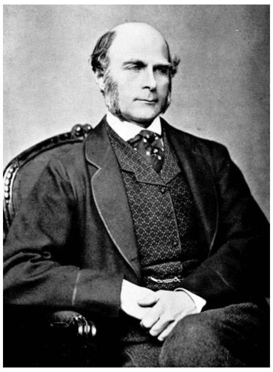
LOIS À DENSITÉ (PARTIE 3) : LES LOIS NORMALES



Un billet de dix Deutsche Mark, avec Carl Friedrich Gauss (1777-1855).



← Laplace (1749 – 1827) →



Galton (1822 – 1911)

← Quételet (1796 – 1874) →

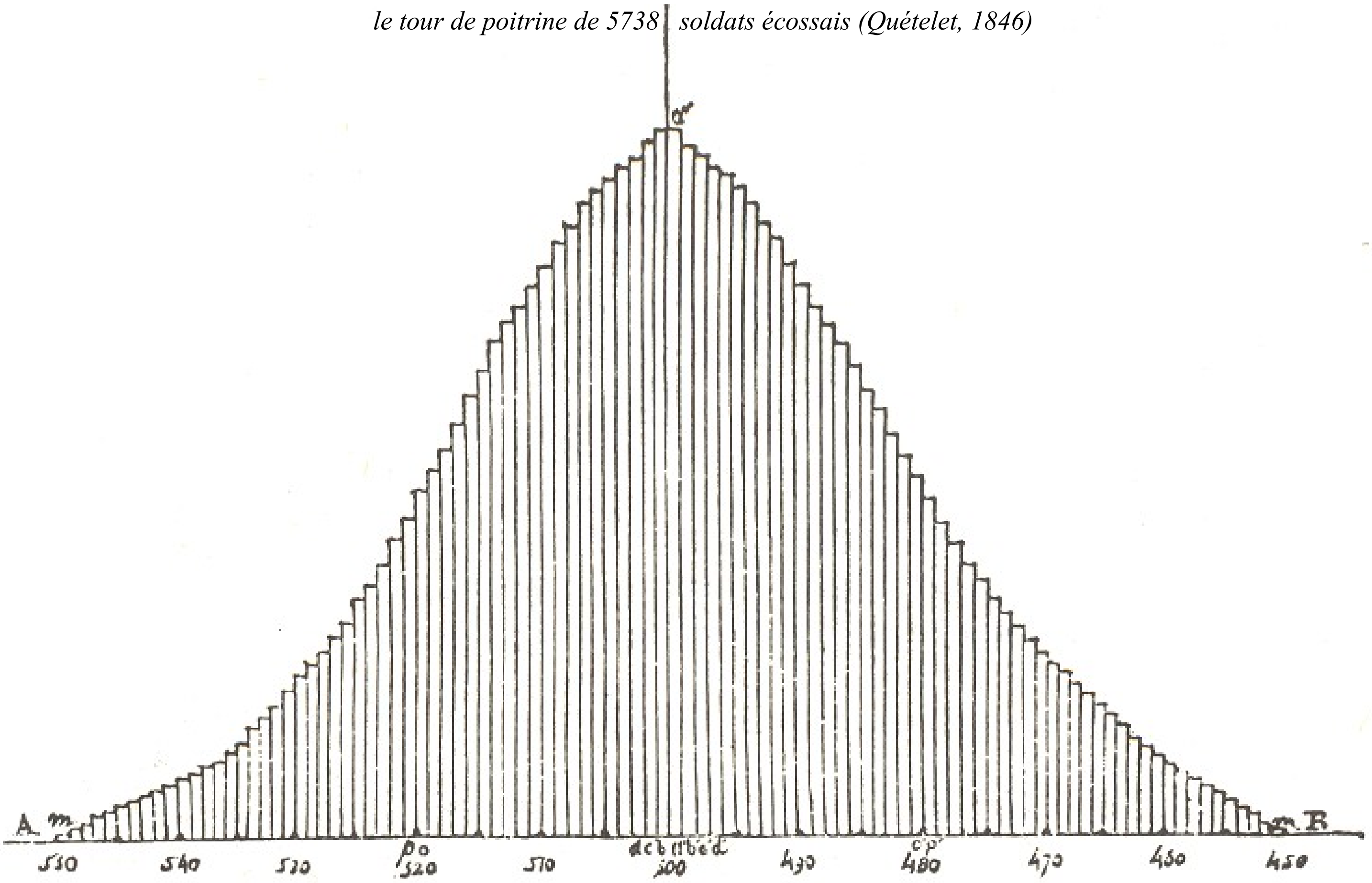


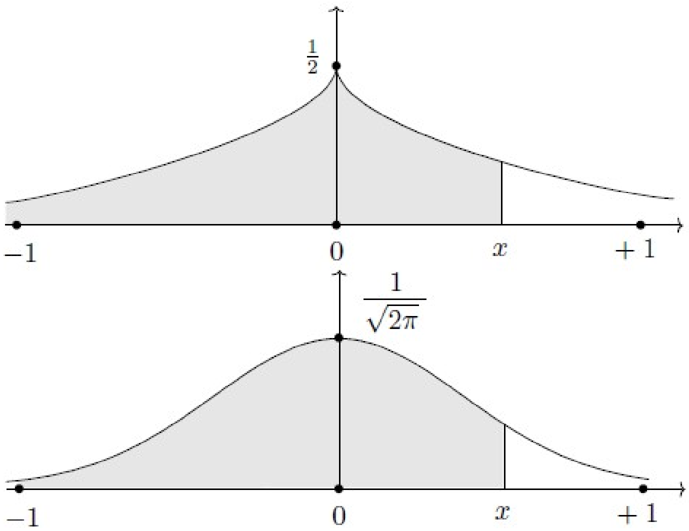
Pearson (1857 – 1936)

De Moivre (1667 – 1754)

Gauss (1777 – 1855)

le tour de poitrine de 5738 soldats écossais (Quételet, 1846)





Première et deuxième lois de Laplace

