Chapitre 12

Lois à densité

I. Variables aléatoires à densité

1) Introduction

Dans les situations étudiées jusqu'à présent, à une expérience aléatoire, on associait un univers fini Ω={ $w_1, w_2, ..., w_n$ } à une loi de probabilité P.
Toute variable aléatoire X ne prenait alors qu'un nombre fini de valeurs x_1 , x_2 , …, x_k .
Cependant, certaines expériences aléatoires conduisent à des variables aléatoires qui prennent toutes les valeurs d'un intervalle I de ℝ.

Exemples :

- On tire sur une cible circulaire de 1 m de rayon, sans jamais la manquer. La variable aléatoire qui indique la distance, en mètre, du point d'impact au centre prend toutes les valeurs de $[0;1]$.
- Lors d'un appel téléphonique, la durée de communication est une variable aléatoire qui prend ses valeurs dans un intervalle de temps.

Dans ces conditions, il n'est plus possible de définir la loi de X en dressant le tableau des probabilités de chacun des événements { $X=x_i$ } puisqu'il y en a une infinité.
Une autre approche est alors nécessaire.
On s'intéresse aux événements du type « X prend ses valeurs dans un intervalle J » noté { X∈J }.
Il s'agit alors de définir la probabilité $P(X \in J)$.

Remarque :
La notation { X∈J } est abusive car X n'est pas un nombre. Il s'agit de l'ensemble des issues de Ω auxquelles on associe un réel appartenant à J : { $w \in \Omega \mid X(w) \in J$ }.

2) Variable aléatoire continue

Définitions :

- L'ensemble des issues d'une expérience aléatoire s'appelle l'**univers** de l'expérience.
  On le note Ω.
- Une **variable aléatoire** est une fonction définie sur Ω et à valeurs dans ℝ.
  On la note X.

Exemple :
On lance deux dés cubiques équilibrés dont les faces sont numérotées de 1 à 6. On peut définir une variable aléatoire qui, à chaque lancer, associe la somme des nombres obtenus. Dans ce cas, l'univers Ω est l'ensemble des couples $(i;j)$ où *i* et *j* prennent des valeurs entières comprises entre 1 et 6. Comme cette variable aléatoire ne prend qu'un nombre fini de valeurs (entiers compris entre 2 et 12), elle est dite **discrète**.

Définition :
Soit Ω l'univers d'une expérience aléatoire.
Une variable aléatoire X définie sur Ω, qui prend comme valeur tous les nombres réels d'un intervalle I de ℝ, est dite **continue**.

Exemple :
Une entreprise produit des ampoules à basse consommation.
On peut définir une variable aléatoire X, qui à chaque ampoule produite par cette entreprise, associe sa durée de vie, en heures. Cette durée n'est pas nécessairement un nombre entier d'heures et théoriquement, on ne connaît pas la durée maximale de vie d'une ampoule.
Cette variable aléatoire X est donc continue et l'intervalle I est l'intervalle $[0;+\infty[$.

3) *Fonction densité*

Une fois la variable aléatoire définie, on s'intéresse à sa **loi de probabilité**.

Variable aléatoire discrète

Dans le cas d'une variable aléatoire discrète, la loi de probabilité associée est généralement donnée sous la forme d'un tableau. Elle peut aussi être donnée par l'intermédiaire d'une formule générale ; c'est le cas pour les variables aléatoires qui suivent une loi binomiale.

Exemple :
Dans le 1er exemple, la loi de probabilité de la variable aléatoire discrète est définie par le tableau suivant :

s	2	3	4	5	6	7	8	9	10	11	12
$P(X=s)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

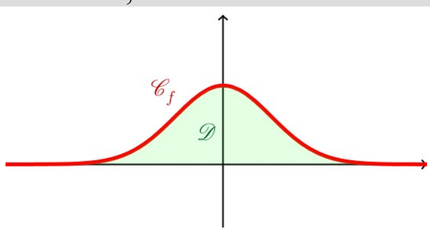
On vérifie que la somme des probabilités des événements $\{X=s\}$ (avec $2\leqslant s\leqslant 12$) est bien égale à 1.

Variable aléatoire continue

Dans le cas d'une variable aléatoire continue, on a recours à une fonction définie sur ℝ appelée **densité de probabilité**.

Définition :
Une fonction f définie sur ℝ est appelée **densité**, ou densité de probabilité, si :

- f est positive sur ℝ : pour tout nombre réel x, $f(x)\geqslant 0$.
- f est continue sur ℝ (sauf éventuellement en un nombre fini de points).
- L'aire du domaine $\mathscr{D}$, domaine délimité par la courbe $\mathscr{C}_f$ représentative de la fonction f dans un repère orthogonal, et par l'axe des abscisses, est égale à 1.



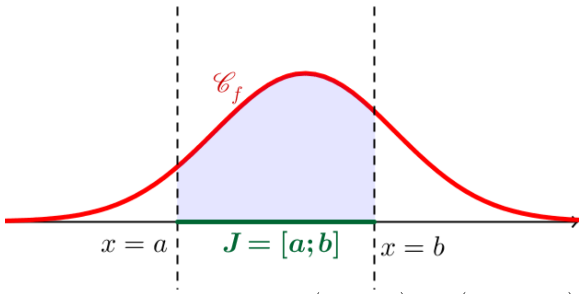
Remarque :
Dans le cas d'une variable aléatoire discrète, la somme des probabilités des événements $\{X=k\}$ est égale à 1. Dans le cas d'une variable aléatoire continue, cette propriété se traduit par l'aire du domaine $\mathscr{D}$ est égale à 1.

4) Probabilité d'un événement

Définition :
Soit Ω l'univers d'une expérience aléatoire et X une variable aléatoire définie sur Ω, continue de densité f .

La probabilité de l'événement $\{X\in J\}$, où J est un **intervalle** de ℝ, est notée $P(X\in J)$ et elle est définie comme l'aire du domaine suivant :

$$\{M(x;y)\ ;\ x\in J \text{ et } 0\leqslant y\leqslant f(x)\}$$

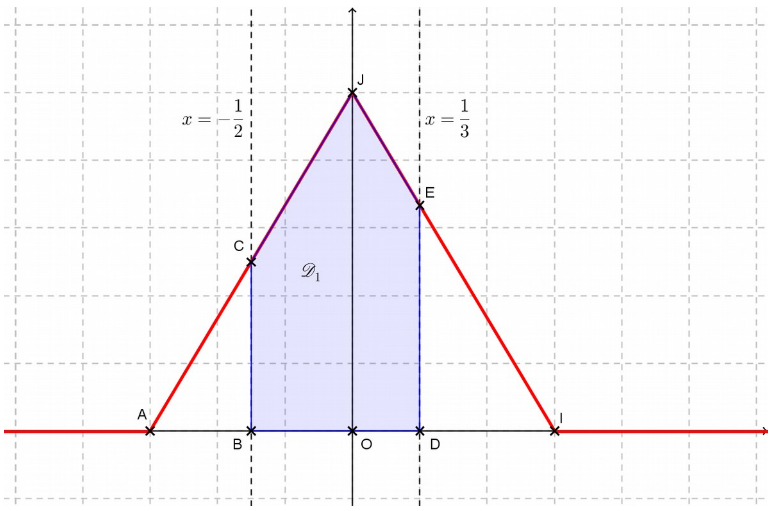


Aire du domaine bleu : $P(X\in J)=P(a\leqslant x\leqslant b)$

Remarques :
- Par cette définition, la probabilité de l'événement $\{X\in\mathbb{R}\}$ correspond à l'aire du domaine $\mathscr{D}$, aire du domaine délimité par la courbe $\mathscr{C}_f$ et l'axe des abscisses. Par la troisième caractéristique de la densité f , cette aire est égale à 1. On a alors $P(X\in]-\infty;+\infty[)=1$.
- La fonction densité d'une variable aléatoire continue définit à elle seule la loi de probabilité de la variable aléatoire continue à laquelle elle est associée.

Exemple :
Soit X une variable aléatoire continue dont la densité f est définie sur ℝ par :

$$f(x)=\begin{cases}1+x \text{ si } -1\leqslant x<0\\ 1-x \text{ si } 0\leqslant x\leqslant 1\\ 0 \text{ sinon}\end{cases}$$



On désire calculer la probabilité de l'événement $\left\{X\in\left[-\frac{1}{2};\frac{1}{3}\right]\right\}$, ce qui se note $P\left(X\in\left[-\frac{1}{2};\frac{1}{3}\right]\right)$ ou encore $P\left(-\frac{1}{2}\leqslant X\leqslant\frac{1}{3}\right)$.

Par définition, cette probabilité correspond à l'aire du domaine $\mathscr{D}_1$ colorié en bleu où $\mathscr{C}_f$ est la courbe représentative de la densité f dans le repère orthogonal $(O;I,J)$.

En notant A, B, C, D et E les points de coordonnées respectives $(-1;0)$, $\left(-\frac{1}{2};0\right)$, $\left(-\frac{1}{2};\frac{1}{2}\right)$, $\left(\frac{1}{3};0\right)$ et $\left(\frac{1}{3};\frac{2}{3}\right)$, l'aire de ce domaine $\mathscr{D}_1$ est égale à l'aire du triangle AIJ à laquelle il faut soustraire l'aire du triangle ABC et celle du triangle DEI.
Par suite, on a :

$$\mathscr{A}_{\mathscr{D}1}=\mathscr{A}_{\text{AIJ}}-\mathscr{A}_{\text{ABC}}-\mathscr{A}_{\text{DEI}}=1-\frac{1}{8}-\frac{2}{9}=\frac{47}{72}$$

Remarques :

- L'aire du domaine $\mathscr{D}_1$ est déterminée au moyen d'arguments géométriques. Mais on peut également déterminer cette aire à l'aide d'un calcul d'intégrale. En effet, comme la densité est une fonction positive et **continue** sur l'intervalle $\left[-\frac{1}{2};\frac{1}{3}\right]$, on a la relation suivante :

$$\mathscr{A}_{\mathscr{D}1}=\int_{-\frac{1}{2}}^{\frac{1}{3}} f(x)\,\mathrm{d}x$$

  Par la relation de Chasles, il en découle :

$$\mathscr{A}_{\mathscr{D}1}=\int_{-\frac{1}{2}}^{0}(1+x)\,\mathrm{d}x+\int_{0}^{\frac{1}{3}}(1-x)\,\mathrm{d}x=\left(\frac{1}{2}-\frac{1}{8}\right)+\left(\frac{1}{3}-\frac{1}{18}\right)=\frac{47}{72}.$$

  D'une manière générale, pour tous nombres réels a et b tels que $a\leqslant b$:

$$P(a\leqslant X\leqslant b)=\int_a^b f(x)\,\mathrm{d}x$$

- On constate graphiquement que la fonction f est bien positive et continue sur $\mathbb{R}$. En outre, l'aire du domaine $\mathscr{D}$, délimité par la courbe $\mathscr{C}_f$ et par l'axe des abscisses, correspond ici à l'aire du triangle AIJ qui est égale à 1.
  On a ainsi bien vérifiée que la fonction f est une densité.
- Par définition, la probabilité de l'événement $\{X\in[-1;1]\}$ est égale à l'aire du domaine délimité par la courbe $\mathscr{C}_f$, l'axe des abscisses et les droites d'équation $x=-1$ et $x=1$. Ce domaine est le triangle AIJ dont l'aire est égale à 1.
  Ainsi, il est certain que X prend ses valeurs dans cet intervalle.
  On peut d'ailleurs démontrer que l'intervalle $[-1;1]$ est le plus petit intervalle tel que la probabilité X y prenne ses valeurs est égale à 1. En dehors de cet intervalle, on constate que la densité est nulle.
  D'une manière générale, une variable aléatoire continue prend ses valeurs sur l'intervalle I où la densité ne s'annule pas. Cet intervalle I est appelé le **support de densité**.

Propriété :
Soit k un nombre réel et X une variable aléatoire continue de densité f .
La probabilité que la variable aléatoire X prenne la valeur k est nulle : $P(X=k)=0$.

Démonstration :
L'événement $\{X=k\}$ s'écrit aussi $\{X \in \{k\}\}$ ou encore $\{X \in [k;k]\}$. Par suite, $P(X=k)=P(X \in [k;k])$.
Par définition, la probabilité de cet événement est alors l'aire du domaine délimité par l'axe des abscisses, la courbe $\mathscr{C}_f$ et les droites d'équation $x=k$. Par conséquent :

$$P(X=k)=\int_k^k f(x)\,\mathrm{d}x=0$$

Conséquence :
Dans le cadre du calcul de probabilité d'un événement $\{X \in J\}$ où X est une **variable aléatoire continue** et J un intervalle de $\mathbb{R}$, les éventuelles inégalités larges peuvent être remplacées par des inégalités strictes et vice-versa.

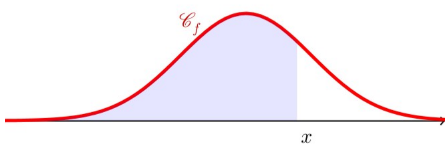
Exemple :
Dans l'exemple précédent. On désire calculer $P\left(X \in \left]-\frac{1}{2};\frac{1}{3}\right]\right)$ ou encore $P\left(-1<X<-\frac{1}{2}\right)$.

- $P\left(X \in \left]-\frac{1}{2};\frac{1}{3}\right]\right) = P\left(-\frac{1}{2}<X\leqslant\frac{1}{3}\right) = P\left(-\frac{1}{2}\leqslant X\leqslant\frac{1}{3}\right) = \mathscr{A}_{\mathscr{D}1} = \frac{47}{72}$
- $P\left(-1<X<-\frac{1}{2}\right) = P\left(-1\leqslant X\leqslant-\frac{1}{2}\right) = \mathscr{A}_{\text{ABC}} = \frac{1}{8}$

5) *Fonction de répartition*

Définition :
Soit X une variable aléatoire continue de densité f dont la courbe représentative est notée $\mathscr{C}_f$.
La **fonction de répartition** associée à la variable aléatoire X est la fonction F définit sur $\mathbb{R}$ par :

$$F(x)=P(X\leqslant x)=P(X \in]-\infty;x])$$



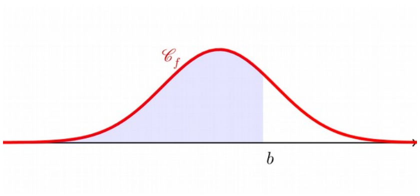
Aire du domaine bleu :

$$P(X\leqslant x)=F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

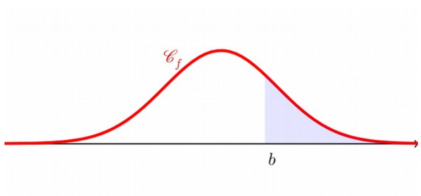
Remarques :
- Comme $P(X=x)=0$ pour tout nombre réel x, on peut remplacer dans la définition précédente l'inégalité large par une inégalité stricte. Ainsi, pour tout réel x, $F(x)=P(X\leqslant x)=P(X<x)$.
- L'image d'un nombre réel x par la fonction de répartition F, $F(x)$, est naturellement un nombre positif (aire d'un domaine). En outre, comme l'aire du domaine délimité par la courbe $\mathscr{C}_f$ et l'axe des abscisses est égale à 1, ce nombre est inférieur ou égal à 1. Ainsi, pour tout nombre réel x, $0\leqslant F(x)\leqslant 1$.
- La fonction de répartition est une fonction continue et croissante sur $\mathbb{R}$.

