Lois normales. Intervalle de fluctuation. Estimation.

1. Loi centrée réduite

1.1. Nouvelle fonction de densité sur ℝ

Soit φ la fonction définie sur ℝ par $\boxed{\varphi(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x^2}}$.

φ est continue et dérivable sur ℝ.

$\left(\mathrm{e}^{-\frac{1}{2}x^2}\right)'=-\frac{1}{2}\times 2x\times \mathrm{e}^{-\frac{1}{2}x^2}$.

Donc, pour tout x réel, $\varphi'(x)=\frac{-x}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x^2}$.

Pour tout nombre réel x, le signe de $\varphi'(x)$ est le signe de $(-x)$.

$\lim\limits_{x\to-\infty}-\frac{1}{2}x^2=-\infty$ et $\lim\limits_{X\to-\infty}\mathrm{e}^X=0$ donc $\lim\limits_{x\to-\infty}\varphi(x)=0$.

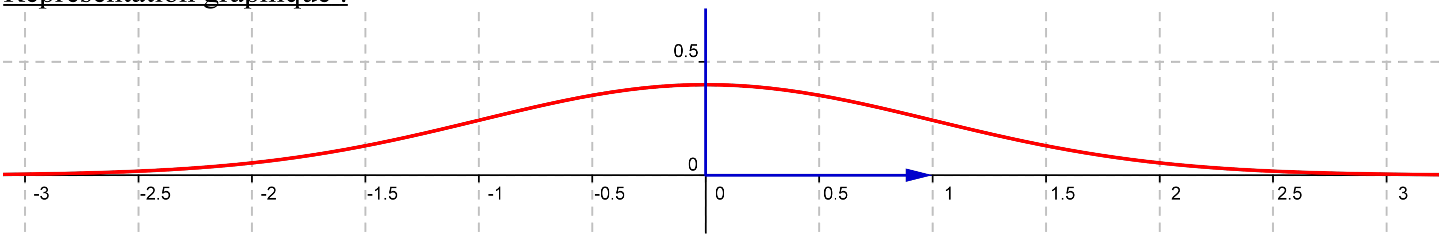
De même, $\lim\limits_{x\to+\infty}\varphi(x)=0$.

Tableau de variations :

x	$-\infty$		0		$+\infty$
Signe de φ ' (x)		$+$	0	$-$	
Variations de φ(x)	0	↗	φ(0)	↘	0

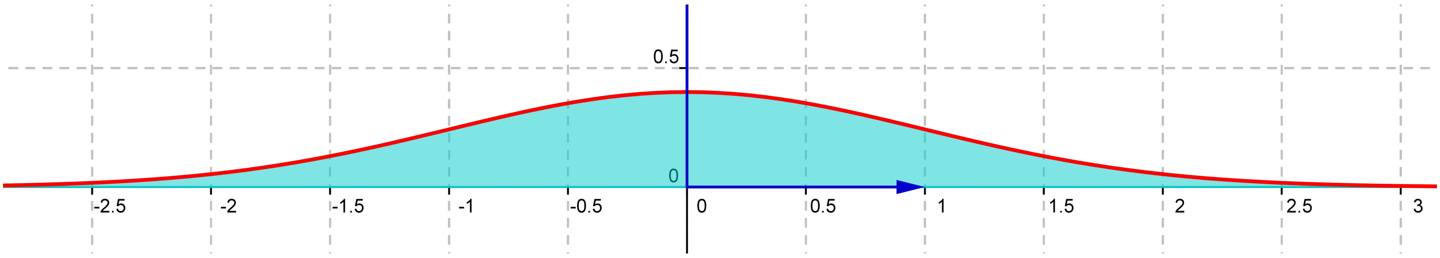
$\varphi(0)=\frac{1}{\sqrt{2\pi}}\approx 0,4$

Représentation graphique :



Remarques :

La fonction φ est paire donc la courbe représentative de φ sur ℝ est symétrique par rapport à l'axe $(y'y)$.
On admet que l'aire en unité d'aire de la partie du plan comprise entre la courbe et l'axe des abscisses est égale à 1.



On peut noter : $\boxed{\int_{-\infty}^{+\infty} \varphi(x)\,dx = 1}$

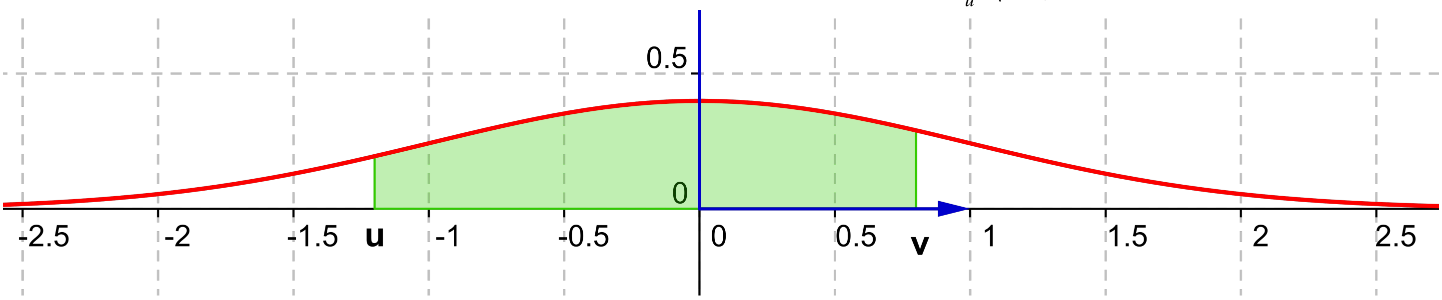
Conséquence :

φ est continue et positive sur ℝ et $\int_{-\infty}^{+\infty} \varphi(x)\,dx = 1$ donc φ est une densité de probabilité sur ℝ.

1.2. Valeurs remarquables

On dit qu'une variable aléatoire X à valeurs réelles suit la loi normale centrée réduite notée $\mathcal{N}(0,1)$ lorsque X admet pour densité de probabilité la fonction φ définie sur ℝ par $\boxed{\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}$.

Donc, si u et v sont deux nombres réels vérifiant $u \leqslant v$ alors $P(u \leqslant X \leqslant v) = \int_u^v \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx$



$P(u \leqslant X \leqslant v)$ est l'aire en unité d'aire de la partie du plan colorée en vert.

1.3. Remarques

a) Problème : on ne sait pas écrire une primitive de φ sur ℝ avec des fonctions usuelles.