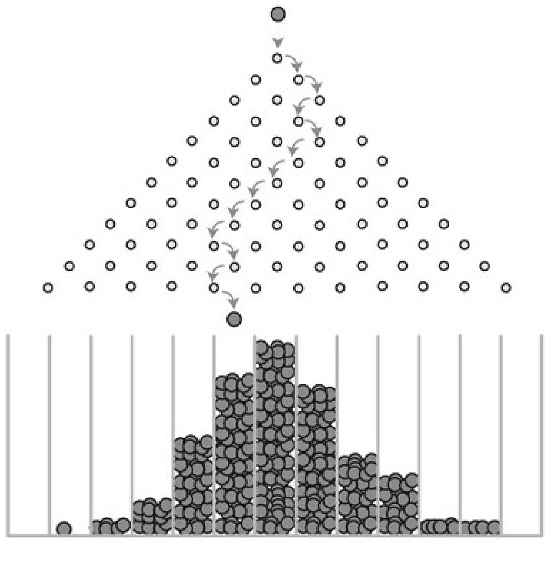


Planche de Galton

I. Loi normale centrée réduite

Admis

THÉORÈME DE MOIVRE-LAPLACE.

Soit $p \in \ ]0;1[$ (réel fixé). Soit X_n une v.a.r. suivant la loi binomiale $\mathcal{B}$ (n ; p), où $n \in \mathbb{N}^*$.

Alors, pour tous nombres réels a et b tels que $a < b$:

$$\lim_{n \to +\infty} p(a \leqslant Z_n \leqslant b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

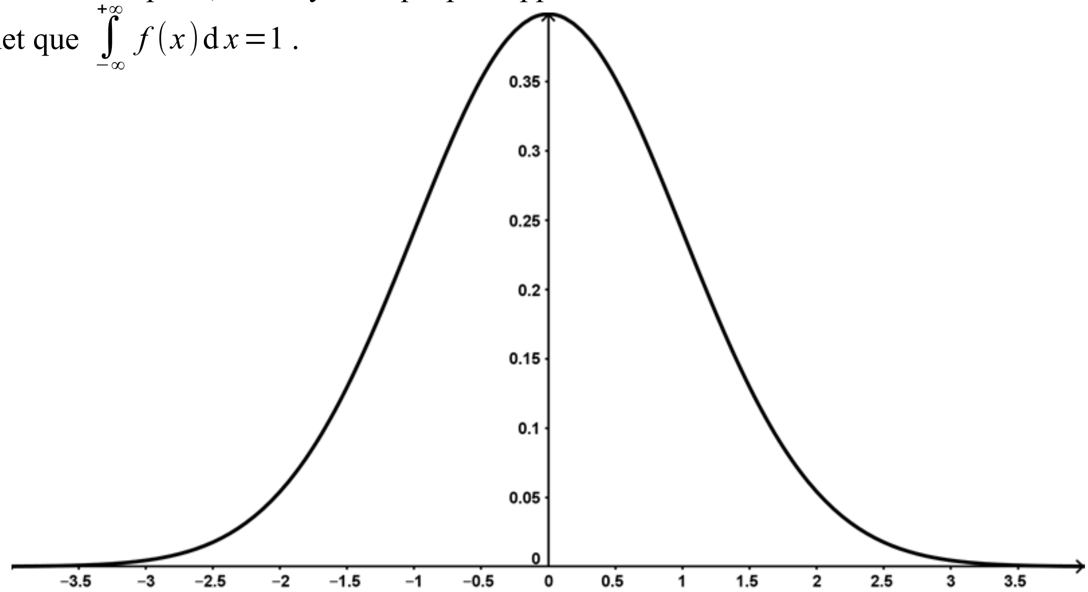
où Z_n est la v.a.r. définie par $Z_n = \dfrac{X_n - n\,p}{\sqrt{n\,p(1-p)}}$.