- La probabilité de l'événement $\{X \in J\}$ où J est un intervalle de ℝ peut toujours s'exprimer à l'aide de la fonction de répartition. En effet :

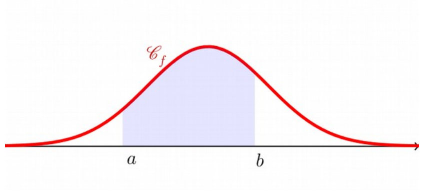
Dans le cas où J est de la forme $]-\infty;b]$ ou $]-\infty;b[$, b étant un nombre réel, la probabilité correspondante est :
$P(X\leqslant b)=P(X<b)=F(b)$



Dans le cas où J est de la forme $[b;+\infty[$, la probabilité correspondante est :
$P(X\geqslant b)=P(X>b)=1-F(b)$



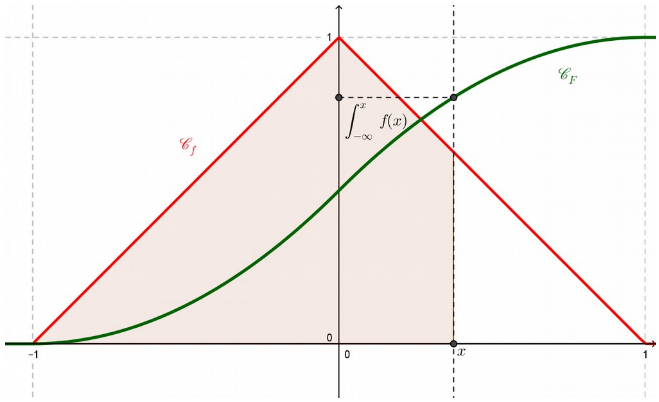
Dans le cas où J est de la forme $[a;b]$, la probabilité correspondante est :
$P(a\leqslant X\leqslant b)=P(a<X\leqslant b)$
$=P(a\leqslant X<b)=P(a<X<b)$
$=F(b)-F(a)$



Ainsi, la loi de probabilité d'une variable aléatoire continue est entièrement déterminée **soit** par la connaissance de la fonction de densité f , **soit** par la connaissance de la fonction de répartition *F*.
Cela s'explique par le lien entre ces deux fonctions. En effet, sur tout intervalle *K* où la fonction de répartition est **dérivable**, on a, pour tout nombre réel $x\in K$, $F'(x)=f(x)$.

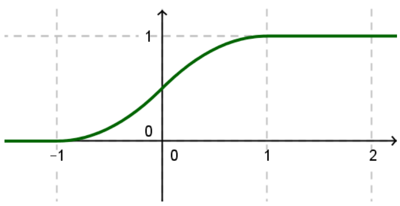
Exemple :
Dans le cas de l'exemple précédent, on peut démontrer que la fonction de répartition *F* est définie sur ℝ par :

$$F(x)=\begin{cases} 0 \text{ si } x<-1 \\ \dfrac{(1+x)^2}{2} \text{ si } -1\leqslant x<0 \\ 1-\dfrac{(1-x)^2}{2} \text{ si } 0\leqslant x\leqslant 1 \\ 1 \text{ si } x>1 \end{cases}$$



Sa courbe représentative $\mathscr{C}_F$ dans un repère orthogonal $(O;I,J)$ est donnée par la figure ci-dessus.

On constate que la fonction de répartition *F* est bien continue et croissante sur ℝ et que, pour tout nombre réel *x*, $0\leqslant F(x)\leqslant 1$.



Remarque :
La fonction de répartition permet de :
- à partir de calculs d'images, calculer une probabilité :
$$P(X\leqslant a)=F(a)$$
- à partir de calculs d'antécédents, déterminer un intervalle :
Pour une probabilité *p* donnée, on cherche *a* tel que $P(X\leqslant a)=F(a)=p$

II. Loi uniforme

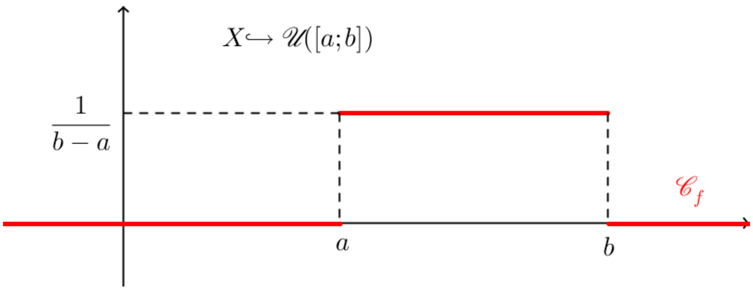
Modèle : tirage d'un nombre au hasard dans $[a;b]$.
Exemples : instant d'arrivée au hasard dans un intervalle de temps $[a;b]$ et, plus généralement, une mesure comprise au hasard entre deux valeurs *a* et *b*.

1) Définition

Définition :
Soit *a* et *b* deux nombres réels distincts tels que $a<b$.
Dire qu'une variable aléatoire continue X suit une loi uniforme sur l'intervalle $[a;b]$ signifie que sa densité f est définie sur ℝ par :

$$f(x)=\begin{cases}\dfrac{1}{b-a} \text{ si } a\leqslant x\leqslant b\\ 0 \text{ sinon}\end{cases}$$

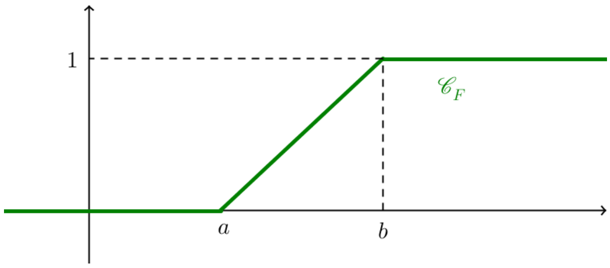
On note :
X suit la loi $\mathscr{U}([a;b])$.



Remarque :
Comme le support de la densité f est l'intervalle $I=[a;b]$, une telle variable aléatoire X prend alors ses valeurs dans cet intervalle.

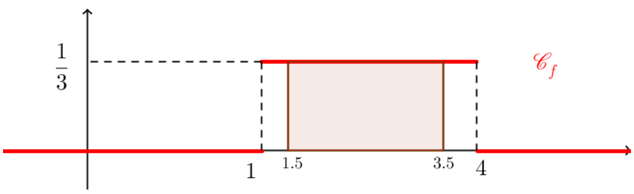
Propriété :
Soit *a* et *b* deux nombres réels distincts tels que $a<b$.
Si une variable aléatoire X suit une loi uniforme sur l'intervalle $[a;b]$, alors la fonction de répartition *F* associée à cette variable est définie pour tout nombre réel *x* par :

$$F(x)=\begin{cases}0 \text{ si } x<a\\ \dfrac{x-a}{b-a} \text{ si } a\leqslant x\leqslant b\\ 1 \text{ si } x>b\end{cases}$$



Remarque :
Sur la figure, on constate que la fonction de répartition *F* est bien continue, croissante sur ℝ et que pour tout nombre réel *x*, $0\leqslant F(x)\leqslant 1$.

Exemple :
On considère X une variable aléatoire continue suivant la loi uniforme $\mathscr{U}([1;4])$.
On désire calculer la probabilité de l'événement $\{X\in[1{,}5;3{,}5]\}$.



Par définition, cette probabilité est l'aire du domaine délimité par la courbe représentative de la densité $\mathscr{C}_f$, l'axe des abscisses et les droites d'équation $x=1{,}5$ et $x=3{,}5$.
Par suite :

$$P(1{,}5\leqslant X\leqslant 3{,}5)=\int_{1{,}5}^{3{,}5}\frac{1}{3}\,\mathrm{d}\,x=\frac{1}{3}\times(3{,}5-1{,}5)=\frac{2}{3}$$

On peut aussi retrouver ce résultat à l'aide de la fonction de répartition :

$$P(1{,}5\leqslant X\leqslant 3{,}5)=F(3{,}5)-F(1{,}5)=\frac{3{,}5-1}{4-1}-\frac{1{,}5-1}{4-1}=\frac{2}{3}$$

Ce calcul nécessite de connaître l'expression analytique de la fonction de répartition et de ses propriétés.

2) ***<u>Espérance</u>***

Après la détermination de la loi d'une probabilité d'une variable aléatoire, on s'intéresse au calcul de certaines de ses caractéristiques telles que l'espérance.

- Dans le cas d'une variable aléatoire discrète Y, l'espérance, notée E(Y) est définie par :

$$E(Y)=\sum_{\mathrm{i}=1}^{n} y_i\,P(Y=y_i)$$

où $(y_i)_{1\leqslant \mathrm{i}\leqslant n}$ désignent les *n* valeurs prises par la variable aléatoire Y.

<u>Exemple :</u>

Dans le 1er exemple du cours (le lancer de deux dés cubiques), l'espérance de la variable aléatoire considérée est égale à : $\frac{252}{36}=7$.

L'interprétation de ce résultat est la suivante : si on lance les deux dés cubiques un grand nombre de fois, la somme des nombres obtenus est en moyenne égale à 7.

- Dans le cas d'une variable aléatoire continue, cette somme n'a pas de sens. En effet, une telle variable peut prendre comme valeur tous les nombres réels d'un intervalle I. On a alors recours à la définition qui suit (prolongement naturel du cas discret).

<u>Définition :</u>

Soit *a* et *b* deux nombres réels distincts tels que $a<b$.

L'espérance d'une variable aléatoire continue X, à valeurs dans l'intervalle $I=[a\,;b]$, dont la densité associée est notée f , est définie par :

$$E(X)=\int_a^b x\,f(x)\,\mathrm{d}\,x$$

<u>Remarque :</u>

Cette définition de l'espérance est uniquement valable pour les variables aléatoires continues dont le support de la densité associée est un intervalle de la forme $[a\,;b]$, *a* et *b* étant deux nombres réels tels que $a<b$. C'est le cas pour les variables aléatoires qui suivent une loi uniforme.

Propriété :

Soit a et b deux nombres réels distincts tels que $a<b$.

L'espérance d'une variable aléatoire X suivant une loi uniforme sur l'intervalle $[a;b]$ est donnée par :

$$E(X)=\frac{a+b}{2}$$

Démonstration :

Par définition de l'espérance et de la densité d'une loi uniforme, on a :

$$E(X)=\int_a^b x\,f(x)\,\mathrm{d}x=\int_a^b \frac{x}{b-a}\,\mathrm{d}x=\left[\frac{x^2}{2(b-a)}\right]_a^b=\frac{b^2-a^2}{2(b-a)}=\frac{(b-a)(b+a)}{2(b-a)}=\frac{a+b}{2}$$

III. Loi exponentielle

Modèle : loi de durée de vie sans vieillissement.

Exemples : la durée de vie d'un système non sujet à l'usure du temps (fonctionnement de composants électroniques), temps d'attente d'un événement accidentel (tremblement de terre), désintégration d'un noyau radioactif...

1) Définition

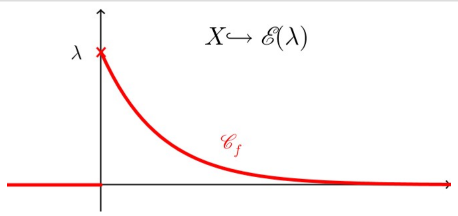
Définition :

Soit λ un réel strictement positif.

Dire qu'une variable aléatoire continue X suit une **loi exponentielle de paramètre λ** signifie que sa densité f est définie sur ℝ par :

$$f(x)=\begin{cases} 0 \text{ si } x<0 \\ \lambda e^{-\lambda x} \text{ si } x\geqslant 0 \end{cases}$$

On note :
X suit la loi $\mathcal{E}(\lambda)$.



Remarques :

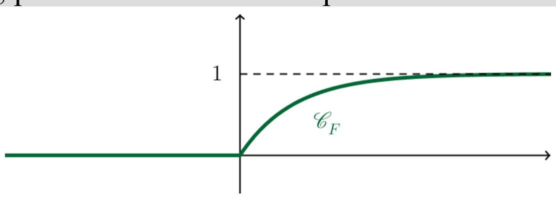
- Comme le support de la densité f est l'intervalle $I=[0;+\infty[$, une telle variable aléatoire prend alors ses valeurs dans cet intervalle.
- Le point M d'abscisse 0 qui appartient à la courbe représentative de la densité a pour ordonnée λ. Cette remarque s'avérera utile dans certains cas pour déterminer le paramètre d'une loi exponentielle.

Propriété :

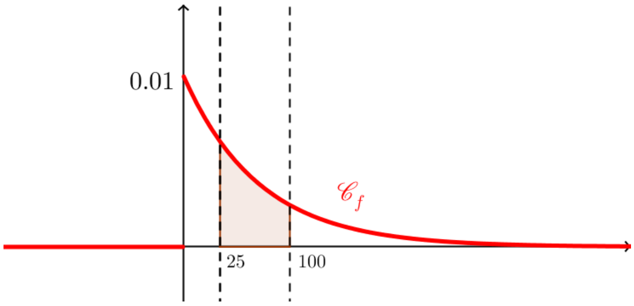
Soit λ un réel strictement positif.

Si une variable aléatoire X suit une loi exponentielle de paramètre λ, alors la fonction de répartition F associée à cette variable aléatoire est définie, pour tout nombre réel x par :

$$F(x)=\begin{cases} 0 \text{ si } x<0 \\ 1-e^{-\lambda x} \text{ si } x\geqslant 0 \end{cases}$$



Exemple :
On considère X une variable aléatoire continue suivant la loi exponentielle de paramètre 0,01. On désire calculer la probabilité de l'événement $\{X \in [25;100]\}$.



Par définition, cette probabilité est l'aire du domaine délimité par la courbe représentative de la densité $\mathscr{C}_f$, l'axe des abscisses et les droites d'équation $x=25$ et $x=100$.
Par suite :

$$P(25 \leqslant X \leqslant 100)=\int_{25}^{100} 0{,}01 e^{-0{,}01x} \mathrm{d}x=\left[-e^{-0{,}01x}\right]_{25}^{100}=e^{-0{,}25}-e^{-1}\simeq 0{,}4109$$

On peut aussi retrouver ce résultat à l'aide de la fonction de répartition :

$$P(25 \leqslant X \leqslant 100)=F(100)-F(25)=(1-e^{-0{,}01\times 100})-(1-e^{-0{,}01\times 25})=e^{-0{,}25}-e^{-1}\simeq 0{,}4109$$

Propriété (durée de vie sans vieillissement) :
Soit λ un réel strictement positif.
Si une variable aléatoire X suit une loi exponentielle de paramètre λ, alors cette variable aléatoire vérifie la propriété dite de **durée de vie sans vieillissement**, qui s'énonce ainsi :

$$P_{X\geqslant t}(X \geqslant t+h)=P(X \geqslant h)$$

pour tous réels *t* et *h* positifs.