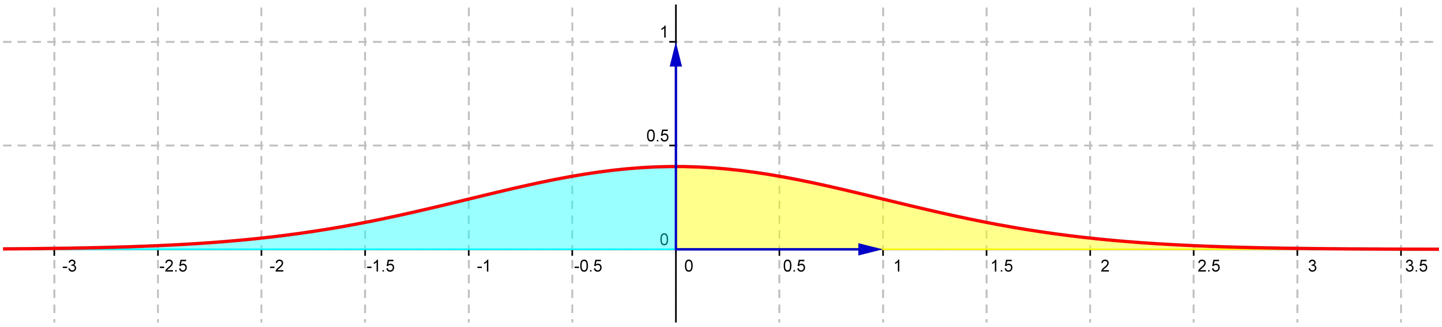
b) La courbe représentative de φ est symétrique par rapport à l'axe des ordonnées donc :

on a l'aire en U.A comprise entre la courbe et l'axe des abscisses sur l'intervalle $]-\infty; 0]$ est égale à l'aire en U.A comprise entre la courbe et l'axe des abscisses sur l'intervalle $[0; +\infty[$.

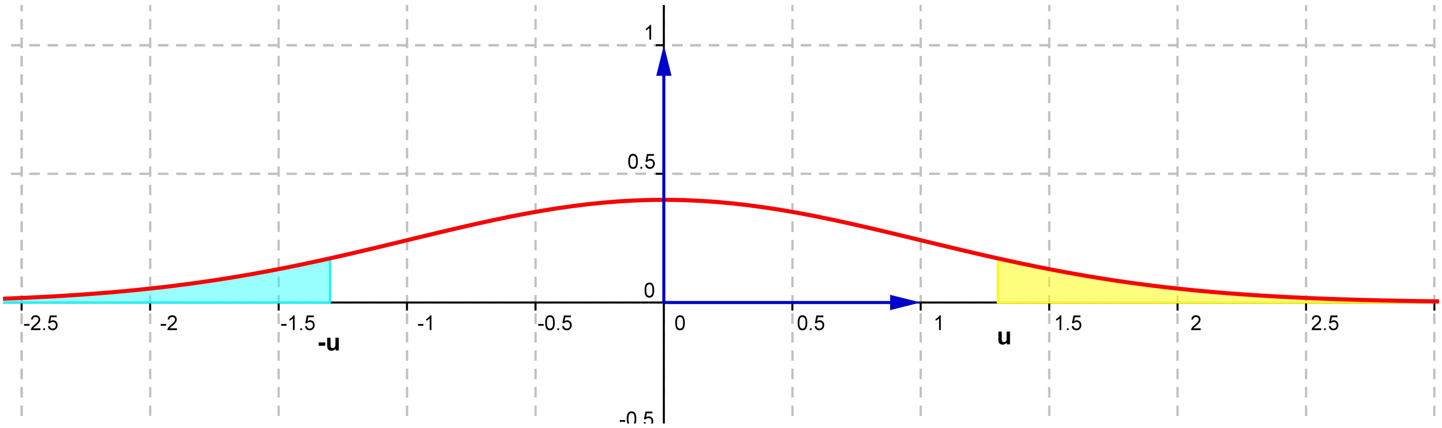
Donc, $P(X \leqslant 0) = P(0 \leqslant X)$

Or, $P(X \leqslant 0) + P(X \geqslant 0) = 1$

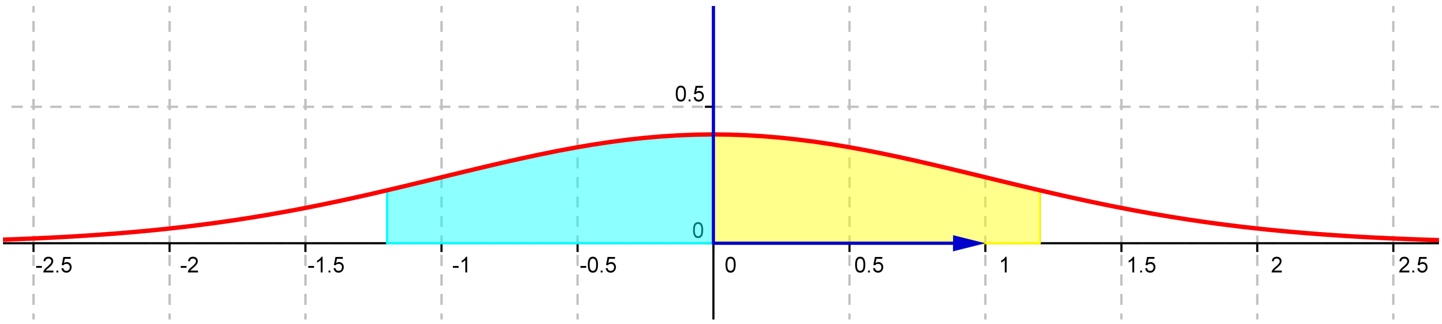
Donc, $\boxed{P(X \leqslant 0) = P(X \geqslant 0) = \frac{1}{2}}$



c) $\boxed{P(X \leqslant -u) = P(u \leqslant X)}$



d) Si $u \geqslant 0$ alors $\boxed{P(-u \leqslant X \leqslant 0) = P(0 \leqslant X \leqslant u)}$



On a donc $P(-u \leqslant X \leqslant u) = 2P(0 \leqslant X \leqslant u)$ et $P(X \leqslant -u) + P(-u < uX \leqslant u) + P(X \leqslant u) = 1$

Donc, $\boxed{P(-u \leqslant X \leqslant u) = 1 - 2P(X \leqslant -u)}$

1.4. Fonction de répartition

a) *Définition*

Si X suit une loi normale, réduite $\mathcal{N}(0;1)$ **sa fonction de répartition** ϕ est définie pour tout x réel par $\boxed{\phi(x) = P(X \leqslant x)}$.

b) *Remarques*

φ est continue sur $\mathbb{R}$ donc φ admet des primitives sur $\mathbb{R}$. On note G la primitive de φ sur $\mathbb{R}$ telle que $G(0) = 0$.

Donc, $G(x)=\int_0^x \varphi(t)\,\mathrm{d}t \qquad (G'(x)=\varphi(x))$

Si $x\geqslant 0$, alors $G(x)$ est **l'aire** en U.A de la partie de plan comprise entre **la courbe** représentative de φ et **l'axe des abscisses** sur $[0;x]$.