Autrement dit, ce théorème justifie que sous certaines conditions sur les paramètres n et p, la probabilité d'un événement associé à une loi binomiale peut être approché par la probabilité d'un événement associé à la loi normale centrée réduite.
On parle alors d'***approximation d'une loi binomiale par une loi normale***.

Remarque : en fait, ce phénomène n'est pas spécifique à la loi binomiale… une loi normale intervient dans « la modélisation de nombreux phénomènes aléatoires possédant de nombreuses causes indépendantes dont les effets s'ajoutent, sans que l'une d'elles soit dominante ». Donc dans une situation de répétition d'expériences identiques et indépendantes. [voir compléments]

DÉFINITION. On dit qu'une variable aléatoire X suit ***la loi normale centrée réduite*** si sa densité est la fonction définie sur $\mathbb{R}$ par $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. On note X ~ $\mathcal{N}$(0 ; 1).

Remarques :
- f est continue et paire, donc symétrique par rapport à l'axe des ordonnées.
- on admet que $\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx = 1$.



• l'espérance de X_n est $n\,p$, notée μ, et son écart-type est $\sqrt{np(1-p)}$, noté σ.
Alors, d'après le théorème de Moivre-Laplace :

si n est assez grand, on peut approcher $p\left(a \leqslant \dfrac{X_n - \mu}{\sigma} \leqslant b\right)$ par $p(a \leqslant T \leqslant b)$ où T ~ $\mathcal{N}$(0 ; 1).

PROPRIÉTÉ.

Soit X une v.a.r. suivant la loi normale centrée réduite. Alors : $E(X)=0$; $V(X)=1$; $\sigma(X)=1$.

Démonstrations : $V(X)=1$ est admis, donc $\sigma(X)=1$ également. Démontrons que $E(X)=0$:

PROPRIÉTÉS.

Soit T une variable aléatoire suivant la loi normale centrée réduite $\mathcal{N}(0;1)$.

- Pour tout réel $u>0$: $p(T\leqslant -u)=p(T\geqslant u)$ et $p(-u\leqslant T\leqslant u)=2\,p(T\leqslant u)-1$.
- $p(T\geqslant 0)=p(T\leqslant 0)=\dfrac{1}{2}$.

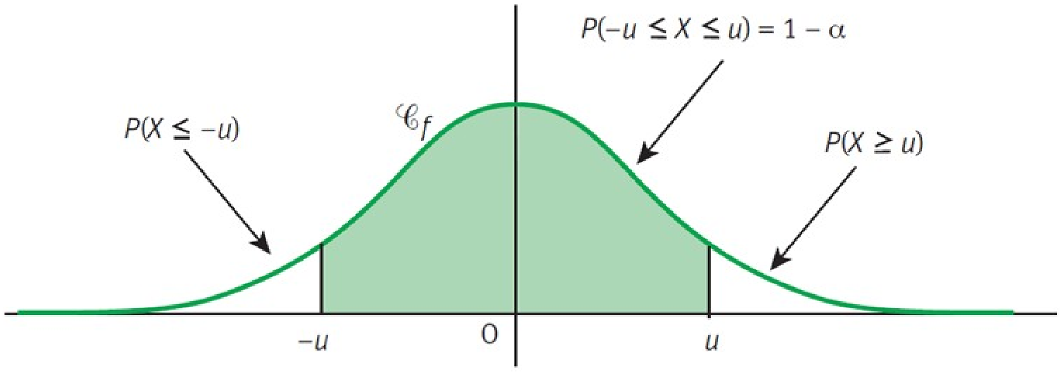
Démonstration : • $p(T\leqslant -u)=p(T\geqslant u)$ est une conséquence de la symétrie de la courbe, tout comme $p(T\geqslant 0)=p(T\leqslant 0)=\dfrac{1}{2}$.

• $p(-u\leqslant T\leqslant u)=$

ROC

PROPRIÉTÉ.