Démonstration :
Par définition d'une probabilité conditionnelle, on a :

$$P_{X\geqslant t}(X \geqslant t+h)=\frac{P(\{X \geqslant t+h\}\cap\{X \geqslant t\})}{P(X \geqslant t)}=\frac{P(X \geqslant t+h)}{P(X \geqslant t)}$$

Or l'événement contraire de l'événement $\{X \geqslant t+h\}$ (resp $\{X \geqslant t\}$) est l'événement $\{X < t+h\}$ (resp $\{X<t\}$).
Il en découle :

$$P_{X\geqslant t}(X \geqslant t+h)=\frac{1-P(X<t+h)}{1-P(X<t)}=\frac{1-F(t+h)}{1-F(t)}=\frac{1-(1-e^{-\lambda(t+h)})}{1-(1-e^{-\lambda t})}=e^{-\lambda(t+h)+\lambda t}=e^{-\lambda h}$$

Or $\mathrm{e}^{-\lambda h}=1-(1-\mathrm{e}^{-\lambda h})=1-F(h)=1-P(X<h)=P(X \geqslant h)$.
D'où le résultat.

Remarque :
La durée de vie d'un appareil est dite « sans vieillissement » lorsque la probabilité qu'il fonctionne encore pendant une durée *h* (au moins) ne dépend que de *h* et pas de la durée *t* de son fonctionnement passé.

2) Espérance

Définition :
Soit a un nombre réel. L'espérance d'une variable aléatoire continue X à valeurs dans l'intervalle $I=[a;+\infty[$, dont la densité associée est notée f , est définie par :

$$E(X)=\lim_{x\to+\infty}\int_a^x t\,f(t)\,\mathrm{d}t$$

Remarque :
La définition précédente de l'espérance est à utiliser pour les variables aléatoires continues dont le support de la densité associée est un intervalle de la forme $[a;+\infty[$, a étant un nombre réel.
Tel est le cas pour les variables aléatoires continues qui suivent une loi exponentielle (a=0).

Propriété :
L'espérance d'une variable aléatoire X suivant une loi exponentielle de paramètre $\lambda>0$ est donnée par :

$$E(X)=\frac{1}{\lambda}$$

Démonstration :
Soit x un nombre réel strictement positif.
On appelle g la fonction définie sur $\mathbb{R}$ par $g(t)=t\,f(t)$, f désignant la densité de la loi exponentielle de paramètre $\lambda>0$. Comme cette fonction g est continue sur l'intervalle $[0;x]$, elle admet sur cet intervalle des primitives.
En outre, comme pour tout $t\geqslant 0$, $(t e^{-\lambda t})'=\mathrm{e}^{-\lambda t}-\lambda t e^{-\lambda t}$, il en découle :

$$\int_0^x g(t)\mathrm{d}t=\int_0^x t\lambda e^{-\lambda t}\mathrm{d}t=\int_0^x e^{-\lambda t}\mathrm{d}t-\int_0^x (t e^{-\lambda t})'\,\mathrm{d}t=\left[-\frac{e^{-\lambda t}}{\lambda}\right]_0^x-\left[t e^{-\lambda t}\right]_0^x=\frac{1}{\lambda}-\frac{e^{-\lambda x}}{\lambda}-x e^{-\lambda x}$$

Par passage à la limite quand x tend vers $+\infty$, on obtient le résultat.

IV. Loi normale

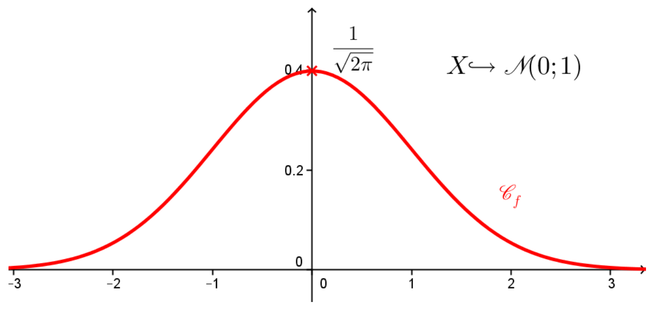
Exemples : phénomènes naturels très fréquents qui résultent de l'addition de plusieurs causes indépendantes (par exemple la taille d'un individu, le taux de cholestérol, les erreurs de mesure,...).

1) Loi normale centrée réduite

Définition :
Dire qu'une variable aléatoire continue X suit la **loi normale centrée réduite** signifie que sa densité f est définie sur $\mathbb{R}$ par :

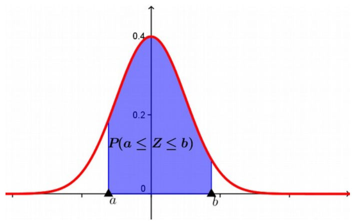
$$f(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$$

On note :
X suit la loi $\mathcal{N}(0;1)$



Remarques :

- Comme $f(x)=f(-x)$ pour tout réel *x*, la courbe représentative de la densité f est symétrique par rapport à l'axe des ordonnées.
  Elle est généralement désignée par l'expression « courbe en cloche ».
- Il n'est pas possible de déterminer les primitives de la densité f à l'aide de fonctions usuelles. On ne peut pas expliciter la fonction de répartition associée à une variable aléatoire qui suit la loi normale centrée réduite.

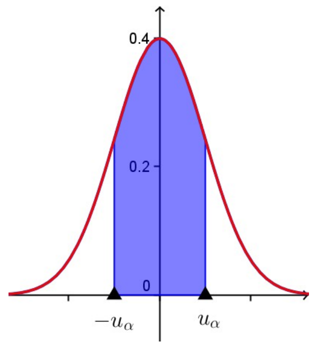


- Pour tous nombres a et b avec $a<b$, on a donc $P(a\leqslant X\leqslant b)=\int_a^b f(t)\,\mathrm{d}t$

Propriété :
Pour tout nombre réel α inclus dans l'intervalle $]0;1[$, il existe un unique réel positif u_α tel que :

$$P(-u_\alpha\leqslant X\leqslant u_\alpha)=1-\alpha$$

où X désigne une variable aléatoire suivant la loi normale centrée réduite.



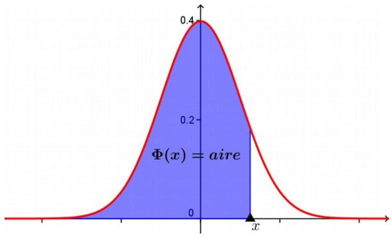
Démonstration :
Soit α un nombre réel inclus dans l'intervalle $]0;1[$. La densité f associée à la variable aléatoire X étant continue sur ℝ, on peut définir la fonction F sur $[0;+\infty[$ par $F(b)=\int_0^b \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x$, b étant un nombre réel positif.
Cette fonction F est l'unique primitive de f sur $[0;+\infty[$ qui s'annule en 0. Elle est continue et strictement croissante sur $[0;+\infty[$. Par propriété de la densité et par symétrie de la courbe représentative $\mathscr{C}_f$, il découle $\lim_{b\to+\infty}F(b)=\frac{1}{2}$.
L'image de 0 par F étant 0, F prend ses valeurs dans l'intervalle $\left[0;\frac{1}{2}\right]$. Comme le nombre réel α est strictement compris entre 0 et 1, le nombre réel $\frac{1-\alpha}{2}$ est strictement compris entre 0 et $\frac{1}{2}$.
Ainsi, d'après le théorème des valeurs intermédiaires, il existe un unique nombre réel strictement positif que l'on note u_α tel que $F(u_\alpha)=\frac{1-\alpha}{2}$.
Par symétrie de la courbe $\mathscr{C}_f$ et par linéarité de l'intégrale, on obtient le résultat :

$$P(-u_\alpha\leqslant X\leqslant u_\alpha)=\int_{-u_\alpha}^{u_\alpha}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x=\int_{-u_\alpha}^{0}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x+\int_0^{u_\alpha}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x=2\int_0^{u_\alpha}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x=2F(u_\alpha)=2\frac{1-\alpha}{2}=1-\alpha$$

Cas particuliers :

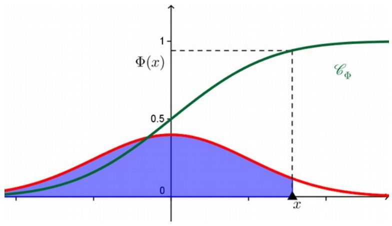
- Pour $\alpha=0{,}05$, une valeur approchée du nombre réel $u_{0,05}$ est 1,96 :
  $$u_{0,05}\simeq 1{,}96$$
- Pour $\alpha=0{,}01$, une valeur approchée du nombre réel $u_{0,01}$ est 2,58 :
  $$u_{0,01}\simeq 2{,}58$$

Fonction de répartition



Définition :
X étant une variable aléatoire qui suit la loi normale, on pose pour tout nombre x :

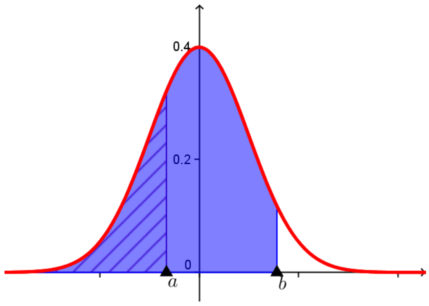
$$\Phi(x)=P(X\leqslant x)$$



Remarque :
La fonction ϕ s'appelle la **fonction de répartition** de X.

Propriété :
Si une variable aléatoire X suit la loi normale, alors pour tous nombres a et b tels que $a\leqslant b$,

$$P(a\leqslant X\leqslant b)=\Phi(b)-\Phi(a)$$



Démonstration :
ϕ(a) est l'aire du domaine sous la courbe, situé à gauche de l'abscisse a.
ϕ(b) est l'aire du domaine sous la courbe, situé à gauche de l'abscisse b.
Par soustraction, on obtient l'aire du domaine sous la courbe situé entre a et b.
On aurait pu aussi utiliser la formule d'intégration :

$$P(a\leqslant X\leqslant b)=\int_a^b f(t)\,\mathrm{d}t=\Phi(b)-\Phi(a)$$

Propriété :
Pour tout nombre réel x,

$$\Phi(-x)=1-\Phi(x)$$

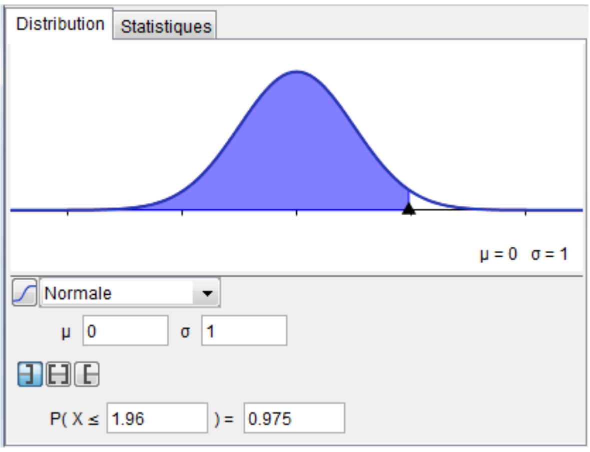
Démonstration :
$\Phi(-x)=P(Z\leqslant -x)$ par définition, et $P(Z\leqslant x)=P(Z<x)$.
Or, $P(Z<-x)=P(Z>x)$ par symétrie de la courbe en cloche par rapport à l'axe des ordonnées : donc $\Phi(-x)=P(Z>x)$.
« $Z>x$ » est l'événement contraire de « $Z\leqslant x$ », donc $P(Z>x)=1-P(Z\leqslant x)$. D'où le résultat.

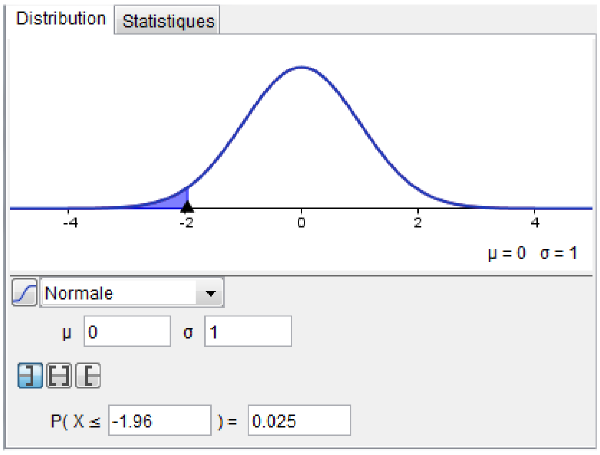
Remarque :

On détermine u_α en utilisant la fonction de répartition ϕ et en recherchant un antécédent.

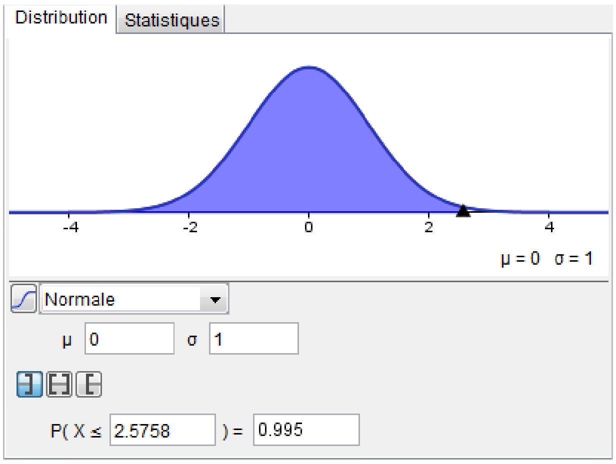
On utilise $P(X \leqslant u_\alpha) = \Phi(u_\alpha) = 1 - \frac{\alpha}{2}$

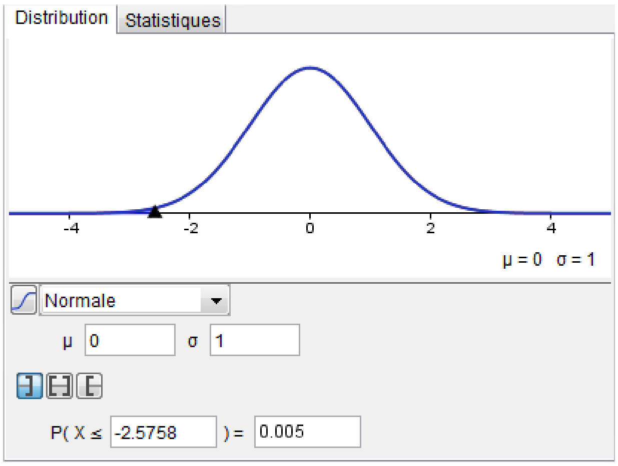
- $\alpha = 0{,}05$





- $\alpha = 0{,}01$





- **Calculatrice**

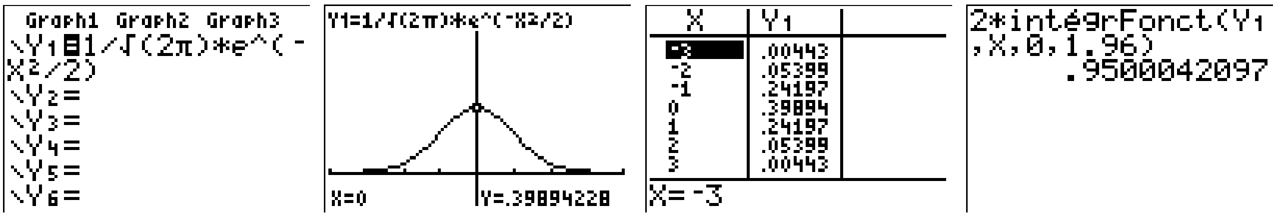
Densité :

