Carl Friedrich Gauss (1777 - 1855)

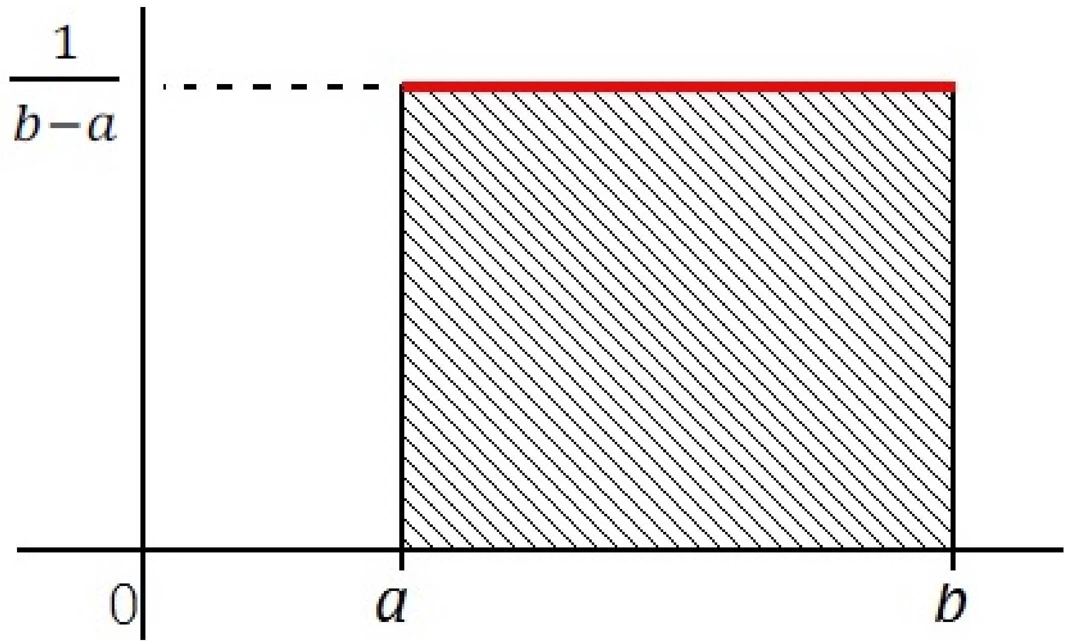
Mathématicien, astronome et physicien allemand. Surnommé « le prince des mathématiciens », il est considéré comme l'un des plus grands mathématiciens de tous les temps. Son nom reste associé aujourd'hui à la célèbre courbe en cloche représentant la densité de probabilité d'une variable aléatoire normale réduite.

I. Loi uniforme sur $[a, b]$

Définition et propriété

Soit a et b deux réels tels que $a < b$

Une variable aléatoire X suit la loi uniforme sur l'intervalle $[\boldsymbol{a}, \boldsymbol{b}]$ lorsque sa densité de probabilité est la fonction constante égale à $\frac{1}{b-a}$ sur $[a, b]$.



Propriété :

Soit X une variable aléatoire qui suit la loi uniforme sur l'intervalle $[a, b]$, on note f sa densité

X admet une espérance

$$\boldsymbol{E(X) = \int_a^b t \times f(t)dt = \frac{a+b}{2}}$$

Exemple 1.

Thomas a dit qu'il passerait voir Anita à un moment quelconque entre 18h30 et 21h00. Quelle est la probabilité qu'il arrive pendant son feuilleton préféré qui dure de 19h à 19h30 ?

II. Loi exponentielle

Définition :

Soit $\lambda > 0$ un réel

Une variable aléatoire suit la loi exponentielle de paramètre $\boldsymbol{\lambda}$ sur $[0;+\infty[$ lorsque sa densité de probabilité est la fonction

$$\forall \boldsymbol{x} \in [\mathbf{0};+\infty[,\ \ \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\lambda}\, \boldsymbol{e}^{-\boldsymbol{\lambda x}}.$$

Propriété **de durée de vie sans vieillissement :**

Soit T une variable aléatoire suivant une loi exponentielle, pour tous réels t et h positifs, $\boldsymbol{P_{T\geq t}(T \geq t+h) = P(T \geq h)}$.

Démonstration (ROC) :

Interprétation :
La probabilité que le phénomène dure au moins $t+h$ heures sachant qu'il a déjà duré t heures sera la même que la probabilité de durer h heures à partir de sa mise en fonction initiale.
En d'autres termes, le fait que le phénomène ait duré pendant $\boldsymbol{t}$ heures ne change rien à son espérance de vie à partir du temps $\boldsymbol{t}$.
(Par exemple, sachant qu'un système a atteint $t=50$ années, la probabilité qu'il dépasse $t+h=80$ ans est égale à la probabilité d'atteindre $h=30$ ans).