Soit T une variable aléatoire suivant la loi normale centrée réduite $\mathcal{N}(0;1)$.
Pour tout $\alpha \in]0;1[$, il existe un unique réel $u_\alpha > 0$ tel que $p(-u_\alpha \leqslant T \leqslant u_\alpha) = 1-\alpha$.



Démonstration :

Soit T une variable aléatoire suivant la loi normale centrée réduite $\mathcal{N}(0;1)$. Soit $\alpha \in]0;1[$.

1. Démontrer que pour tout réel x strictement positif :
$p(-x \leqslant T \leqslant x) = 2F(x)$ où F est la primitive de f qui s'annule en 0.

2. Démontrer que : $0 < \frac{1-\alpha}{2} < \frac{1}{2}$.

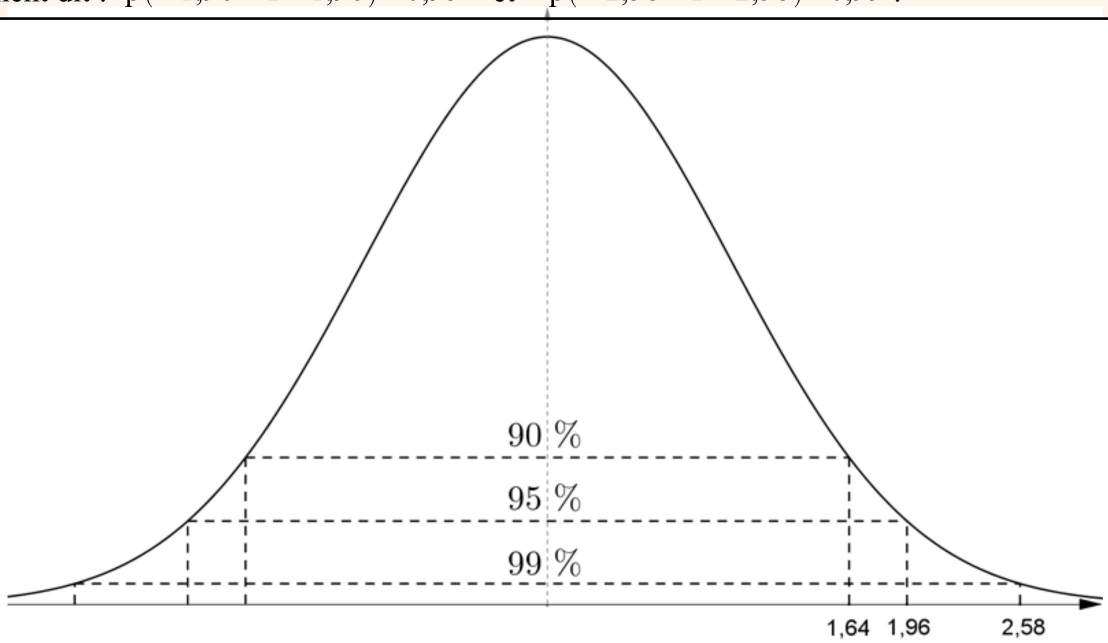
3. a) On admet que F est continue et strictement croissante sur $[0;+\infty[$.
Démontrer qu'il existe un unique réel strictement positif, noté u_α, tel que : $F(u_\alpha) = \frac{1-\alpha}{2}$.

b) En déduire qu'il existe un unique réel $u_\alpha > 0$ tel que $p(-u_\alpha \leqslant T \leqslant u_\alpha) = 1-\alpha$.

CAS PARTICULIERS. (À CONNAÎTRE[1])

$u_{0,05} \approx 1,96$ et $u_{0,01} \approx 2,58$.

Autrement dit : $p(-1,96 \leqslant T \leqslant 1,96) \approx 0,95$ et $p(-2,58 \leqslant T \leqslant 2,58) \approx 0,99$.