- **NormPD(X)=** $f(X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{X^2}{2}}$ **=normalFdp(X)**

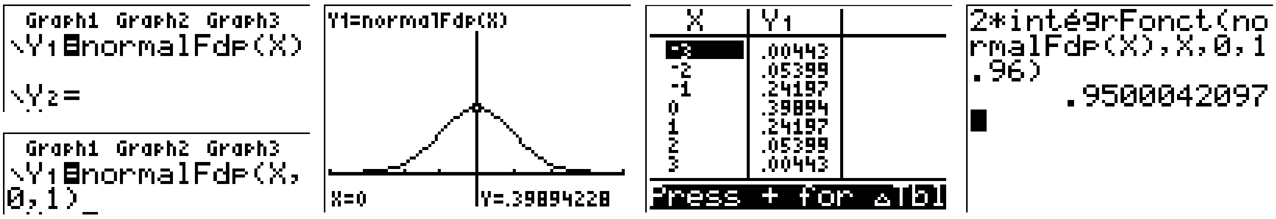
Fonction de répartition :

- **NormCD(A,B)=** $P(A \leqslant X \leqslant B) = F(B) - F(A) = \int_A^B \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ **=normalFRep(A,B)**
- **InvNormCD(P)** renvoit t tel que $P(X \leqslant t) = P = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ **:FracNormale(P)**

Fonction densité

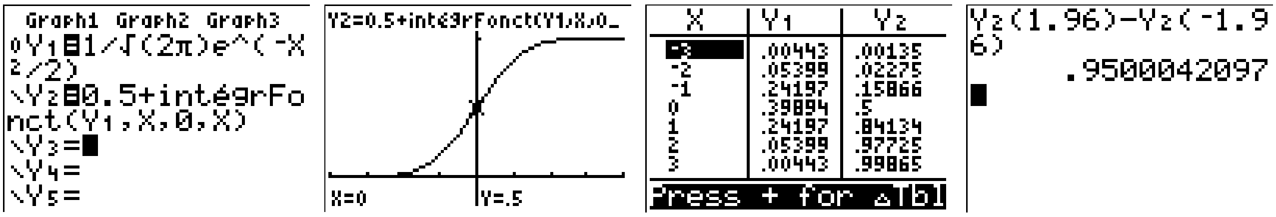


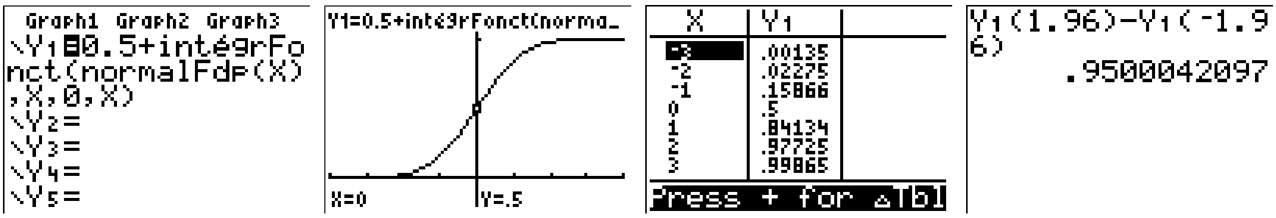
ou

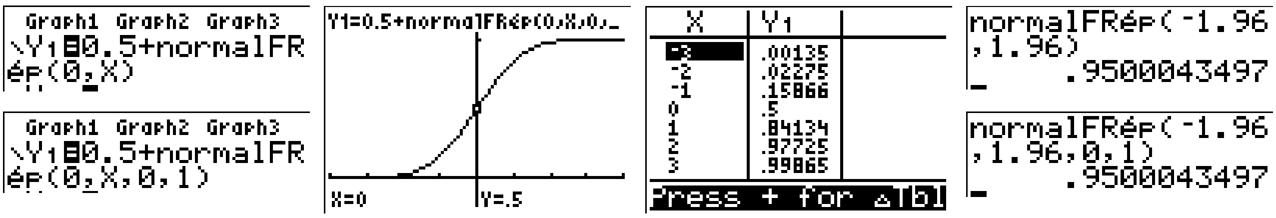


Fonction de répartition

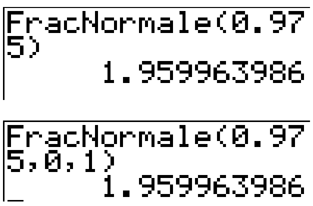
Calcul de probabilités



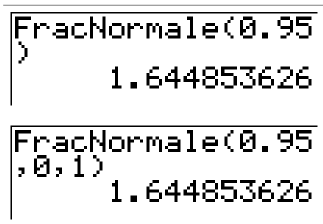




Calcul d'intervalle

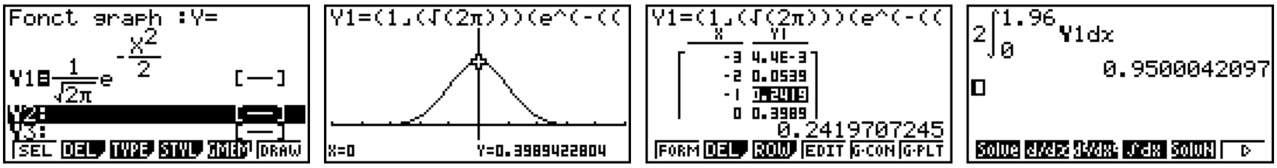


$$P(X\leqslant 1{,}959963986)=\Phi(1{,}959963986)=\int_{-\infty}^{1{,}959963986}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x\simeq 0{,}975$$

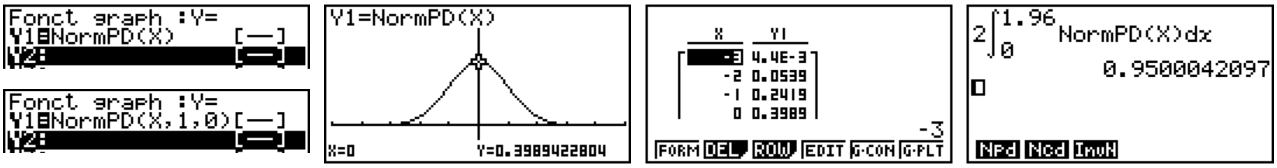


$$P(X\leqslant 1{,}644853626)=\Phi(1{,}644853626)=\int_{-\infty}^{1{,}644853626}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x\simeq 0{,}95$$

Fonction densité

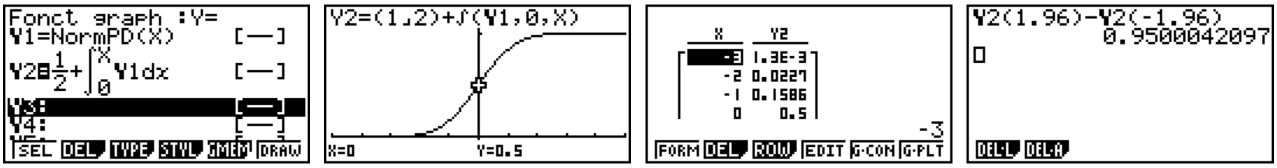


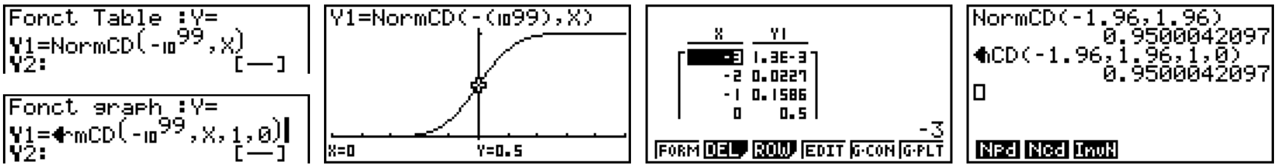
ou



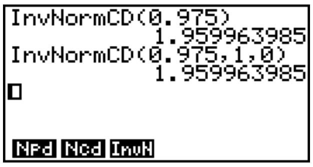
Fonction de répartition

- Calcul de probabilités

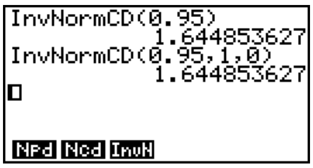




- Calcul d'intervalle



$$P(X \leqslant 1{,}959963986) = \Phi(1{,}959963986) = \int_{-\infty}^{1{,}959963986} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \,\mathrm{d}x \simeq 0{,}975$$



$$P(X \leqslant 1{,}644853626) = \Phi(1{,}644853626) = \int_{-\infty}^{1{,}644853626} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \,\mathrm{d}x \simeq 0{,}95$$

Espérance

<u>Définition :</u>
L'espérance d'une variable aléatoire continue X à valeurs dans ℝ dont la densité associée est notée f , est définie par :

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

<u>Propriété :</u>
Soit X une variable aléatoire continue suivant la loi normale centrée réduite $\mathcal{N}(0;1)$.
Alors $E(X)=0$.

Démonstration :
Soit x et y deux nombres réels respectivement négatif et positif.
On appelle g la fonction définie sur ℝ par $g(t)=t\,f(t)$, f désignant la densité associée à la variable aléatoire X. Cette fonction g étant continue sur ℝ, elle admet des primitives également définies sur ℝ, à une constante additive près par : $-\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}$

Par conséquent, $\int_0^y g(t)\,\mathrm{d}t=\left[-\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\right]_0^y=\frac{1}{\sqrt{2\pi}}\left(1-\mathrm{e}^{-\frac{y^2}{2}}\right)$.

Par passage à la limite, $\lim_{y\to+\infty}\int_0^y g(t)\,\mathrm{d}t=\frac{1}{\sqrt{2\pi}}$.

De même, $\int_x^0 g(t)\,\mathrm{d}t=\frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{x^2}{2}}-1\right)$.

Par passage à la limite, $\lim_{x\to-\infty}\int_x^0 g(t)\,\mathrm{d}t=-\frac{1}{\sqrt{2\pi}}$.

Ainsi $E(X)=-\frac{1}{\sqrt{2\pi}}+\frac{1}{\sqrt{2\pi}}=0$.

<u>Remarque :</u>
Similairement au cas des variables aléatoires discrètes, on peut définir la variance et l'écart-type d'une variable aléatoire continue.
La variance est notée $V(X)$ et elle est définie par $V(X)=E((X-E(X))^2)$.
L'écart-type est noté $\sigma(X)$ et il est égal à $\sqrt{V(X)}$.
Si X suit la loi normale centrée réduite, on admettra que la variance $V(X)$ est égale à 1.
Les deux paramètres de la loi normale centrée réduite correspondent en fait à l'espérance et à l'écart-type (au carré) de la variable aléatoire continue X.

2) *Loi normale*

Définition :

Dire qu'une variable aléatoire continue X suit la **loi normale** d'espérance μ et d'écart-type σ signifie que la variable aléatoire continue $\frac{X-\mu}{\sigma}$ suit la loi normale centrée réduite.

On note :

X suit la loi $\mathcal{N}(\mu\ ;\ \sigma^2)$.

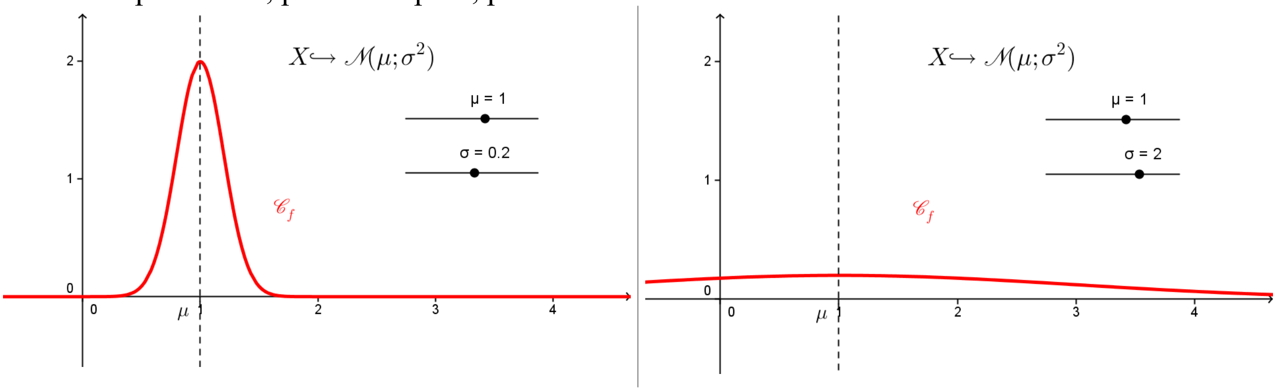
Remarques :

- On dit aussi loi de Laplace-Gauss, loi de Gauss ou loi gaussienne.

- La densité associée à une variable aléatoire continue suivant la loi normale d'espérance μ et d'écart-type σ a une expression similaire à celle de la loi normale centrée réduite.

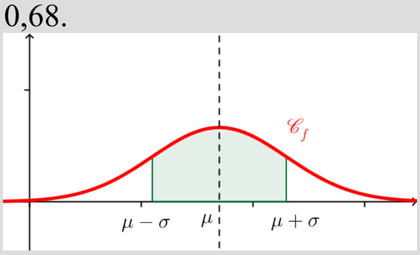
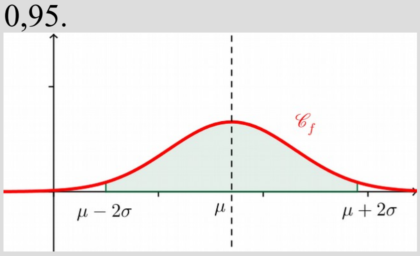
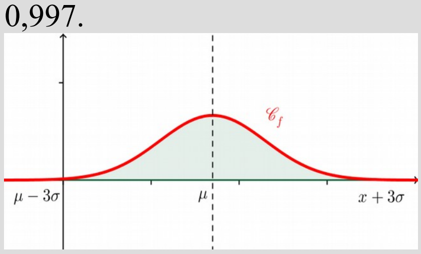
  En effet, cette densité est définie sur ℝ par $f(x)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$.

  On remarque que, pour tout nombre réel x, $f(\mu-x)=f(\mu+x)$.
  Par conséquent, la courbe représentative de cette densité, désignée aussi par le terme « courbe en cloche », admet la droite d'équation $x=\mu$ pour axe de symétrie.
  L'écart-type σ a, quant à lui, un impact sur la forme de la « cloche ».
  En particulier, plus σ est petit, plus la « cloche » est « haute ».