Si $x\leqslant 0$, alors $G(x)=-\int_x^0 \varphi(t)\,\mathrm{d}t$ est **l'opposé de l'aire** en U.A de la partie de plan comprise entre **la courbe** représentative de φ et **l'axe des abscisses** sur $[x;0]$.

On a aussi :
$\boxed{\lim_{x\to+\infty} G(x)=\frac{1}{2}}$ (c'est l'aire en U.A de la partie de plan comprise entre la courbe et l'axe des abscisses sur $[0;+\infty[$) et $\boxed{\lim_{x\to-\infty} G(x)=\frac{1}{2}}$.

c) *Égalités*

Pour tout réel x , on a $\boxed{P(X\leqslant x)=\phi(x)=\frac{1}{2}+G(x)}$ avec $G(x)=\int_0^x \varphi(t)\,\mathrm{d}t$

ϕ est dérivable sur ℝ et $\phi'(x)=G'(x)=\varphi(x)$

Conclusion : ϕ est **une primitive** de φ sur ℝ.

d) *Conséquences*

- Si u et v sont deux nombres réels vérifiant $u\leqslant v$ alors :
  $\boxed{P(u\leq X\leq v)=\int_u^v \varphi(x)\,\mathrm{d}x=[\phi(x)]_u^v=\phi(v)-\phi(u)}$

- $\phi(-x)=P(X\leqslant -x)$
  Si $x>0$ alors $P(X\leqslant -x)+P(-x\leqslant X\leqslant x)+P(x\leqslant X)=1$
  Or, $P(x\leqslant X)=P(x\leqslant -x)=\phi(-x)$
  Donc, $\boxed{P(-x\leqslant X\leqslant x)=1-2\phi(x)}$

- $P(X\leqslant x)+P(x\leqslant X)=1$
  Or, $P(x\leqslant X)=P(X\leqslant -x)=\phi(-x)$
  Donc, $\boxed{\phi(x)+\phi(-x)=1}$

- Si $x\geqslant 0$ alors $P(-x\leqslant X\leqslant x)=1-2\phi(x)=1-2(1-\phi(x))=2\phi(x)-1$
  $\boxed{P(-x\leqslant X\leqslant x)=2\phi(x)-1}$

e) *Étude de la fonction*

ϕ est une primitive de φ sur $\mathbb{R}$.
Donc, ϕ est dérivable sur $\mathbb{R}$ et pour tout nombre réel x : $\phi'(x)=\varphi(x)>0$

$\phi(x)=\frac{1}{2}+G(x)$

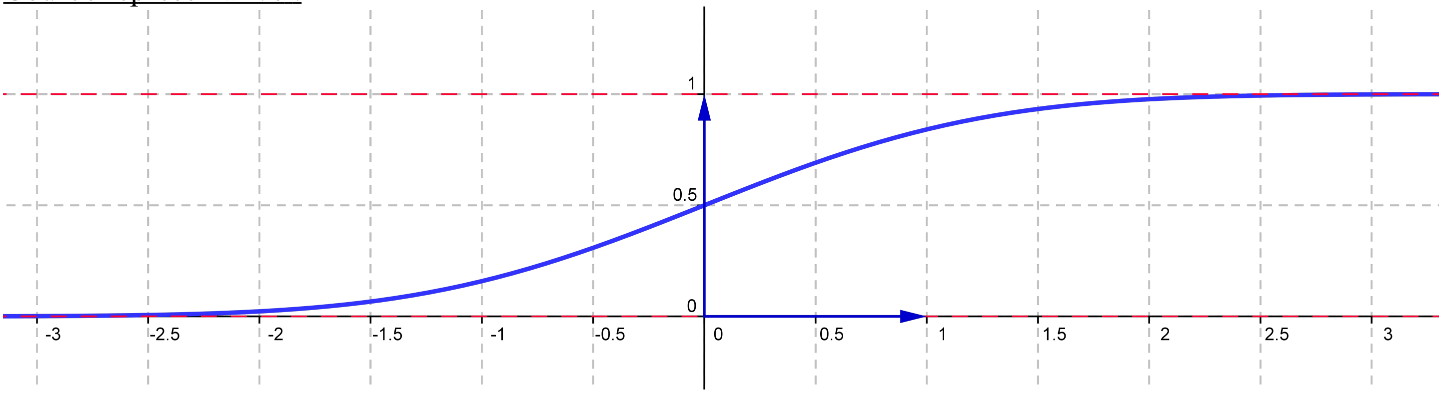
Or, $\lim_{x\to+\infty} G(x)=\frac{1}{2}$ donc $\lim_{x\to+\infty} \phi(x)=1$

et $\lim_{x\to-\infty} G(x)=-\frac{1}{2}$ donc $\lim_{x\to-\infty} \phi(x)=0$

Tableau de variations :

x	$-\infty$	$+\infty$
Signe de $\Phi'(x)$	$+$	
Variations de $\Phi(x)$	0 ↗	1

Courbe représentative :



La droite d'équation $y=1$ est une asymptote à la courbe représentative de ϕ en $+\infty$ et $y=0$ en $-\infty$.

1.5. Calculs numériques

a) On peut facilement donner des valeurs approchées des probabilités du type $P(a\leqslant X\leqslant b)$ en utilisant une calculatrice ou un logiciel. Par exemple, calculons $P(1\leqslant X\leqslant 2{,}5)$