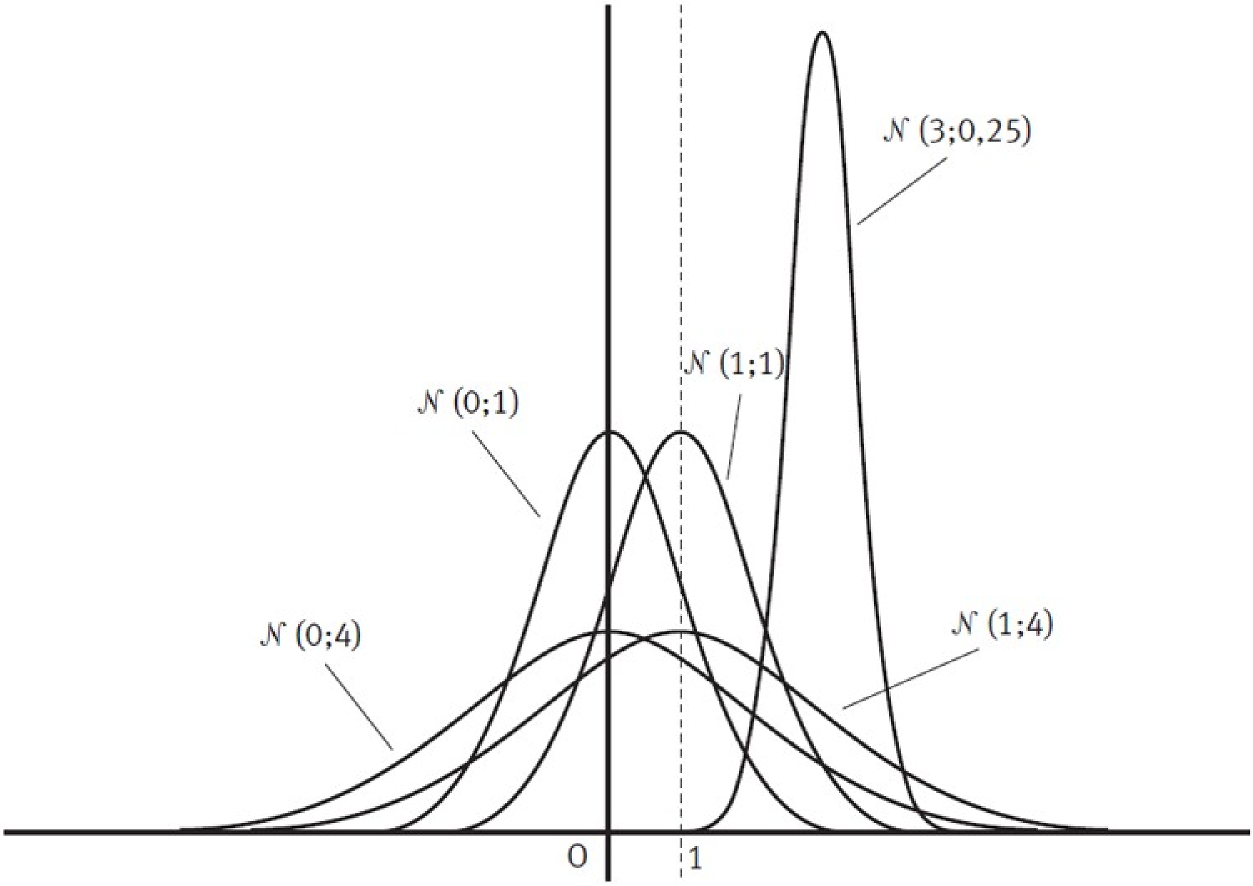
1 Nous verrons un peu plus loin comment retrouver ces valeurs à la calculatrice.

II. Loi normale

DÉFINITION. On dit qu'une variable aléatoire X suit ***la loi normale d'espérance*** μ ***et d'écart-type σ*** si la v.a.r. $\frac{X-\mu}{\sigma}$ suit la loi normale centrée réduite. On note X ~ $\mathcal{N}(\mu, \sigma^2)$.

On observe que chaque courbe est symétrique par rapport à la droite d'équation $x=\mu$.

L'écart-type a un effet sur la forme de « la cloche ».
Plus σ est petit, plus elle est « haute » et plus les valeurs sont resserrées autour de la moyenne μ.