Propriété :

Soit X une variable aléatoire continue suivant la loi normale $\mathcal{N}(\mu\ ;\ \sigma^2)$.

La probabilité de l'événement $\{X\in[\mu-\sigma;\mu+\sigma]\}$ est approximativement égale à 0,68. 	La probabilité de l'événement $\{X\in[\mu-2\sigma;\mu+2\sigma]\}$ est approximativement égale à 0,95. 	La probabilité de l'événement $\{X\in[\mu-3\sigma;\mu+3\sigma]\}$ est approximativement égale à 0,997. 

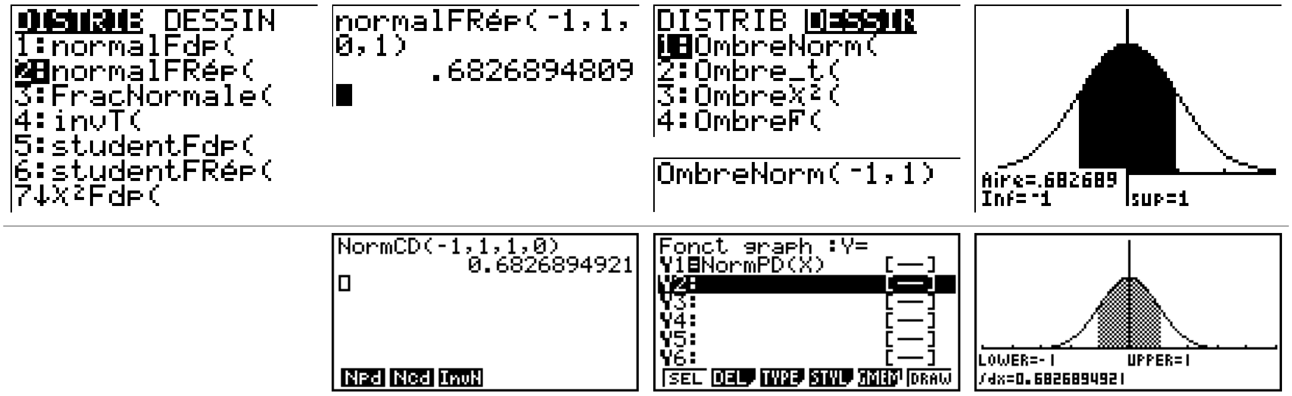
Démonstration :

$P(X \in [\mu-\sigma;\mu+\sigma]) = P\left(-1 \leqslant \frac{X-\mu}{\sigma} \leqslant 1\right)$.

Comme X suit la loi normale $\mathcal{N}(\mu\ ;\ \sigma^2)$, la variable aléatoire $\frac{X-\mu}{\sigma}$ suit la loi normale $\mathcal{N}(0\ ;\ 1)$.

Ainsi, par définition, la probabilité de l'événement $\left\{-1 \leqslant \frac{X-\mu}{\sigma} \leqslant 1\right\}$ est l'aire du domaine délimité par la courbe représentative de la densité associée à la loi normale centrée réduite, l'axe des abscisses et les droites d'équation $x=-1$ et $x=1$. ce qui se traduit à l'aide d'une intégrale par l'égalité $P\left(-1 \leqslant \frac{X-\mu}{\sigma} \leqslant 1\right) = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathrm{d}x$.

Or, on ne sait pas expliciter une primitive de la densité associée à la loi normale centrée réduite.
Il n'est donc pas possible de déterminer la valeur exacte de cette probabilité. On peut tout de même en obtenir une valeur approchée en utilisant (par exemple) la calculatrice qui nous renvoie ici une valeur légèrement supérieure à 0,68.



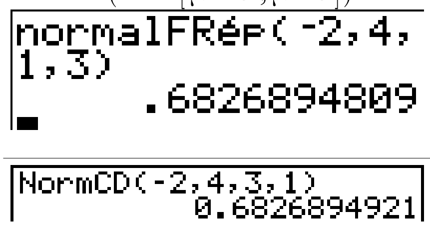
Ainsi quel que soit le réel μ, quel que soit le réel strictement positif σ, on a :

$$P(X \in [\mu-\sigma;\mu+\sigma]) \geqslant 0{,}68 .$$

Calculatrice :

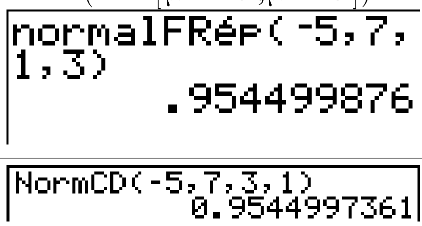
Pour $\mu=1$ et $\sigma=3$

$P(X \in [\mu-\sigma;\mu+\sigma])$



$P(X \in [-2;4]) \simeq 0{,}68$

$P(X \in [\mu-2\sigma;\mu+2\sigma])$

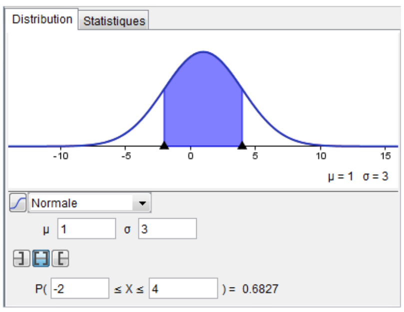


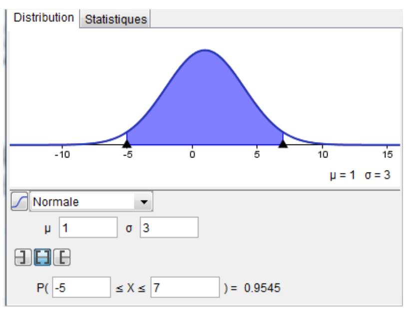
$P(X \in [-5;7]) \simeq 0{,}95$

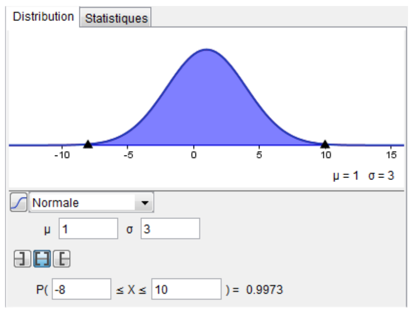
$P(X \in [\mu-3\sigma;\mu+3\sigma])$

$P(X \in [-8;10]) \simeq 0{,}997$

Logiciel :

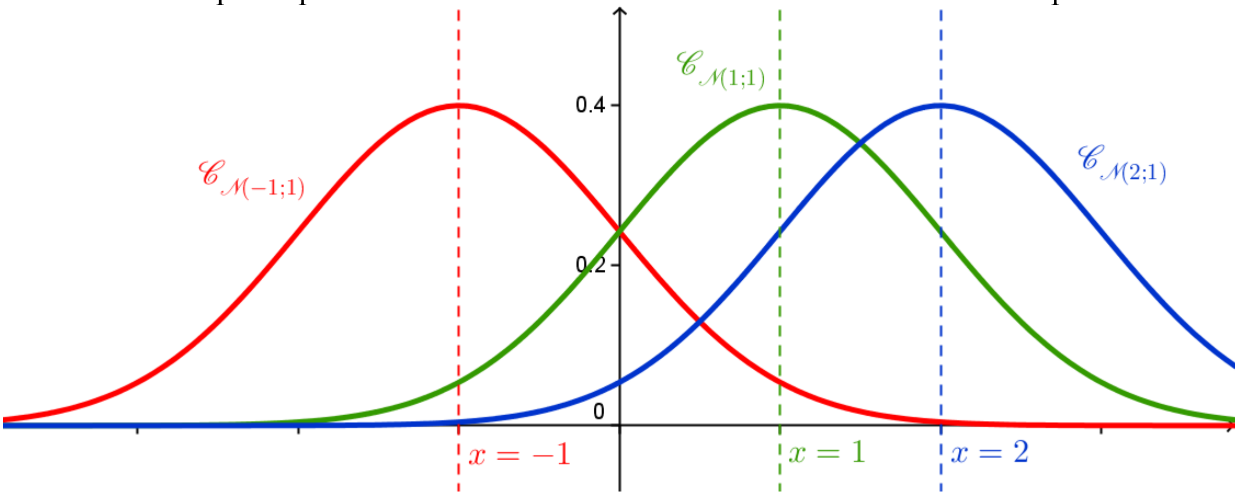




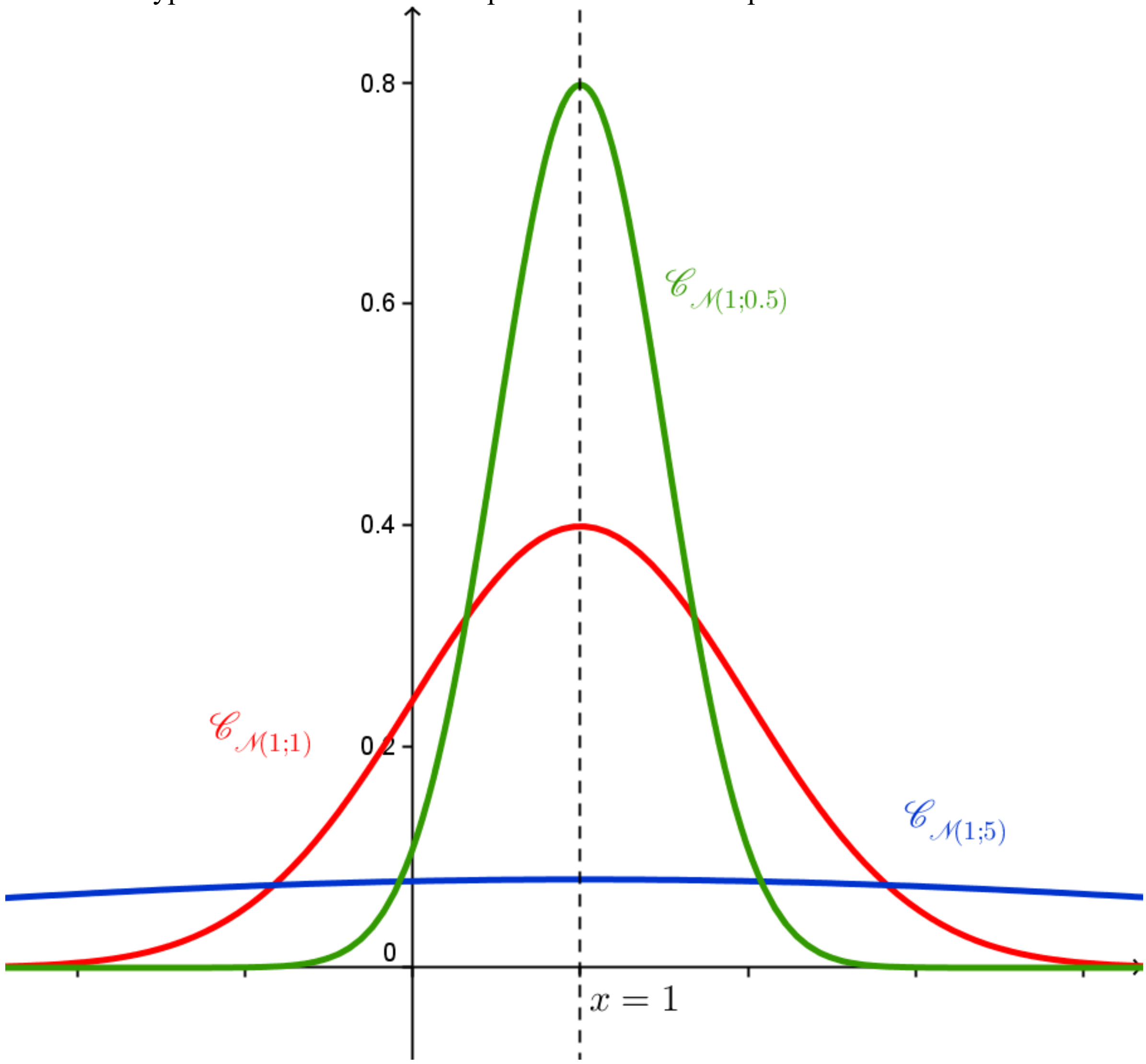


Remarque :

- L'espérance est un indicateur de position :
  les valeurs prises par la variable aléatoire sont concentrées autour de cette espérance.



- L'écart-type est un indicateur de dispersion autour de l'espérance.



3) Approximation normale d'une loi binomiale

Théorème de Moivre-Laplace :
n est un entier naturel non nul et p un nombre réel fixé appartenant à l'intervalle $[0;1]$.
Soit X_n une variable aléatoire suivant la loi binomiale $\mathscr{B}(n;p)$.
Alors pour tous nombres réels a et b tels que $a<b$:

$$\lim_{n \to +\infty} P\left(a \leqslant \frac{X_n - np}{\sqrt{np(1-p)}} \leqslant b\right) = P(a \leqslant Z \leqslant b) = \int_a^b \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \mathrm{d}x$$

où Z suit la loi normale centrée réduite $\mathscr{N}(0;1)$.

Propriété :
On considère que la limite dans le théorème de Moivre-Laplace est pratiquement atteinte lorsqu'on a simultanément $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$. Dans ces conditions :
soit X_n une variable aléatoire suivant la loi binomiale $\mathscr{B}(n;p)$.
Alors pour tous nombres réels a et b tels que $a<b$:

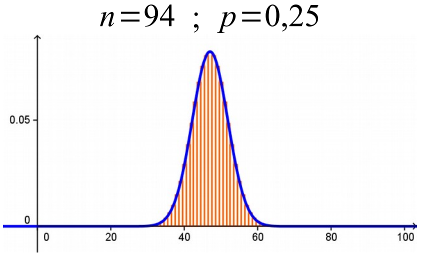
$$P\left(a \leqslant \frac{X_n - np}{\sqrt{np(1-p)}} \leqslant b\right) \simeq P(a \leqslant Z \leqslant b)$$

où Z suit la loi normale centrée réduite $\mathscr{N}(0;1)$.

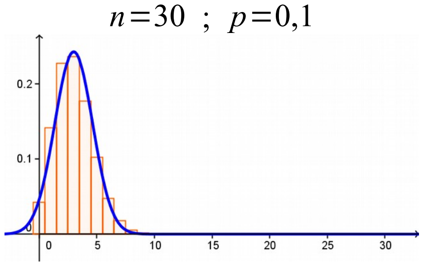
Remarque :
Ce théorème justifie que, sous certaines conditions sur les paramètres n et p, la probabilité d'un événement associé à une loi binomiale peut être approchée par une probabilité d'un événement associé à la loi normale centrée réduite.
On parle d'**approximation d'une loi binominale par une loi normale**.
L'intérêt de cette approximation est, entre autres, de simplifier les calculs numériques. Ce phénomène n'est pas spécifique à la loi binomiale. Il s'observe dès lors que l'on est dans une situation de répétition d'expériences identiques et indépendantes.