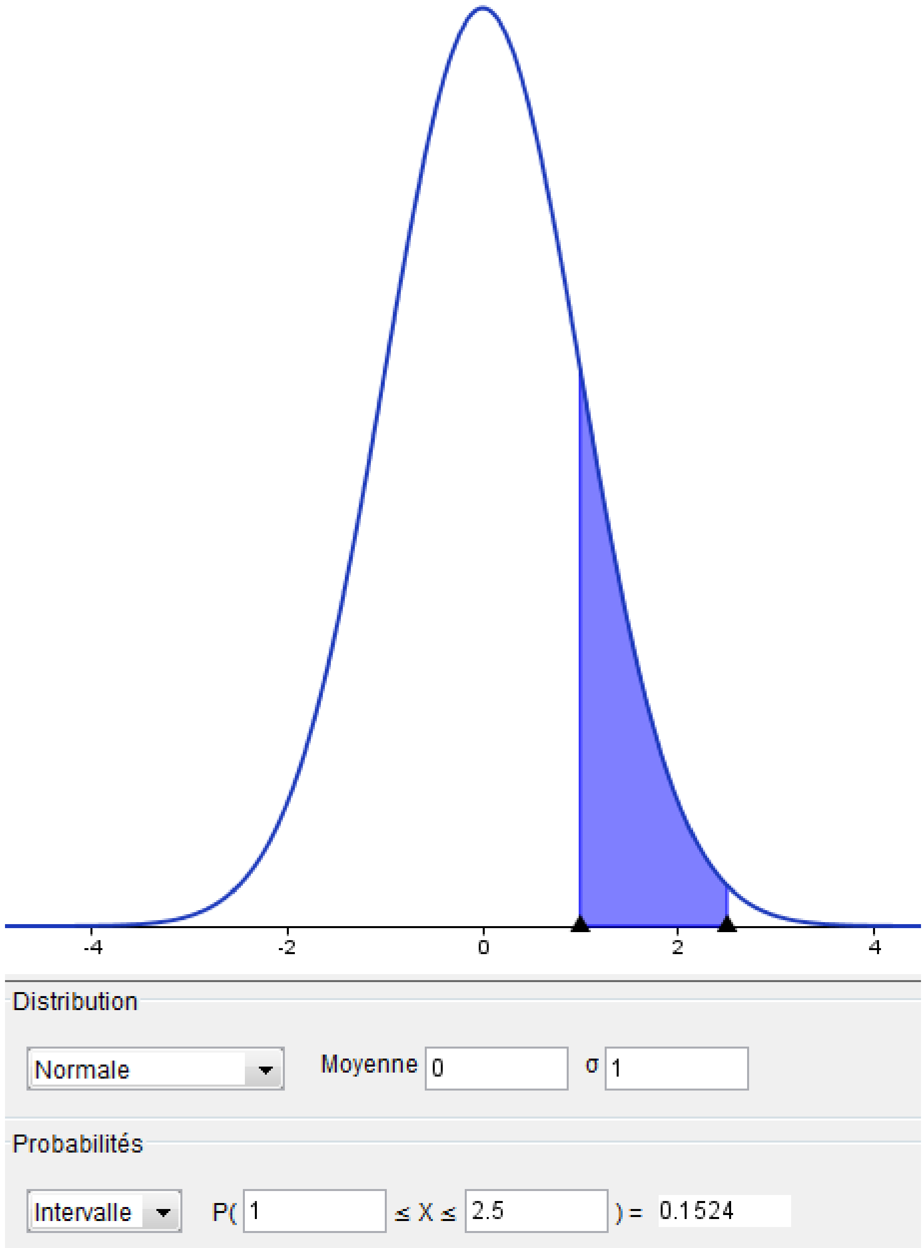
- *Casio Graph 35+*

Par la touche OPTN choisir Stat, Dist, Norm, Normcd **Norm CD(1,2.5,1,0)**
Résultat : 0,1524

- *TI 82, STAT, 83*

Par la touche 2nde Var choisir normal Frép **normal Frep(1,2.5,0,1)**
Résultat : 0,1524

- *géogébra*



b) On peut donner des valeurs approchées de $\phi(a)$ pour a réel.

- Si $a \geqslant 0$, alors $\phi(a) = \frac{1}{2} + P(0 \leqslant X \leqslant a)$

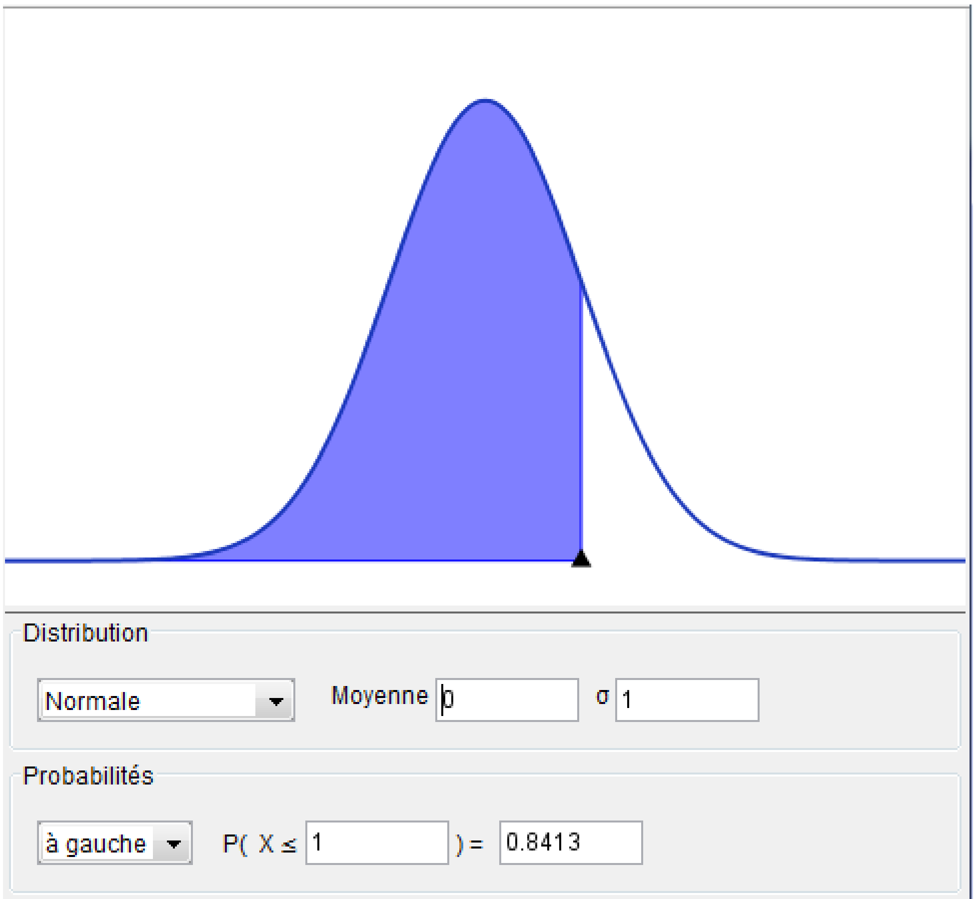
  La calculatrice donne une valeur approchée de $P(0 \leqslant X \leqslant a)$
  Exemple : $a = 0{,}5$
  $\phi(a) = 0{,}5 + P(0 \leqslant X \leqslant 0{,}5)$
  $\phi(a) = 0{,}5 + 0{,}1915$
  $\phi(a) = 0{,}6915$

- Si $a \leqslant 0$, alors $\phi(a) = 0{,}5 - P(a \leqslant X \leqslant 0)$
  Or, $P(a \leqslant X \leqslant 0) = P(0 \leqslant X \leqslant -a)$
  Donc $\phi(a) = 0{,}5 - P(0 \leqslant X \leqslant -a)$

Exemple : $a=-0,2$
$P(0\leqslant X\leqslant 0,2)=0,0793$
$\phi(-0,2)=0,5-0,0793=0,4207$

- Avec géogébra, on a directement $\phi(a)$

Pour $a=0,5$



c) Il est possible que l'on interdise à un examen l'utilisation de la calculatrice ou d'un logiciel (on peut penser à l'examen de fin de première année de médecine)
Dans ce cas, on peut donner au candidat un extrait d'une table donnant pour les nombres décimaux avec au plus deux chiffres après la virgule compris entre 0 et 3 (ou 4) une valeur approchée à 10^{-4} près.