__Remarque__ : si X ~ $\mathcal{N}(\mu, \sigma^2)$ alors on a $E(X)=\mu$ et $\sigma(X)=\sigma$.

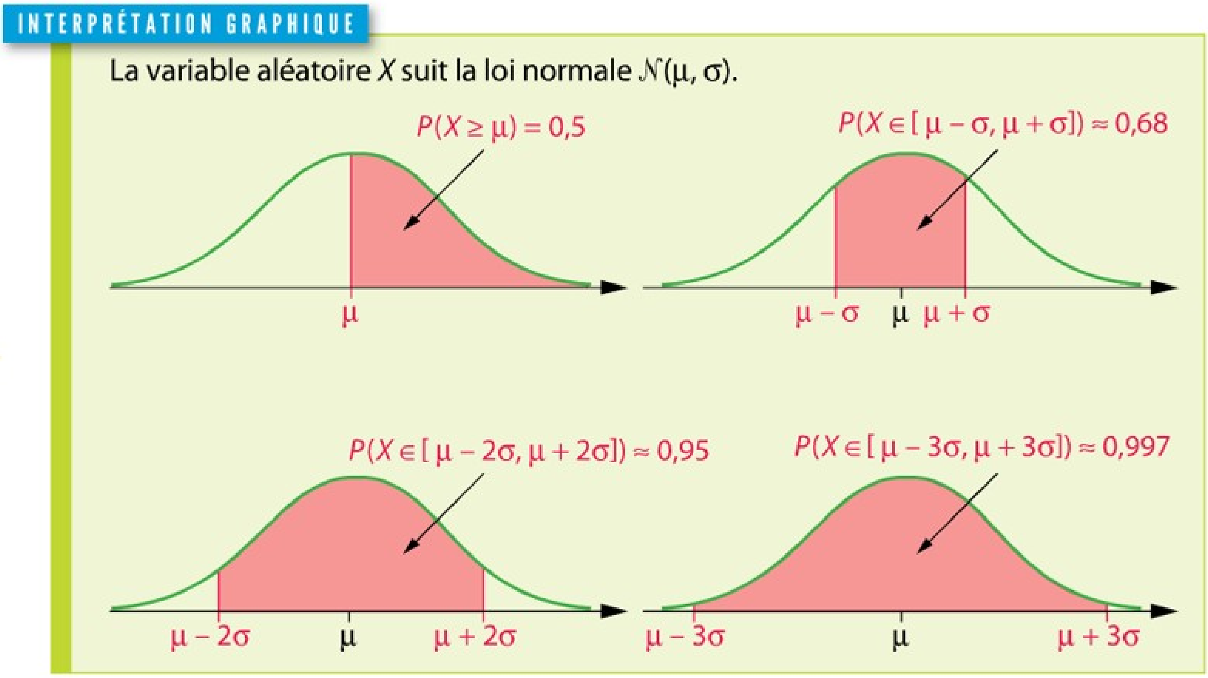
En effet : $T=\frac{X-\mu}{\sigma}$ suit la loi normale centrée réduite, donc $X=\sigma T+\mu$
et donc $E(X)=E(\sigma T+\mu)=\sigma E(T)+\mu=\sigma\times 0+\mu=\mu$.
De même : $V(X)=V(\sigma T+\mu)=\sigma^2 V(T)=\sigma^2\times 1=\sigma^2$ d'où $\sigma(X)=\sigma$.

PROPRIÉTÉS. Soit $X \sim \mathcal{N}(\mu, \sigma^2)$.

Alors :
- $p(\mu-\sigma \leqslant X \leqslant \mu+\sigma) \approx 0{,}68$
- $p(\mu-2\sigma \leqslant X \leqslant \mu+2\sigma) \approx 0{,}95$
- $p(\mu-3\sigma \leqslant X \leqslant \mu+3\sigma) \approx 0{,}997$.

<u>Démonstrations</u> : On pose $T \sim \mathcal{N}(0,1)$.

$p(\mu-\sigma \leqslant X \leqslant \mu+\sigma) = p\left(-1 \leqslant \dfrac{X-\mu}{\sigma} \leqslant 1\right) = p(-1 \leqslant T \leqslant 1) \approx 0{,}68$. Idem pour les autres.



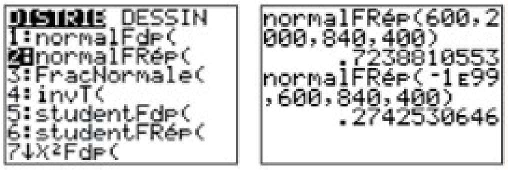
Source : Collection Sigma, éditions Foucher, avril 2013

III. Avec la calculatrice

LOIS DE PROBABILITÉ AVEC LA CALCULATRICE

Loi normale avec une calculatrice TI

On accède au menu « distribution » par 2nde/Distrib (ou DISTR).
Pour calculer $P(a \leq X \leq b)$, on utilise normalFRép ou normalcdf.
Si l'une des bornes a ou b est absente, on la remplace par −1E99 ou 1E99.



Loi normale avec une calculatrice Casio

On se place dans le Menu STAT.
On accède à la loi normale par DIST/NORM.
Pour calculer $P(a \leq X \leq b)$, on utilise Ncd.
Si l'une des bornes a ou b est absente, on la remplace par −1E99 ou 1E99.

Loi binomiale avec une calculatrice TI

On accède au menu « distribution » par 2nde/Distrib (ou DISTR).
Pour calculer $P(X = k)$, on utilise l'instruction binomFdp ou binompdf.
Pour calculer $P(X \leq k)$, on utilise l'instruction binomPRép ou binomcdf.