Remarques :
- Pour les grandes valeurs de n, la loi binomiale $\mathscr{B}(n;p)$ est très proche de la loi normale de même espérance np et de même variance $np(1-p)$.
- Quand on a $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$, l'erreur sur les probabilités calculées est très faible. On ne fera l'approximation que lorsque ces trois conditions sont vérifiées.

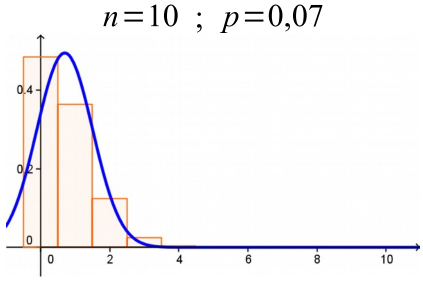
$n=94$; $p=0,25$



$n=30$; $p=0,1$



$n=10$; $p=0,07$



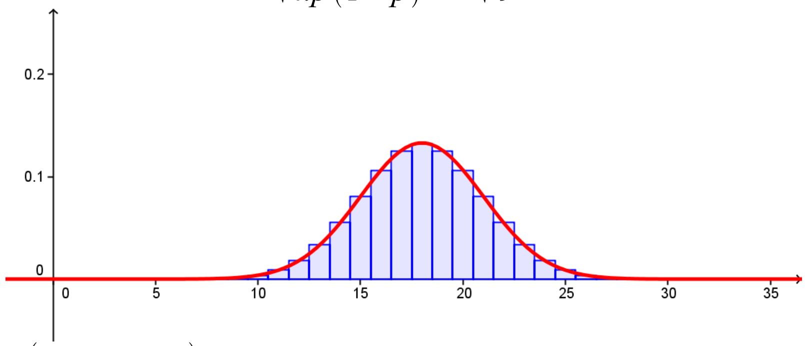
Exemple :

$n=36$, $p=0{,}5$. X suit la loi binomiale $\mathscr{B}(36\,;0{,}5)$.

Les trois conditions sont remplies : $n\geqslant 30$, $np\geqslant 5$ et $n(1-p)\geqslant 5$.

De plus, $E(X)=np=18$ et $V(X)=np(1-p)=9$.

On peut donc centrer et réduire X : $Z=\dfrac{X-np}{\sqrt{np(1-p)}}=\dfrac{X-18}{\sqrt{9}}=\dfrac{X-18}{3}$.



- Calculons $P(17\leqslant X\leqslant 19)$.
  Par l'approximation normale, il faut d'abord l'écrire $P(16{,}5\leqslant X\leqslant 19{,}5)$. En effet, on remplace un diagramme en bâtons par un histogramme, donc chaque segment vertical est remplacé par un rectangle vertical de même hauteur et de largeur 1.
  C'est ce qu'on appelle la **correction de continuité**.
  Donc $16{,}5\leqslant X\leqslant 19{,}5 \Leftrightarrow \dfrac{16{,}5-18}{3}\leqslant\dfrac{X-18}{3}\leqslant\dfrac{19{,}5-18}{3} \Leftrightarrow -0{,}5\leqslant Z\leqslant 0{,}5$
  $$P(16{,}5\leqslant X\leqslant 19{,}5)=P(-0{,}5\leqslant Z\leqslant 0{,}5)=\Phi(0{,}5)-\Phi(-0{,}5)\simeq 0{,}6915-0{,}3085=0{,}383$$
- Proposons un intervalle I tel que la probabilité de l'événement $\{X\in I\}$ soit égale à 0,95.
  Par l'approximation normale, $\mathscr{B}(36\,;0{,}5)$ est très proche de $\mathscr{N}(18\,;3)$ on sait que la probabilité de l'événement $\{X\in[18-2\times 3\,;18+2\times 3]\}$ est approximativement égale à 0,95.
  Donc $I\simeq[12\,;24]$.
  Il s'agit ici d'un intervalle centré sur $E(X)$.

Calculatrice :

- **Calcul de probabilité**

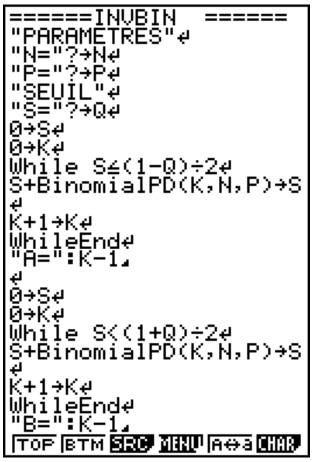
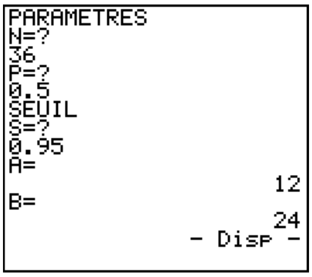
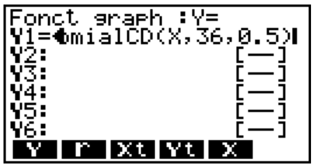
<table>
<tr><th colspan="3">Calcul à partir de la loi binomiale</th><th>Approximation</th></tr>
<tr>
<td rowspan="2">PROGRAM:BINOMIAL
:Disp "PARAMETRE
S"
:Prompt N,P
:Disp "BORNES"
:Prompt A,B
:0→S
:For(K,A,B)
:S+(N Combinaiso
n K)*P^K*(1-P)^(
N-K)→S
:End
:Disp S</td>
<td rowspan="2">prgmBINOMIAL

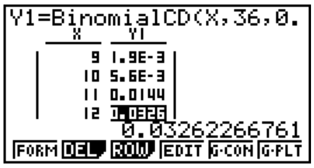
PARAMETRES
N=?36
P=?0.5
BORNES
A=?17
B=?19
.382280683
Done</td>
<td rowspan="2">binomFRép(36,0.5
,19)-binomFRép(3
6,0.5,16)
.3822806932</td>
<td>Loi normale
normalFRép(16.5,
19.5,18,3)
.3829249356</td>
</tr>
<tr><td>Loi normale centrée réduite
normalFRép(-0.5,
0.5,0,1)
.3829249356</td></tr>
</table>

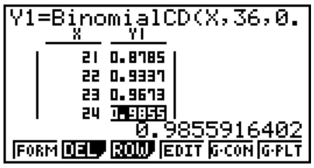
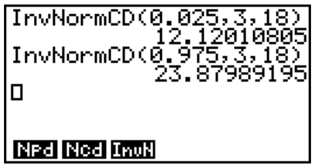
Calcul à partir de la loi binomiale			Approximation
======BINOMIAL====== "PARAMETRES"↵ "N="?→N↵ "P="?→P↵ "BORNES"↵ "A="?→A↵ "B="?→B↵ 0→S↵ For A→K To B↵ S+BinomialPD(K,N,P)→S ↵ Next↵ "PROBA=":S◢ TOP BTM SRC MENU A⇔a CHAR	PARAMETRES N=? 36 P=? 0.5 BORNES A=? 17 B=? 19 PROBA= 0.382280683 - Disp -	◀nomialCD(16,36,0.5) 0.382280683 □ DEL·L DEL·A	Loi normale BinomialCD(19,36,0.5▷ 0.382280683 ◀CD(16.5,19.5,3,18) 0.3829249225 □ Npd Ncd InvN Loi normale centrée réduite BinomialCD(19,36,0.5▷ 0.382280683 NormCD(16.5,19.5,3,1▷ 0.3829249225 NormCD(-0.5,0.5) 0.3829249225 □ Npd Ncd InvN

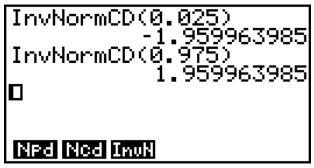
- **Détermination d'intervalle**

Calcul à partir de la loi binomiale			Approximation
PROGRAM:INVBINOM :Disp "PARAMETRE S" :Prompt N,P :Disp "SEUIL" :Input "S=",Q :0→S :0→K :While S≤(1-Q)/2 :S+(N Combinaiso n K)*P^K*(1-P)^(N-K)→S :K+1→K :End :0→R :0→I :While R<(1+Q)/2 :R+(N Combinaiso n I)*P^I*(1-P)^(N-I)→R :I+1→I :End :Disp "A=",K-1 :Disp "B=",I-1	prgmINVBINOM PARAMETRES N=?36 P=?0.5 SEUIL S=0.95 A= 12 B= 24 Done	Graph1 Graph2 Graph3 \Y1=binomFRép(36 ,0.5,X) \Y2= \Y3= \Y4= \Y5= \Y6= X Y1 6 3.5E-5 7 1.6E-4 8 6E-4 9 .00197 10 .00567 11 .01441 12 .03262 Y1=.032622667151 X Y1 20 .79748 21 .87851 22 .93375 23 .96738 24 .98559 25 .99433 26 .99803 Y1=.985591640213	Loi normale FracNormale(0.02 5,18,3) 12.12010804 FracNormale(0.97 5,18,3) 23.87989196 Loi normale centrée réduite FracNormale(0.02 5,0,1) -1.959963986 FracNormale(0.97 5,0,1) 1.959963986 $-1{,}96 \leqslant \frac{X-18}{3} \leqslant 1{,}96$ $-5{,}88 \leqslant X-18 \leqslant 5{,}88$ $12{,}12 \leqslant X \leqslant 23{,}88$

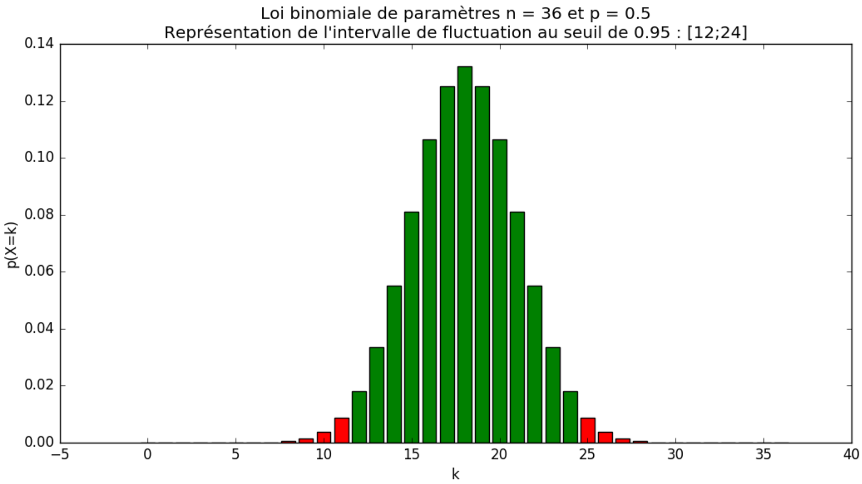
<table>
<tr><th colspan="3">Calcul à partir de la loi binomiale</th><th>Approximation</th></tr>
<tr>
<td>
</td>
<td>
</td>
<td>




</td>
<td>Loi normale


Loi normale centrée réduite


$-1{,}96\leqslant\frac{X-18}{3}\leqslant 1{,}96$
$-5{,}88\leqslant X-18\leqslant 5{,}88$
$12{,}12\leqslant X\leqslant 23{,}88$</td>
</tr>
</table>



Remarques :

- On rappelle que lorsque X suit une loi binomiale $\mathscr{B}(n;p)$, la loi de probabilité est donnée par :

$$p(X=k)=\binom{n}{k}p^k(1-p)^{n-k} \text{ et donc } p(X\leqslant k)=\sum_{i=0}^{k}\binom{n}{i}p^i(1-p)^{n-i}$$

- La détermination de probabilités ou d'intervalles à partir des propriétés de la loi binomiale ont le mérite de donner des résultats exacts mais le problème réside dans la complexité des calculs (surtout lorsque n est très grand ou p très petit, sans compter les coefficients binomiaux qui représentent rapidement des nombres très grands). Ses calculs sont donc difficiles à effectuer. Les approximations par la loi normale sont donc pratiques et la marge d'erreur (d'autant plus petite que n est grand) reste acceptable.
- Pour information :
    - $\binom{n}{k}=\frac{n!}{k!(n-k)!}$.
    - Formule de Stirling : pour n grand, $n!\simeq\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

Python (loi binomiale) :

Les calculs sont valables pour des valeurs de n et p limitées

```python
#On importe la fonction factorielle
from math import factorial

# Combinaison : version recursive
def combinRec(n, k):
    """Nombre de combinaisons de k objets parmi n (calcul récursif)"""
    if k == 0 or k == n:
        return 1
    return combinRec(n-1, k-1) + combinRec(n-1, k)

# Combinaison : version itérative
def combin(n, k):
    """Nombre de combinaisons de k objets parmi n (calcul itératif)"""
    return factorial(n) / (factorial(k) * factorial(n - k))

# Loi binomiale
def binom(k, n, p):
    """p(X=k) ou X est une v.a. qui suit la loi binomiale B(n,p)"""
    return combin(n,k) * pow(p,k) * pow(1-p,n-k)

# Loi binomiale cumulée
def binomCum(k, n, p):
    """p(X<=k) ou X est une v.a. qui suit la loi binomiale B(n,p)"""
    s = 0
    for i in range(0,k+1):
        s += binom(i,n,p)
    return s

# Inverse binomiale
def invBin(n, p, s):
    """Intervalle centré [a;b] tel que p(a<=X<=b) >= s
    ou X est une v.a. qui suit la loi binomiale B(n,p)"""
    # Calcul de a
    somme = 0
    k = 0
    while somme <= (1-s)/2:
        somme += binom(k,n,p)
        k += 1
    a = k - 1
    # Calcul de b
    somme = 0
    k = 0
    while somme < (1+s)/2:
        somme += binom(k,n,p)
        k += 1
    b = k - 1
    return a, b
```

```
>>> binom(17,36,0.5)+binom(18,36,0.5)+binom(19,36,0.5)
0.38228068297030404
>>> binomCum(19,36,0.5)-binomCum(16,36,0.5)
0.38228068297030404
>>> invBin(36,0.5,0.95)
(12, 24)
>>>
```