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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

Par exemple :
$\phi(1,22)=0,8888$

$\phi(1,35)=0,9115$

$\phi(-1,37)=1-\phi(1,37)=1-0,9131=0,0869$

$P(1,22\leqslant X\leqslant 1,35)=\phi(1,35)-\phi(-1,22)$
Or, $\phi(-1,22)=1-\phi(1,22)$

$P(1{,}22 \leqslant X \leqslant 1{,}35) = 0{,}8888 + 0{,}9115 - 1 = 0{,}8003$

$P(-1{,}22 \leqslant X \leqslant 1{,}22) = 2\phi(1{,}22) - 1 = 2 \times 0{,}8888 - 1 = 0{,}7776$

1.6. Espérance et variance d'une loi normale centrée réduite

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

Son espérance mathématiques est $\boxed{E(X)=0}$.

Sa variance est $\boxed{V(x)=E((X-E(X))^2)=1}$

Démonstration :

$E(x)=\int_{-\infty}^{+\infty} t\varphi(t)\,\mathrm{d}t$

On considère, pour $x \geqslant 0$, $\int_0^x t\varphi(t)\,\mathrm{d}t$

Or, pour tout réel t , $\varphi'(t)=-t\varphi(t)$, donc une primitive de $t \longmapsto t\varphi(t)$ est $-\varphi(t)$.
Donc,

$\int_0^x t\varphi(t)\,\mathrm{d}t = [-\varphi(t)]_0^x = -\varphi(x)+\varphi(0)$

$\lim_{x\to+\infty} \varphi(x)=0$

Donc, $\lim_{x\to+\infty} \int_0^x t\varphi(t)\,\mathrm{d}t = \varphi(0)$

On considère, pour $y \leqslant 0$, $\int_y^0 t\varphi(t)\,\mathrm{d}t$

$\int_y^0 t\varphi(t)\,\mathrm{d}t = [-\varphi(t)]_y^0 = -\varphi(0)+\varphi(y)$

$\lim_{y\to-\infty} \varphi(y)=0$

Donc, $\lim_{y\to-\infty} \int_y^0 t\varphi(t)\,\mathrm{d}t = -\varphi(0)$

$E(x)=\lim_{y\to-\infty}\int_y^0 t\varphi(t)\,\mathrm{d}t + \lim_{x\to+\infty}\int_0^x t\varphi(t)\,\mathrm{d}t = -\varphi(0)+\varphi(0)=0$

$E(X)=0$

On admet que $V(x)=\sigma^2=1$

Remarque :
Ces résultats justifient la notation $\mathcal{N}(0\,;1)$: 0 pour l'espérance mathématique et 1 pour la variance.

1.7. Théorème

Si X est une variable aléatoire suivant **une loi normale centrée et réduite** $\mathcal{N}(0\,;1)$ alors pour tout nombre réel α appartenant à l'intervalle $]0;1[$, il existe un unique nombre réel strictement positif u_α tel que : $\boxed{P(-u_\alpha \leqslant X \leqslant u_\alpha)=1-\alpha}$

Démonstration :
La démonstration peut faire l'objet d'une restitution organisée des connaissances au baccalauréat.

On considère la fonction H définie sur $[0;+\infty[$ par :

$$H(x)=P(-x\leqslant X\leqslant x)=\int_{-x}^{x}\varphi(t)\,\mathrm{d}t=[\phi(t)]_{-x}^{x}=\phi(x)-\phi(-x)$$

(Rappel : ϕ est une primitive de φ sur $[0;+\infty[$)

H est dérivable sur $[0;+\infty[$
$H'(x)=\phi'(x)-(-\phi'(-x))=\phi'(x)+\phi'(-x)=\varphi(x)+\varphi(-x)=2\varphi(x)>0$
(car φ est une fonction paire et positive)
H est donc strictement croissante sur $[0;+\infty[$.

$\lim\limits_{x\to+\infty}\phi(x)=1$
$\lim\limits_{x\to+\infty}(-x)=-\infty$ et $\lim\limits_{X\to-\infty}\phi(X)=0$ donc $\lim\limits_{x\to-\infty}\phi(-x)=0$
Donc, $\lim\limits_{x\to+\infty}H(x)=1$

$H(0)=0$

H est continue, strictement croissante sur $[0;+\infty[$, $H(0)=0$ et $\lim\limits_{x\to+\infty}H(x)=1$, le théorème des valeurs intermédiaires nous permet d'affirmer que pour tout y tel que $0\leqslant y<1$ il existe x unique tel que $H(x)=y$.

Si $0<\alpha<1$, on pose $y=1-\alpha$ et $y\in]0;1[$. On note u_α l'unique antécédent de $y=1-\alpha$. $u_\alpha\in]0;+\infty[$, et $P(-u_\alpha\leqslant X\leqslant u_\alpha)=1-\alpha$.

1.8. Cas particuliers

- Premier cas : $\alpha=0{,}05$

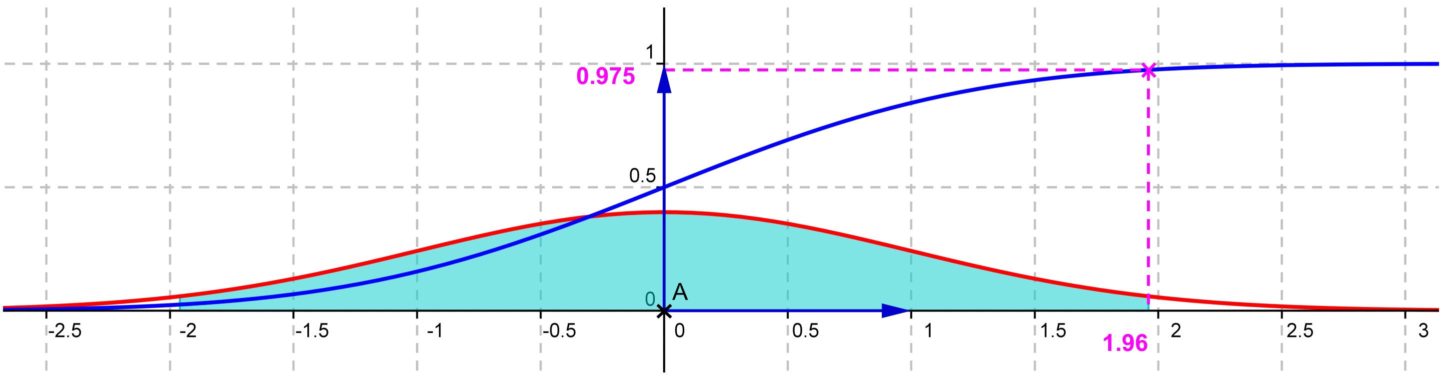
$P(-u_{0,05}\leqslant X\leqslant u_{0,05})=0{,}95$

$$H(x)=\int_{-x}^{x}\varphi(t)\,\mathrm{d}t=\phi(x)-\phi(-x)$$

et, $\phi(-x)=1-\phi(x)$
Donc, $H(x)=2\phi(x)-1$
$H(u_{0,05})=2\phi(u_{0,05})-1$
$2\phi(u_{0,05})-1=0{,}95$

$$\phi(u_{0,05})=\frac{1{,}95}{2}=0{,}975$$

On utilise la calculatrice ou un logiciel ou on regarde une table de valeurs approchées de $\phi(x)$.
On obtient $u_{0,05}=1{,}96$
Ce résultat doit être connu : $\boxed{P(-1{,}96\leqslant X\leqslant 1{,}96)=0{,}95}$