Loi binomiale avec une calculatrice CASIO

On se place dans le Menu STAT.
On accède à la loi binomiale par DIST/BINM.
Pour calculer $P(X = k)$, on utilise l'instruction Bpd.
Pour calculer $P(X \leq k)$, on utilise l'instruction Bcd.

Source : Collection Sigma, éditions Foucher, avril 2013

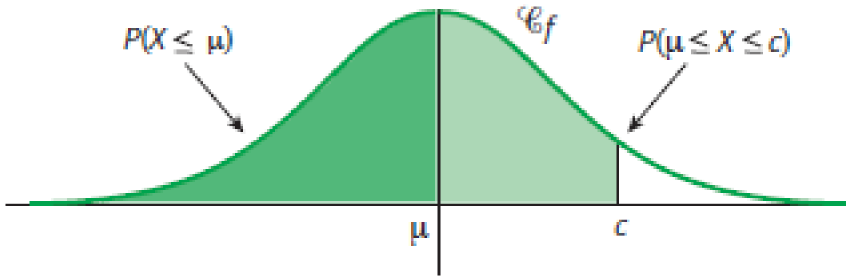
Si vous préférez une vidéo : https://youtu.be/qh31hHKsaIA

ATTENTION : **Bpd** permet de calculer $p(X=k)$ pour X qui suit une loi binomiale.
mais **Npd** permet de calculer $f(k)$ avec f la densité d'une loi normale !
Pour calculer $p(Z=k)$ avec Z qui suit une loi normale, il faut utiliser **Ncd**, avec Lower = k et Upper = k.

Pour calculer $p(X \leqslant c)$, on a donc deux méthodes :

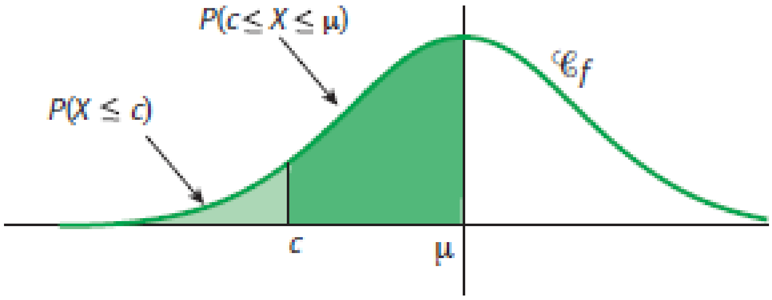
• Utiliser l'approximation $p(X \leqslant c) \approx p(-10^{99} \leqslant X \leqslant c)$.

• Utiliser les propriétés de symétrie de la courbe, selon la position de c par rapport à μ…



Si $c \geq \mu$:

$$P(X \leq c) = P(X < \mu) + P(\mu \leq X \leq c) = \frac{1}{2} + P(\mu \leq X \leq c)$$



Si $c \leq \mu$:

$$P(X \leq c) = P(X \leq \mu) - P(c < X \leq \mu) = \frac{1}{2} - P(c \leq X \leq \mu).$$

<u>Exercice III.1</u> : soit X une v.a.r. suivant la loi normale $\mathcal{N}\left(46{,}75\,;6{,}23^2\right)$. Avec la calculatrice, calculer :

1. $p(40 \leqslant X \leqslant 60)$

2. $p(X \leqslant 60)$ de deux façons différentes

3. $p(X > 40)$ de deux façons différentes

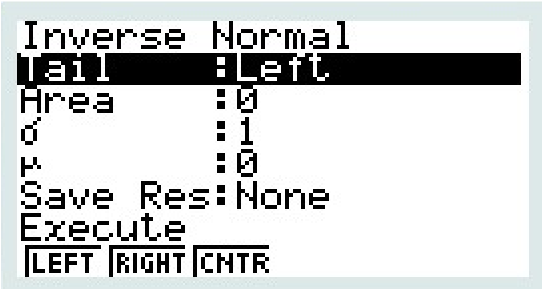
4. $p(X \geqslant 60)$ de deux façons différentes.

Exercice III.2 : soit X une v.a.r. suivant la loi normale $\mathcal{N}\left(46{,}75\,;6{,}23^2\right)$.
Avec l'aide ci-dessous, déterminer une valeur approchée du réel k tel que :

1. $p(X\leqslant k)=0{,}75$ 2. $p(X<k)=0{,}95$ 3. $p(X\geqslant k)=0{,}95$.

A I D E

• **Sur CASIO** : utiliser le menu STAT
Menu DIST puis NORM et InvN. Un menu « Inverse Normal » s'ouvre :



Tail vous permet de choisi *Left*, *Right* ou *Central*.
Cela dépend de si vous voulez déterminer *k* tel que $p(X\leqslant k)$, $p(X\geqslant k)$ ou $p(-k\leqslant X\leqslant k)$.