Python (loi normale) :

```
# On importe les items nécessaires pour définir la loi normale
from math import sqrt, pi, exp

# On définit la fonction de densité de la loi normale
def normale(x, σ=1, μ=0):
    """ Loi Normale par défaut espérance (μ) = 0 et écart-type (σ) = 1"""
    return 1/(σ*sqrt(2*pi))*exp(-0.5*((x-μ)/σ)**2)

# On implémente la méthode des rectangles pour l'intégration numérique
def integRect(f, a, b, n=1000):
    """ Approximation de l'intégrale de f(x) sur [a;b]
    n détermine le nombre de subdivisions sur l'intervalle (par défaut 1000)"""
    somme = 0
    dx = (b-a)/n
    for k in range (n):
        somme += f(a+(k*dx))*dx
    return somme

# Fonction de répartition de la loi normale
def normCum(a, b, σ=1, μ=0, n=1000):
    """p(a<=X<=k) ou X est une v.a. qui suit la loi normale N(σ,μ)
    n détermine le nombre de subdivisions sur l'intervalle (par défaut 1000)"""
    densite = lambda x: normale(x,σ,μ)
    return integRect(densite,a,b,n)

# Inverse normale
def invNorm(s, σ=1, μ=0, n=1000):
    """Intervalle centré [a;b] tel que p(a<=X<=b) >= s
    ou X est une v.a. qui suit la loi normale N(σ,μ)
    n détermine le nombre de subdivisions de l'unité"""
    densite = lambda x: normale(x,σ,μ)
    # Calcul de b
    x = μ
    dx = 1/n
    somme = 0
    k = 0
    while somme < s / 2:
        somme += densite(x+(k*dx))*dx
        k += 1
    b = μ + (k – 1)*dx
    # Calcul de b
    a = μ - (k - 1)*dx
    return a, b
```

```
>>> normale(0)
0.3989422804014327
>>> normale(0,3,1)
0.12579440923099774
>>> integRect(normale,-1.96,1.96)
0.9500039163481592
>>> integRect(normale,-1.96,1.96,10000)
0.9500042067700077
>>> normCum(-1.96,1.96)
0.9500039163481592
>>> normCum(-1.96,1.96,n=10000)
0.9500042067700077
>>> invNorm(0.95)
(-1.957, 1.957)
>>> invNorm(0.95,n=10000)
(-1.9596, 1.9596)
>>> normCum(16.5,19.5,3,18)
0.3829248932092479
>>> invNorm(0.95,3,18)
(12.123999999999999, 23.876)
>>>
```

Annexe 1 : Table de la loi normale

La table suivante donne les valeurs de la fonction de répartition $\Phi(x)=P(X\leqslant x)$, lorsque X suit la loi normale centrée réduite $\mathcal{N}(0;1)$.
Les valeurs en début de lignes donnent la première partie de la variable, les valeurs en début de colonnes donnent la deuxième partie.
Ainsi la case de la deuxième ligne et troisième colonne donne : $\Phi(0{,}12)=0{,}54776$

$\Phi(x)$	**0,00**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,99990	0,99990	0,99990	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997

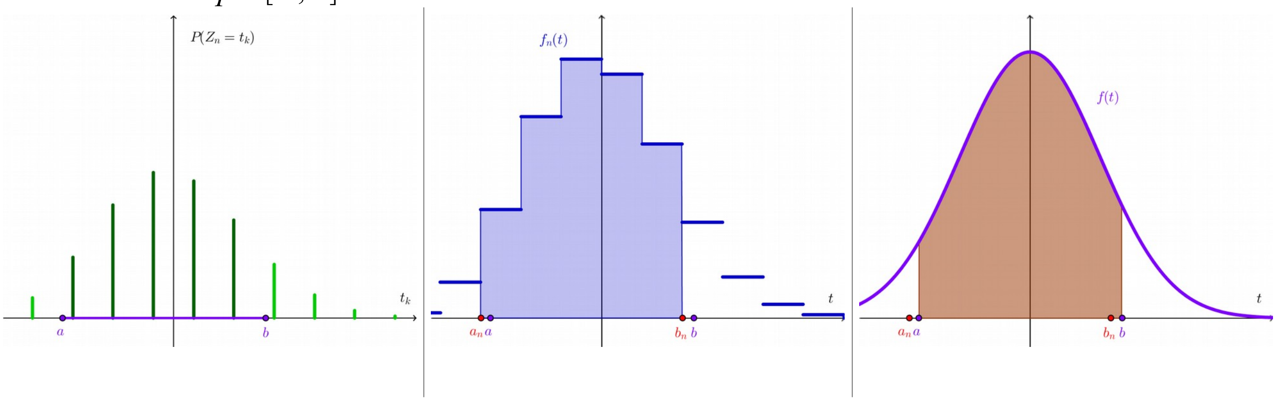
Annexe 2 : Théorème de Moivre Laplace

<u>Théorème de Moivre-Laplace :</u>
n est un entier naturel non nul et *p* un nombre réel fixé appartenant à l'intervalle $[0;1]$.
Soit X_n une variable aléatoire suivant la loi binomiale $\mathscr{B}(n\,;p)$.
Alors pour tous nombres réels *a* et *b* tels que $a<b$:

$$\lim_{n\to+\infty} P\left(a \leqslant \frac{X_n - np}{\sqrt{np(1-p)}} \leqslant b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \,\mathrm{d}x$$

Idée de la démonstration :

- Pour tout $p \in [0;1]$

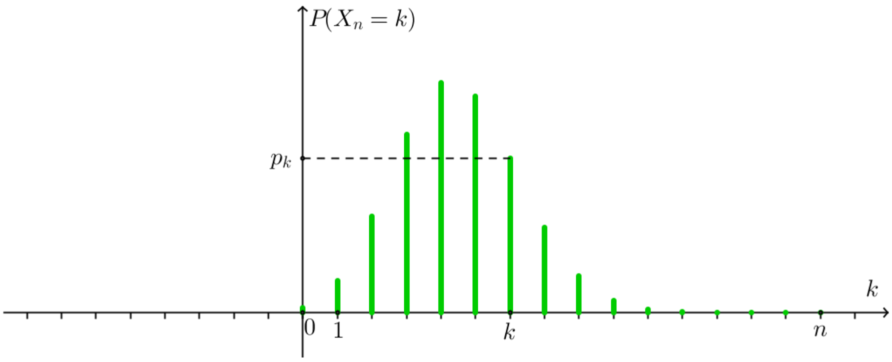


Pour tout $n \in \mathbb{N}^*$, $\mathrm{P}(a \leqslant Z_n \leqslant b) = \int_{a_n}^{b_n} f_n(t)\,\mathrm{d}t$

$$\lim_{n\to+\infty} \left| \int_{a_n}^{b_n} f_n(t)\,\mathrm{d}t - \int_a^b \frac{\mathrm{e}^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t \right| = 0$$

- **Variable centrée réduite**

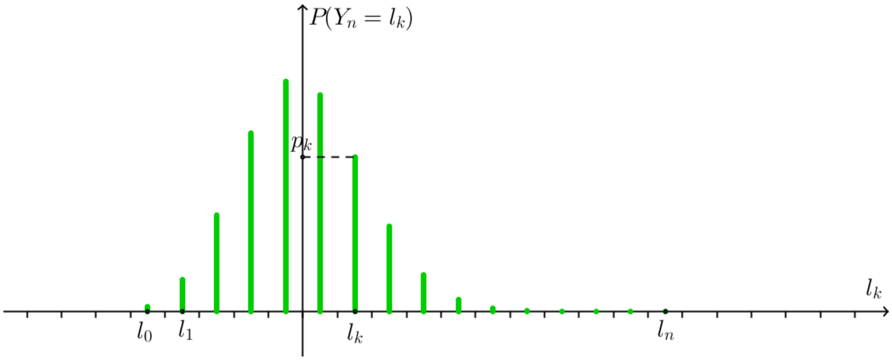
$X_n \in \{0;1;\dots;k;\dots;n\}$ $\quad p(X_n=k)=\binom{n}{k} p^k (1-p)^{n-k}$ $\quad \mu=np \;;\; \sigma=np(1-p)$



On « centre » la variable : $Y_n=X_n-np$

$Y_n \in \{\, l_0 \;;\, l_1 \;;\dots;\, l_k \;;\dots;\, l_n \,\}$
avec $l_k=k-np$

$p(Y_n=l_k)=p(X_n=k)$ $\quad \mu=0 \;;\; \sigma=np(1-p)$



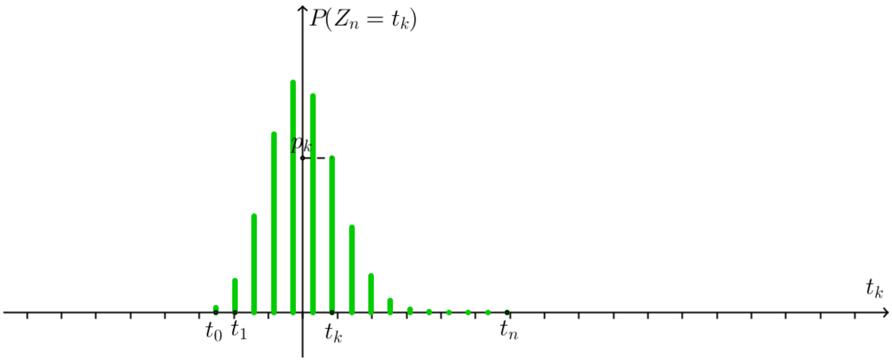
$l_{k+1}=l_k+1$

On « réduit » la variable : $Z_n=\frac{Y_n}{\sqrt{np(1-p)}}=\frac{X_n-np}{\sqrt{np(1-p)}}$

$Z_n \in \{\, t_0 \;;\, t_1 \;;\dots;\, t_k \;;\dots;\, t_n \,\}$
avec

$$t_k=\frac{l_k}{\sqrt{np(1-p)}}=\frac{k-np}{\sqrt{np(1-p)}}$$

$p(Z_n=t_k)=p(Y_n=l_k)=p(X_n=k)$ $\quad \mu=0 \;;\; \sigma=1$



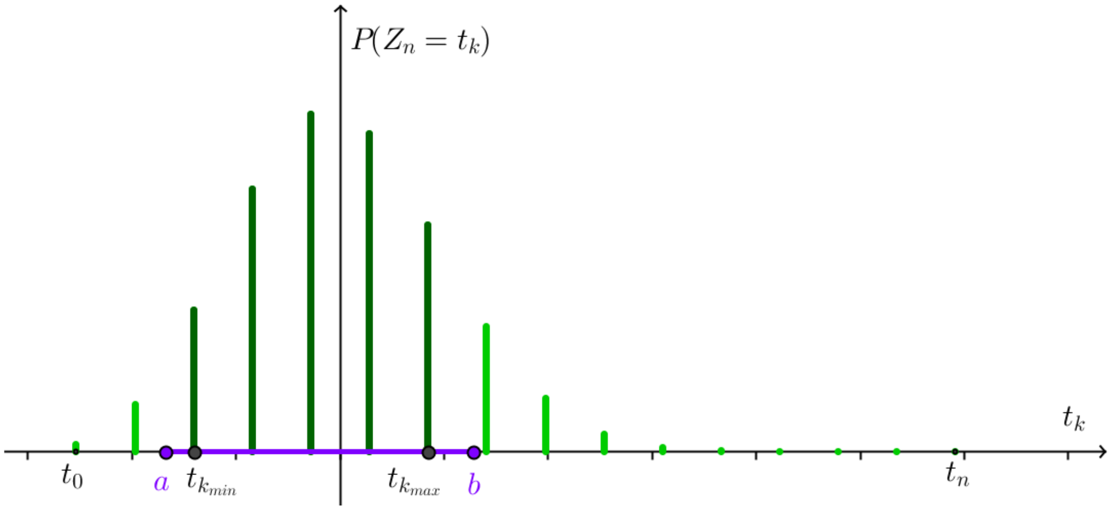
$$t_{k+1}=t_k+\frac{1}{\sqrt{(np(1-p))}}$$

$$t_{k+1}=t_k+\frac{\frac{1}{\sqrt{(p(1-p))}}}{\sqrt{n}}$$

- **Intervalle** $[a;b]$

Pour tout $(a;b) \in \mathbb{R}^2$ tel que $a<b$, on considère :

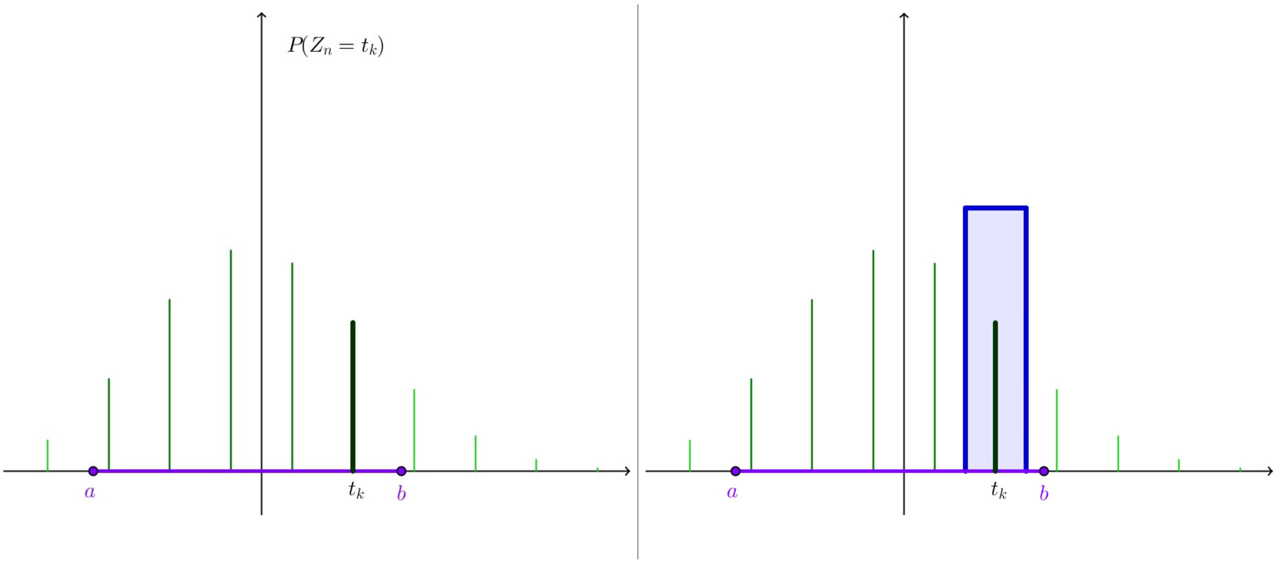
$$k_{min}=min\{k\,;t_k\in[a;b]\} \text{ et } k_{max}=max\{k\,;t_k\in[a;b]\}$$



On a donc :

$$\mathrm{P}(a\leqslant Z_n\leqslant b)=\mathrm{P}(t_{k_{min}}\leqslant Z_n\leqslant t_{k_{max}})=\sum_{k=k_{min}}^{k_{max}}\mathrm{P}(Z_n=t_k)$$

- **Du discret au continu**



On passe du discret au continu en considérant :

$$P(Z_n=t_k)=P\left(t_k-\frac{1}{2\sqrt{np(1-p)}}\leqslant Z_n\leqslant t_k+\frac{1}{2\sqrt{np(1-p)}}\right)$$

On s'intéresse alors à la probabilité sur un intervalle (et non plus en un point).

Il faut ainsi passer d'une barre à un rectangle dont on connaît la largeur : $\frac{1}{\sqrt{np(1-p)}}$.