Deuxième cas : $\alpha = 0{,}01$

$P(-u_{0,01} \leqslant X \leqslant u_{0,01}) = 0{,}99$

$H(x) = \int_{-x}^{x} \varphi(t)\,dt = \phi(x) - \phi(-x)$

et, $\phi(-x) = 1 - \phi(x)$

Donc, $H(x) = 2\phi(x) - 1$

$H(u_{0,01}) = 2\phi(u_{0,01}) - 1$

$2\phi(u_{0,01}) - 1 = 0{,}99$

$\phi(u_{0,01}) = \frac{1{,}99}{2} = 0{,}995$

On utilise la calculatrice ou un logiciel ou on regarde une table de valeurs approchées de $\phi(x)$.

On obtient $u_{0,01} = 2{,}58$

Ce résultat doit être connu : $\boxed{P(-2{,}58 \leqslant X \leqslant 2{,}58) = 0{,}99}$

2. Théorème de Moivre-Laplace

2.1. Propriétés de l'espérance mathématique et de la variance d'une variable aléatoire

X est une variable aléatoire.

$E(X)$ est son espérance mathématique.

$V(X)$ est sa variance.

Si b est un nombre réel alors :

$\boxed{E(X+b) = E(X) + b}$ $\boxed{V(X+b) = V(X)}$

Si on ajoute à chaque valeur de X le nombre b alors on ajoute b à la moyenne.

Si on ajoute à chaque valeur de X le nombre b alors les écarts des nouvelles valeurs à la nouvelle espérance mathématique sont égaux aux précédents.

Cas particuliers :

On pose $\mu = E(X)$

$\boxed{E(X - \mu) = 0}$ $\boxed{V(X - \mu) = V(X)}$

Si a est un réel non nul alors :

$\boxed{E(aX)=a\,E(X)}$ $\boxed{V(aX)=a^2\,V(X)}$

Si on multiplie toutes les valeurs de X par a alors on multiplie par a la moyenne.

Si on multiplie toutes les valeurs de X par a alors on multiplie par a^2 les écarts au carré des nouvelles valeurs de X à la nouvelle espérance mathématique.

Cas particuliers :

Si σ^2 est la variance de X, alors

$\boxed{E\left(\frac{X}{\sigma}\right)=\frac{1}{\sigma}E(X)}$ $\boxed{V\left(\frac{X}{\sigma}\right)=\frac{1}{\sigma^2}V(X)=1}$

Conséquence :

Soit X une variable aléatoire réelle, on note $E(X)=\mu$ et $V(X)=\sigma^2$.

La valeur aléatoire $\frac{X-\mu}{\sigma}$ a pour espérance mathématiques 0 et pour variance 1.

$\boxed{E\left(\frac{X-\mu}{\sigma}\right)=0}$ et $\boxed{V\left(\frac{X-\mu}{\sigma}\right)=1}$.

2.2. Lois binomiales

$n\in\mathbb{N}^*$ $p\in]0;1[$

X_n est la loi binomiale de paramètres n et p $\mathcal{B}(n;p)$ si et seulement si X_n prend les valeurs 0 ; 1 ; 2 ; … ; n avec les probabilités : $\boxed{P(X_n=k)=\binom{n}{p}p^k(1-p)^{n-k}}$ (pour tout entier naturel k compris entre 0 et n).

L'espérance mathématique de X_n est : np

$\boxed{\mu=E(X_n)=np}$ $\boxed{\sigma^2=V(X_n)=np(1-p)}$

On considère la variable aléatoire : $\boxed{Z_n=\frac{X_n-np}{\sqrt{np(1-p)}}=\frac{X_n-\mu}{\sigma}}$

Alors, $\boxed{E(Z_n)=0}$ et $\boxed{V(Z_n)=1}$

La loi de probabilité de Z_n n'est pas une loi binomiale car les valeurs de Z_n ne sont pas des entiers (compris entre 0 et n).

2.3. Théorème de Moivre-Laplace (admis)

Si la variable aléatoire X_n suit **la loi binomiale** $\mathcal{B}(n;p)$ alors la variable Z_n définie par $Z_n=\dfrac{X_n-np}{\sqrt{np(1-p)}}$ **converge vers la loi normale centrée réduite** $\mathcal{N}(0;1)$, c'est à dire que pour les réels a et b tels que $a\leqslant b$:

$$\boxed{\lim_{n\to+\infty} P(a\leqslant Z_n\leqslant b)=\int_a^b \varphi(t)\,\mathrm{d}t=\int_a^b \frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x}$$

3. Loi normale $\mathcal{N}(\mu\,;\sigma^2)$

3.1. Définition

μ est un nombre réel donné.
σ est un nombre réel strictement positif donné.

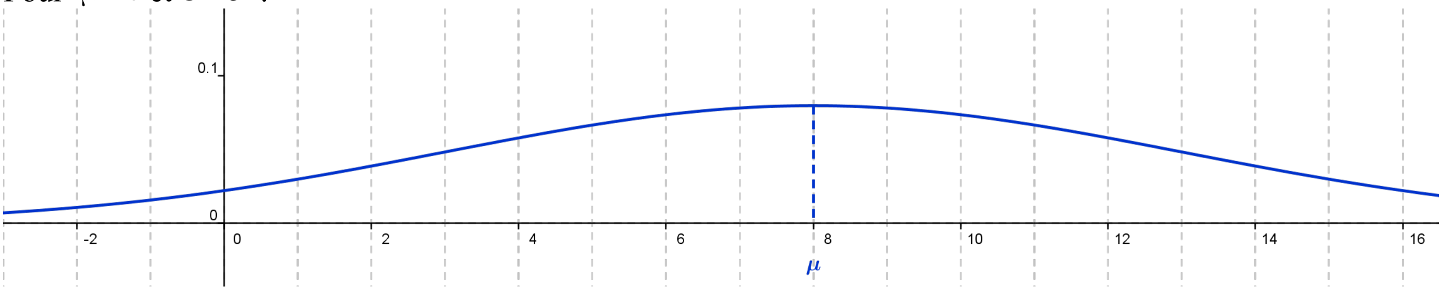
On dit que la variable aléatoire X suit **la loi normale $\mathcal{N}(\mu;\sigma^2)$** si la variable $\dfrac{X-\mu}{\sigma}$ suit **la loi normale centrée et réduite $\mathcal{N}(0;1)$**.