Area correspond à la probabilité que vous souhaitez.

μ et σ sont bien sûr les paramètres de la loi normale $\mathcal{N}(\mu, \sigma^2)$.

Pour voir en vidéo trois exemples avec cette méthode : https://youtu.be/pWehWIpVBPE

• **Sur TI** : https://youtu.be/aipNt2M-c80

B I L A N À R E T E N I R

Soit X_n une v.a.r. suivant la loi binomiale $\mathcal{B}(n\ ;\ p)$.
D'après le théorème de Moivre-Laplace :

- si n est assez grand, on peut approcher $p\left(a \leqslant \frac{X_n - \mu}{\sigma} \leqslant b\right)$ par $p(a \leqslant T \leqslant b)$ où $T \sim \mathcal{N}(0\ ;\ 1)$.

- si n est assez grand, on peut approcher $p(a \leqslant X_n \leqslant b)$ par $p(a \leqslant Z \leqslant b)$ où $Z \sim \mathcal{N}(\mu\ ;\ \sigma^2)$.

Que sait-on aujourd'hui de l'erreur commise en approchant une loi binomiale par une loi normale ?
(hors programme)

Théorème d'Uspensky (1937)

Soient $X_n \sim \mathcal{B}(n\ ;\ p)$ et $Z \sim \mathcal{N}(\mu\ ;\ \sigma^2)$, avec $\mu = np$ et $\sigma = \sqrt{np(1-p)}$
.

Si $np(1-p) \geqslant 25$, autrement dit si $\sigma \geqslant 5$, alors :
pour tous réels a et b tels que $a < b$,

$$|p(a \leqslant X_n \leqslant b) - p(a \leqslant Z \leqslant b)| \leqslant \frac{0{,}588}{\sigma}\ .$$

Théorème de Berry-Esseen-Asof (1941 puis 2011)

Soient $X_n \sim \mathcal{B}(n\ ;\ p)$ et $Z \sim \mathcal{N}(\mu\ ;\ \sigma^2)$, avec $\mu = np$ et $\sigma = \sqrt{np(1-p)}$.
Pour tous réels a et b tels que $a < b$:

$$|p(a \leqslant X_n \leqslant b) - p(a \leqslant Z \leqslant b)| \leqslant \frac{C\left(p^2 + (1-p)^2\right)}{\sigma} \text{ avec } C = 0{,}9568\ .$$

Remarque : les valeurs calculées du coefficient C ont chuté rapidement au cours des années.
La valeur originale d'Esseen (mathématicien suédois) date de 1941 et était 15,18.
En 1972, Van Beek trouve 1,5764.
En 1986, Shiganov trouve 1,531.
En 2007 et 2008, Shevtsova trouve 1,4112 et 1,401.
En 2009, Ilya Tyurin trouve 1,1788.
En 2009, Victor Korolev et Shevtsova trouvent 1,0258.
En 2010, Ilya Tyurin trouve 0,957.
En 2011, Asof trouve 0,9568.

<u>Exercice III.3</u> :

1. Déterminer, selon les valeurs de p, lequel des deux théorèmes est le meilleur.

2. Jusqu'à quelle année le théorème d'Uspensky était toujours meilleur que celui de Berry-Esseen ?

IV. Approcher du discret par du continu… Attention !

Si X suit une loi binomiale de paramètres n et p, on peut vouloir trouver une approximation de $\mathrm{p}(X \leqslant 45)$ à l'aide du théorème de Moivre-Laplace.
La plupart des corrigés sur internet ou dans les manuels scolaires proposent alors d'appliquer le théorème de Moivre-Laplace, en approchant la loi binomiale par la loi normale Z correspondante, c'est-à-dire en remplaçant $\mathrm{p}(X \leqslant 45)$ par $\mathrm{p}(Z \leqslant 45)$.

1. Pourquoi cela n'est finalement pas très judicieux ?

2. Déterminer les paramètres de la loi normale Z qui approche X si n est grand.

3. Prenons $n=255$ et $p=0{,}02$ (situation pas vraiment rare).

a) Calculer $\mathrm{p}(X \leqslant 5)$ en utilisant la loi binomiale.

b) Calculer $\mathrm{p}(Z \leqslant 5)$ et $\mathrm{p}(0 \leqslant Z \leqslant 5)$.

c) Qu'observe-t-on ?

Remarque : on peut démontrer que l'erreur commise en négligeant $\mathrm{p}(Z<0)$ *est inférieure à 0,0127.*

[voir sur mon site]