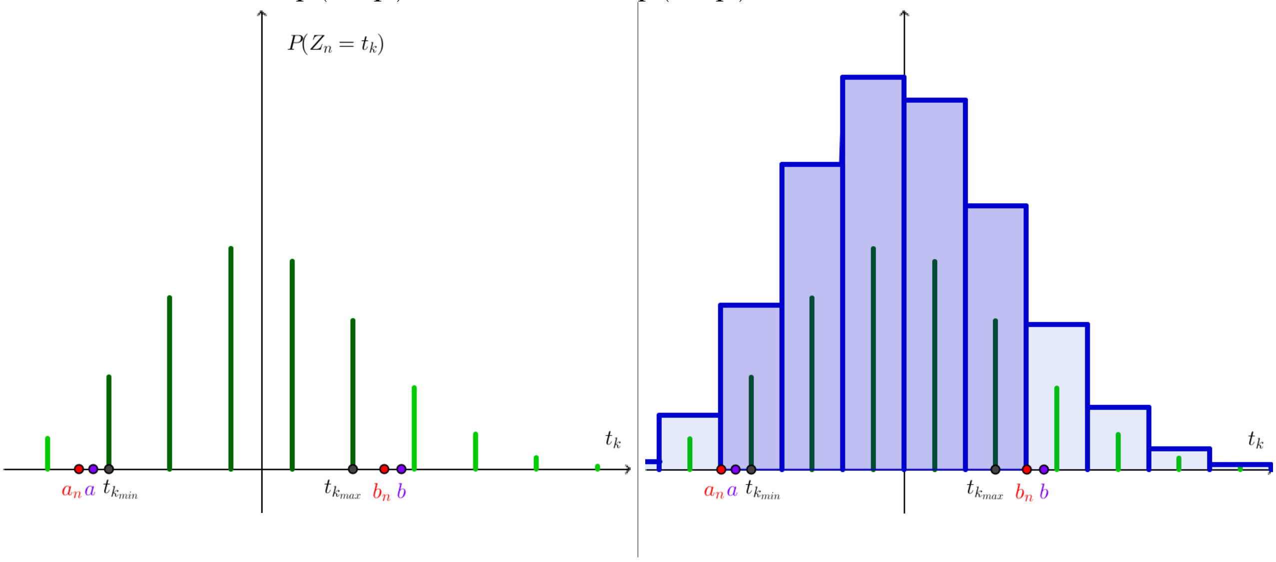
Ainsi pour que l'aire du rectangle soit égale à $P(Z_n=t_k)$ on doit avoir :

$$\mathcal{A}_k=P(Z_n=t_k)=\frac{1}{\sqrt{np(1-p)}}\times\sqrt{np(1-p)}P(Z_n=t_k).$$

C'est ainsi que la hauteur des rectangles est $\sqrt{np(1-p)}P(Z_n=t_k)$.

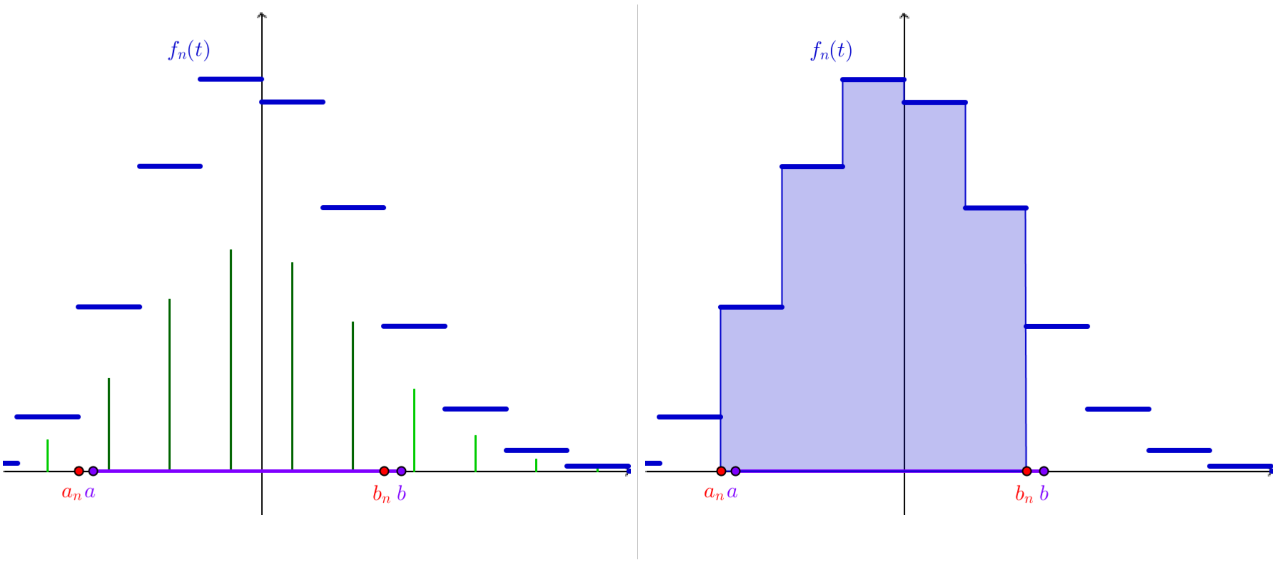
L'axe des ordonnées ne donne alors plus de probabilités.

On pose $a_n = t_{k_{min}} - \dfrac{1}{2\sqrt{np(1-p)}}$ et $b_n = t_{k_{max}} + \dfrac{1}{2\sqrt{np(1-p)}}$



On définit ainsi :

$$f_n(t) = \begin{cases} \sqrt{np(1-p)} \times \mathrm{P}(\mathrm{Z}_n = t_k) \text{ sur } J_k = \left[t_k - \dfrac{1}{2\sqrt{np(1-p)}} ; t_k + \dfrac{1}{2\sqrt{np(1-p)}} \right[\text{ pour tout } k \text{ tel que } 0 \leqslant k \leqslant n \\ 0 \text{ sinon} \end{cases}$$

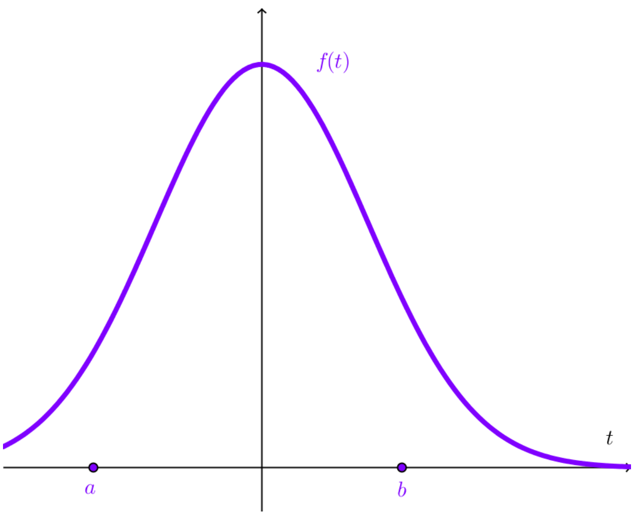


On a donc $\mathrm{P}(\mathrm{Z}_n = t_k) = \displaystyle\int_{t_k - \frac{1}{2\sqrt{np(1-p)}}}^{t_k + \frac{1}{2\sqrt{np(1-p)}}} f_n(t)\,\mathrm{d}t$

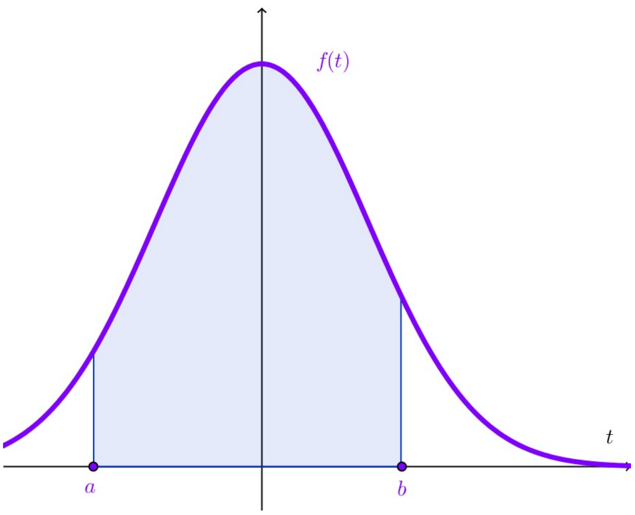
Soit $\mathrm{P}(a \leqslant \mathrm{Z}_n \leqslant b) = \displaystyle\sum_{k=k_{min}}^{k_{max}} \mathrm{P}(\mathrm{Z}_n = t_k) = \sum_{k=k_{min}}^{k_{max}} \int_{t_k - \frac{1}{2\sqrt{np(1-p)}}}^{t_k + \frac{1}{2\sqrt{np(1-p)}}} f_n(t)\,\mathrm{d}t = \int_{a_n}^{b_n} f_n(t)\,\mathrm{d}t$

Intégrale de Gauss

On introduit la fonction $f(t)=\frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}}$.



$$f(t)=\frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}}$$

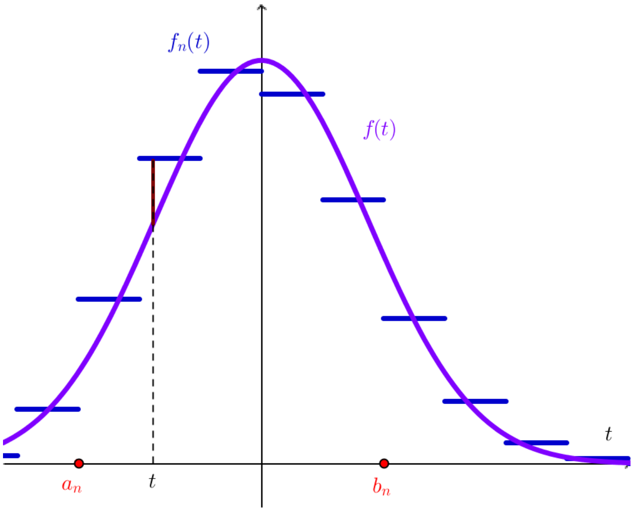


$$\int_a^b \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,dt$$

Convergence

On s'intéresse à $\left| P(a \leqslant Z_n \leqslant b) - \int_a^b \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} dt \right|$.

- Convergence simple de f_n vers f.

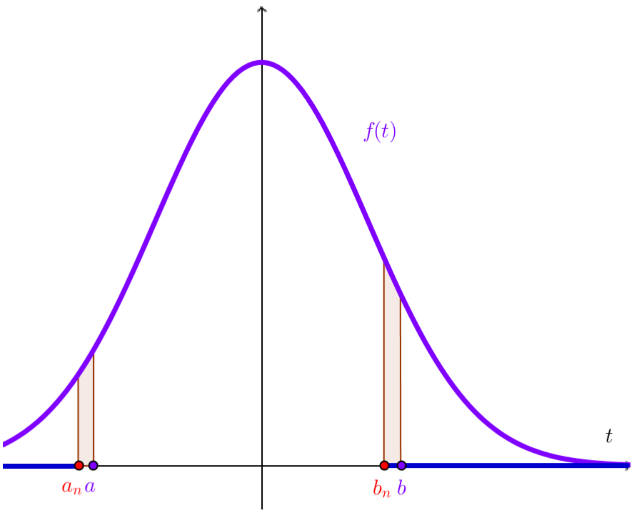


Pour tout $t \in [a_n ; b_n]$, $\left| f_n(t) - \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} \right| \leqslant \frac{C}{\sqrt{n}}$

On obtient alors $\left| \int_{a_n}^{b_n} f_n(t) dt - \int_{a_n}^{b_n} \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} dt \right| \leqslant \int_{a_n}^{b_n} \left| f_n(t) - \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} \right| dt \leqslant \int_{a_n}^{b_n} \frac{C}{\sqrt{n}} dt$

soit $\left| \int_{a_n}^{b_n} f_n(t) dt - \int_{a_n}^{b_n} \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} dt \right| \leqslant \frac{K_1}{\sqrt{n}}$

- Convergence des bornes d'intégrations



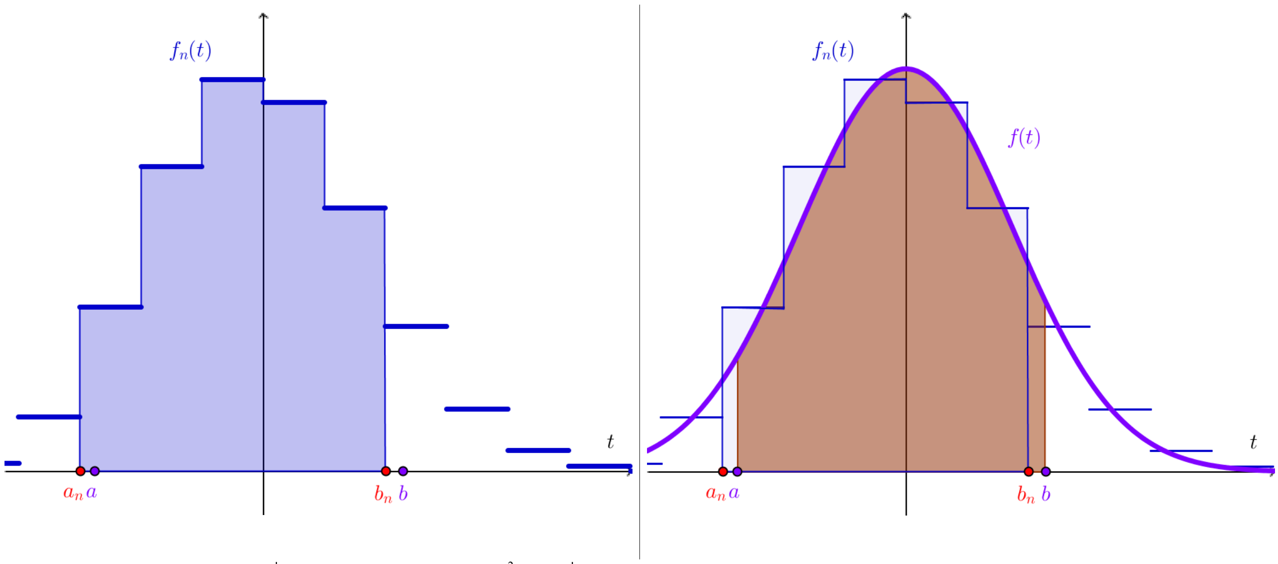
$\lim_{n \to +\infty} a_n = a$ et $\lim_{n \to +\infty} b_n = b$

$\left| \int_{a_n}^{b_n} \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} dt - \int_a^b \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} dt \right| \leqslant \frac{K_2}{\sqrt{n}}$

- Conclusion

$$\left|\mathrm{P}(a\leqslant Z_n\leqslant b)-\int_a^b \frac{\mathrm{e}^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t\right|=\left|\int_{a_n}^{b_n} f_n(t)\,\mathrm{d}t-\int_a^b \frac{\mathrm{e}^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t\right|\leqslant\left|\int_{a_n}^{b_n} f_n(t)\,\mathrm{d}t-\int_{a_n}^{b_n} \frac{\mathrm{e}^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t\right|+\left|\int_{a_n}^{b_n} \frac{\mathrm{e}^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t-\int_a^b \frac{\mathrm{e}^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t\right|$$

Soit $\left|\mathrm{P}(a\leqslant Z_n\leqslant b)-\int_a^b \frac{\mathrm{e}^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t\right|\leqslant\frac{K_1+K_2}{\sqrt{n}}$.



Par conséquent, $\lim_{n\to+\infty}\left|\int_{a_n}^{b_n} f_n(t)\,\mathrm{d}t-\int_a^b \frac{\mathrm{e}^{\frac{-t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t\right|=0$.