3.2. Conséquences

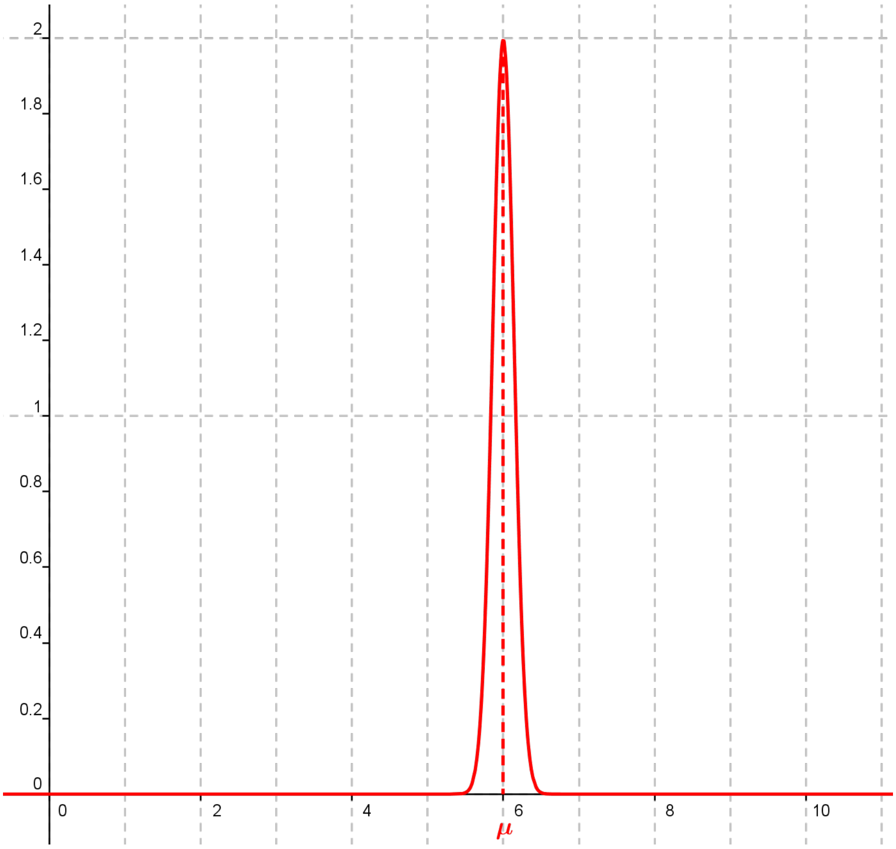
- $E\left(\dfrac{X-\mu}{\sigma}\right)=0$ donc $\boxed{E(X)=\mu}$
- $V\left(\dfrac{X-\mu}{\sigma}\right)=1$ Donc $\dfrac{1}{\sigma^2}V(X-\mu)=1$ donc $\dfrac{1}{\sigma^2}V(x)=1$ donc $\boxed{V(x)=\sigma^2}$
- La fonction de densité est définie sur $\mathbb{R}$ par $f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

La droite d'équations $x=\mu$ est un axe de symétrie de la courbe représentative de f.

Exemples :
Pour $\mu=8$ et $\sigma=5$:



sPour $\mu=6$ et $\sigma=0{,}2$:



- F est la fonction de répartition. Pour tout nombre réel x, $\boxed{F(x)=P(X\leqslant x)=\phi\left(\frac{x-\mu}{\sigma}\right)}$.

3.3. Probabilités des événements : $X\in[\mu-\sigma;\mu+\sigma]$; $X\in[\mu-2\sigma;\mu+2\sigma]$; $X\in[\mu-3\sigma;\mu+3\sigma]$

- $P(X\in[\mu-\sigma;\mu+\sigma])=P(\mu-\sigma\leq X\leq\mu+\sigma)=P(-1\leq\frac{X-\mu}{\sigma}\leq 1)=2\Phi(1)-1=2\times 0{,}8413-1=0{,}6826$

On retiendra : $\boxed{P(X\in[\mu-\sigma;\mu+\sigma])=0{,}68}$

- $P(X\in[\mu-2\sigma;\mu+2\sigma])=P(\mu-2\sigma\leq X\leq\mu+2\sigma)=P(-2\leq\frac{X-\mu}{\sigma}\leq 2)=2\Phi(2)-1=2\times 0{,}9772-1=0{,}9544$

On retiendra : $\boxed{P(X\in[\mu-2\sigma;\mu+2\sigma])=0{,}95}$

- $P(X\in[\mu-3\sigma;\mu+3\sigma])=P(\mu-3\sigma\leq X\leq\mu+3\sigma)=P(-3\leq\frac{X-\mu}{\sigma}\leq 3)=2\Phi(3)-1=2\times 0{,}9987-1=0{,}9974$

On retiendra : $\boxed{P(X\in[\mu-3\sigma;\mu+3\sigma])=0{,}99}$

4. Intervalle de fluctuation asymptotique-Estimation

4.1. Rappels

a) *Échantillonnage*

Dans une population où la fréquence d'un caractère est p ($p\in]0;1[$), la variable aléatoire X_n (n entier naturel supérieur ou égal à 2) qui a tout échantillon de taille n associe le nombre d'individus qui possèdent le caractère étudié suit la loi binomiale de paramètres n et p : $\mathcal{B}(n;p)$.

La variable aléatoire fréquence est : $F_n=\frac{X_n}{n}$ n'est pas une loi binomiale car F_n ne prend pas que des valeurs entières.

b) *Intervalle de fluctuation (cours de 1ère S)*

Un intervalle de fluctuation au seuil de 95% est un intervalle, de la forme $I_n=\left[\frac{a}{n};\frac{b}{n}\right]$ où a et b sont des entiers compris entre 0 et n , tel que : $P\left(\frac{a}{n}\leqslant F_n\leqslant\frac{b}{n}\right)=P(F_n\in I_n)\geqslant 0{,}95$.

Pour obtenir un intervalle $I_n=\left[\frac{a}{n};\frac{b}{n}\right]$ de plus faible amplitude, on détermine les plus petits entiers naturels a et b tels que $P(X_n=a)\geqslant 0{,}025$ et $P(X_n\leqslant b)<0{,}975$.

c) *Intervalle de fluctuation simplifié (cours de 2de)*

Lorsque $n\geqslant 30$, $np\geqslant 5$ et $np(1-p)\geqslant 5$, l'intervalle $I_n=\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$ est un intervalle « simplifié » de $F_n=\frac{X_n}{n}$ au seuil de 95%.

$P(F_n\in I_n)\geqslant 0{,}95$

4.2. Intervalle de fluctuation asymptotique

a) *Théorème*

On suppose $n\geqslant 30$, $np\geqslant 5$ et $np(1-p)\geqslant 5$, $\alpha\in]0;1[$

Soit u_α l'unique réel strictement positif tel que $\int_{-u_\alpha}^{u_\alpha}\varphi(x)\,\mathrm{d}x=1-\alpha$.

Si X_n suit la loi binomiale $\mathcal{B}(n;p)$

si $F_n=\frac{X_n}{n}$

si $I_n=\left[p-u_\alpha\frac{p(1-p)}{\sqrt{n}};p+u_\alpha\frac{p(1-p)}{\sqrt{n}}\right]$

Alors, $\boxed{\lim_{n\to+\infty}(F_n\in I_n)=1-\alpha}$

Démonstration :

Cette démonstration peut faire l'objet d'une restitution organisée des connaissances au baccalauréat.

$Z_n=\frac{X_n-np}{\sqrt{np(1-p)}}$

Le théorème de Moivre-Laplace permet d'affirmer, que pour tous les nombres réels a et b , que :

$$\lim_{n\to+\infty}P(a\leqslant Z_n\leqslant b)=\int_a^b\varphi(x)\,\mathrm{d}x$$

Et, $\lim_{n\to+\infty}P(-u_\alpha\leqslant Z_n\leqslant u_\alpha)=\int_{-u_\alpha}^{u_\alpha}\varphi(x)\,\mathrm{d}x=1-\alpha$

Or, $Z_n = \dfrac{\dfrac{X_n}{n} - p}{\dfrac{\sqrt{p(1-p)}}{\sqrt{n}}} = \dfrac{F_n - p}{\dfrac{\sqrt{p(1-p)}}{\sqrt{n}}}$

Donc, $\lim\limits_{n \to +\infty} P\left(p - u_\alpha \dfrac{\sqrt{p(1-p)}}{\sqrt{n}} \leqslant F_n \leqslant p + u_\alpha \dfrac{\sqrt{p(1-p)}}{\sqrt{n}}\right) = 1 - \alpha$

Soit, $\lim\limits_{n \to +\infty} (F_n \in I_n) = 1 - \alpha$

b) *Définition*

Pour tout $\alpha \in]0;1[$, $\boxed{I_n \in \left[p - u_\alpha \dfrac{\sqrt{p(1-p)}}{\sqrt{n}} \,;\, p + u_\alpha \dfrac{\sqrt{p(1-p)}}{\sqrt{n}}\right]}$ se nomme **intervalle de fluctuation asymptotique** de la variable aléatoire $F_n = \dfrac{X_n}{n}$ au seuil de $1 - \alpha$.
I_n contient F_n avec une probabilité d'autant plus proche de $1 - \alpha$ que n est grand.

c) *Cas particulier*

$\alpha = 5\%$ $\quad\quad 1 - \alpha = 0{,}95$ (95%)

$u_{0,05} = 1{,}96$ (valeur approchée à 1/100 près par défaut)

$\boxed{I_n = \left[p - 1{,}96 \dfrac{\sqrt{p(1-p)}}{\sqrt{n}} \,;\, p + 1{,}96 \dfrac{\sqrt{p(1-p)}}{\sqrt{n}}\right]}$ est un **intervalle de fluctuation asymptotique** de la variable aléatoire $F_n = \dfrac{X_n}{n}$ au seuil de 95%.

d) *Remarque*

On considère la fonction définie sur]0;1[par $h(x) = x(1-x) = -x^2 + x$
h est dérivable sur]0;1[.

x	0		$\frac{1}{2}$		1
Signe de h ' (x)	||	+	0	−	
Variations de h(x)	||	↗	$\frac{1}{4}$	↘	

Le maximum de h sur]0;1[est $\dfrac{1}{4}$.

On peut donc majorer $\sqrt{p(1-p)}$ par $\sqrt{\frac{1}{4}}=\frac{1}{2}$ (la fonction racine carrée est croissante sur $[0;+\infty[$)

$$I_n=\left[p-1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}};p+1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]\subset\left[p-\frac{1{,}96}{2}\frac{1}{\sqrt{n}};p+\frac{1{,}96}{2}\frac{1}{\sqrt{n}}\right]\subset\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$$

On retrouve l'intervalle de fluctuation considéré en seconde, et $\lim\limits_{n\to+\infty}\left(p-\frac{1}{\sqrt{n}}\leqslant F_n\leqslant p+\frac{1}{\sqrt{n}}\right)\geqslant 0{,}95$

4.3. Prise de décision au seuil de 5%

La proportion du caractère C étudié dans une population **est supposée être égale à p** $(p\in]0;1[)$.
la prise de décision consiste, à partir d'un échantillon de taille n , à valider ou non cette hypothèse.

On vérifie que : $n\geqslant 30$; $np\geqslant 5$ et $n(1-p)\geqslant 5$ et on détermine $I_n=\left[p-1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}};p+1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]$.

On calcule la fréquence f sur l'échantillon considéré.
On applique la règle de décision au seuil de 5%.
Si $\boxed{f\notin I_n}$ alors **on rejette l'hypothèse** (la proportion est égale à p).
Si $\boxed{f\in I_n}$ alors **on ne rejette pas l'hypothèse** (on accepte l'hypothèse).

Remarque :
Le théorème précédent permet d'affirmer que la probabilité de rejeter l'hypothèse faite sur p alors qu'elle est vraie est d'environ 0,05.

4.4. Intervalle de confiance

On ne connaît pas la proportion p du caractère C dans la population. On connaît la fréquence f du caractère sur un échantillon de taille n (assez grand : $n\geqslant 30$, $nf\geqslant 5$, $n(1-f)\geqslant 5$)
Peut-on estimer la proportion p à partir de f ?

a) En terminale, on utilise l'intervalle de fluctuation de la fréquence F_n : $J_n=\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$.

$\lim\limits_{n\to+\infty} P(F_n\in J_n)\geqslant 0{,}95$

C'est à dire pour n assez grand $P(F_n\in J_n)\geqslant 0{,}95$
Or, f est une fréquence obtenue sur un échantillon de taille n , donc :

$$P\left(p-\frac{1}{\sqrt{n}}\leqslant f\leqslant p+\frac{1}{\sqrt{n}}\right)\geqslant 0{,}95$$

Or, $p-\frac{1}{\sqrt{n}}\leqslant f\leqslant p+\frac{1}{\sqrt{n}}$

$\Leftrightarrow f-\frac{1}{\sqrt{n}}\leqslant p\leqslant f+\frac{1}{\sqrt{n}}$

et $P\left(f-\frac{1}{\sqrt{n}}\leqslant p\leqslant f+\frac{1}{\sqrt{n}}\right)\geqslant 0{,}95$

b) *Définition*

Soit f une fréquence du caractère C sur un échantillon de taille n .
L'intervalle $\left[f-\frac{1}{\sqrt{n}};f+\frac{1}{\sqrt{n}}\right]$ est **un intervalle de confiance au seuil de 95%** de la proportion inconnue p dans la population.
L'amplitude de l'intervalle est $\frac{2}{\sqrt{n}}$.