Cette loi permet de modéliser la décroissance radioactive ou le temps écoulé entre deux coups de téléphone reçus, ou le temps écoulé entre deux accidents de voiture dans lequel un individu donné est impliqué …

Propriété :

Soit T une variable aléatoire qui suit une loi exponentielle de paramètre λ sur $[0; +\infty[$, on note f sa densité

T admet une espérance

$$\boldsymbol{E(T) = \lim_{x \to +\infty} \int_0^x t \times f(t)dt = \frac{1}{\lambda}}$$

Démonstration (ROC) :

Exemple 2.

La durée de vie, en heures, d'un composant électronique est modélisée par la loi exponentielle de moyenne 20. Quelle est la probabilité que l'un des composants pris au hasard, soit encore en état de marche au bout de 50 h ?

III. Loi normale centrée réduite $\mathcal{N}(0;1)$

Théorème de Moivre-Laplace (admis) :

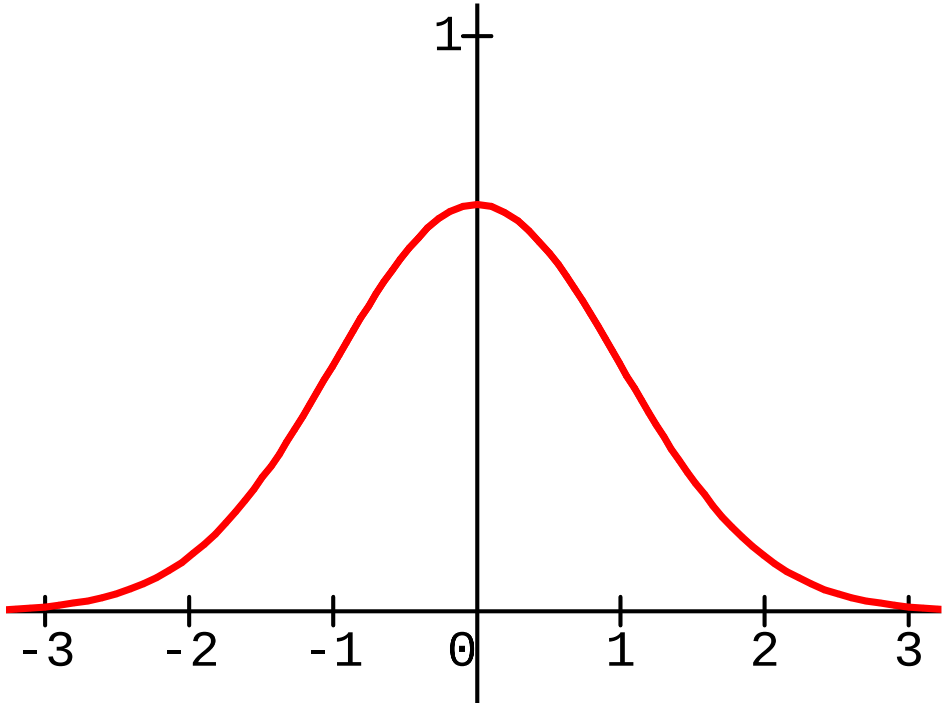
Pour tout entier naturel n, X_n est une variable aléatoire qui suit une loi binomiale $\mathcal{B}(n;p)$.
La variable centrée réduite $Z_n = \frac{X_n - np}{\sqrt{np(1-p)}}$ vérifie pour tous réels a et b :

$$\lim_{n\to+\infty} P(Z_n \in [a,b]) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Définition :
Une variable aléatoire suit la loi normale centrée réduite notée $\mathcal{N}(0;1)$ lorsque pour tout réels a et b tel que $a < b$, on a

$$\boldsymbol{P(a \leq X \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx}$$

La fonction définie sur $\mathbb{R}$ par
$\boldsymbol{f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}$ est appelée densité de la loi $\mathcal{N}(0;1)$.
On dit que la courbe représentant f est une courbe « en cloche ».



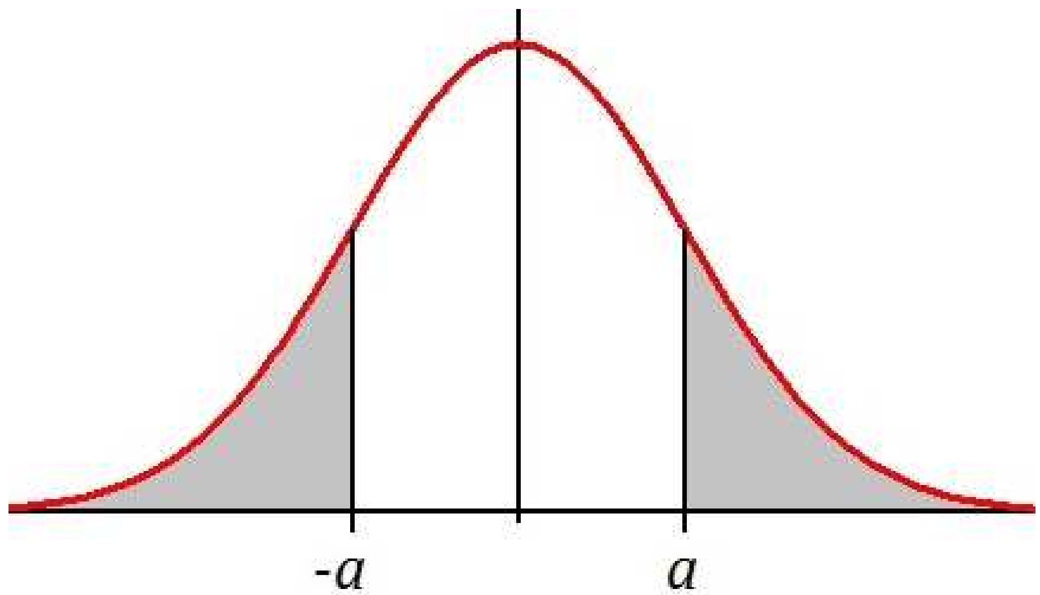
Propriétés :

Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

❶La fonction f est **continue** et **paire** sur $\mathbb{R}$. La courbe est donc symétrique par rapport à l'axe des ordonnées.

❷ Soit T une variable aléatoire qui suit la loi normale centrée réduite, pour tout $a \in \mathbb{R}$, donc

$$P(T \leq -a) = P(T \geq a)$$
$$= 1 - P(T \leq a)$$



Propriété :

Soit T une variable aléatoire qui suit la loi normale $\mathcal{N}(0; 1)$,

pour tout $\alpha \in]0; 1[$, il existe un unique réel positif $u_\alpha > 0$ tel que

$$\boldsymbol{P(-u_\alpha \leq T \leq u_\alpha) = 1 - \alpha}$$

En particulier $\boldsymbol{u_{0,05} \approx 1,96}$ et $\boldsymbol{u_{0,01} \approx 2,58}$

Démonstration (ROC) :

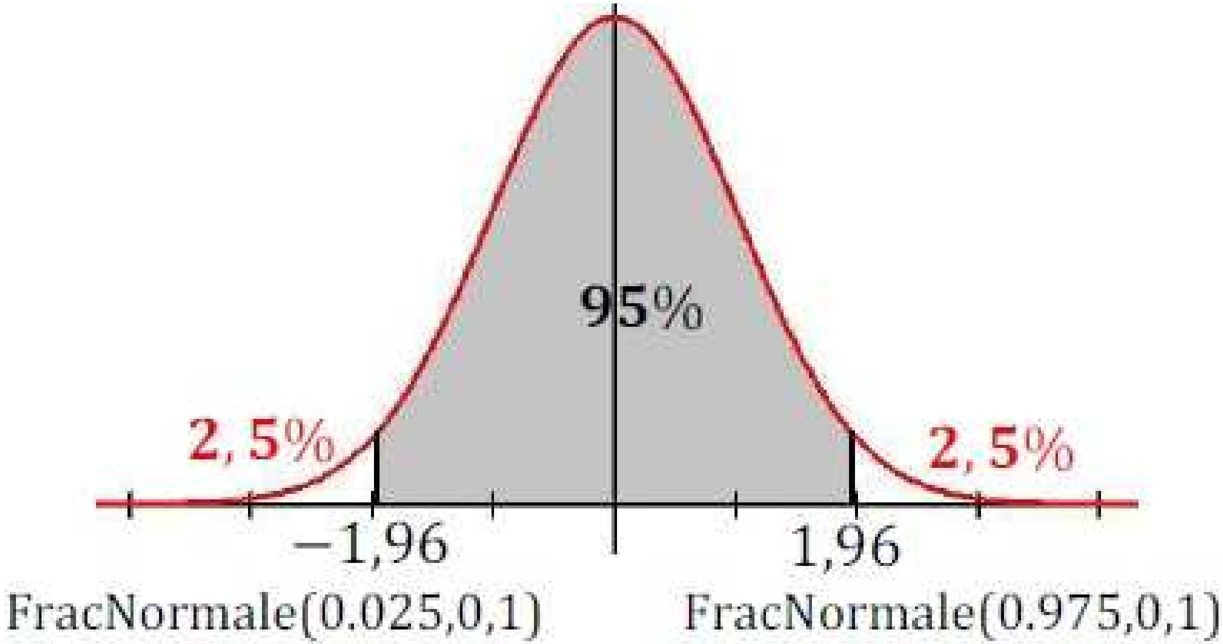
Remarque :

$u_{0,05}$ est le réel tel que :

$$P\left(-u_{0,05} \leq T \leq u_{0,05}\right) = 1 - 0{,}05 = 0{,}95$$

$$\boldsymbol{P(-1,96 \leq T \leq 1,96) = 0,95}$$

Cela donne une idée de la répartition des valeurs de T : environ 95% des réalisations de T se trouvent entre $-1{,}96$ et $1{,}96$.

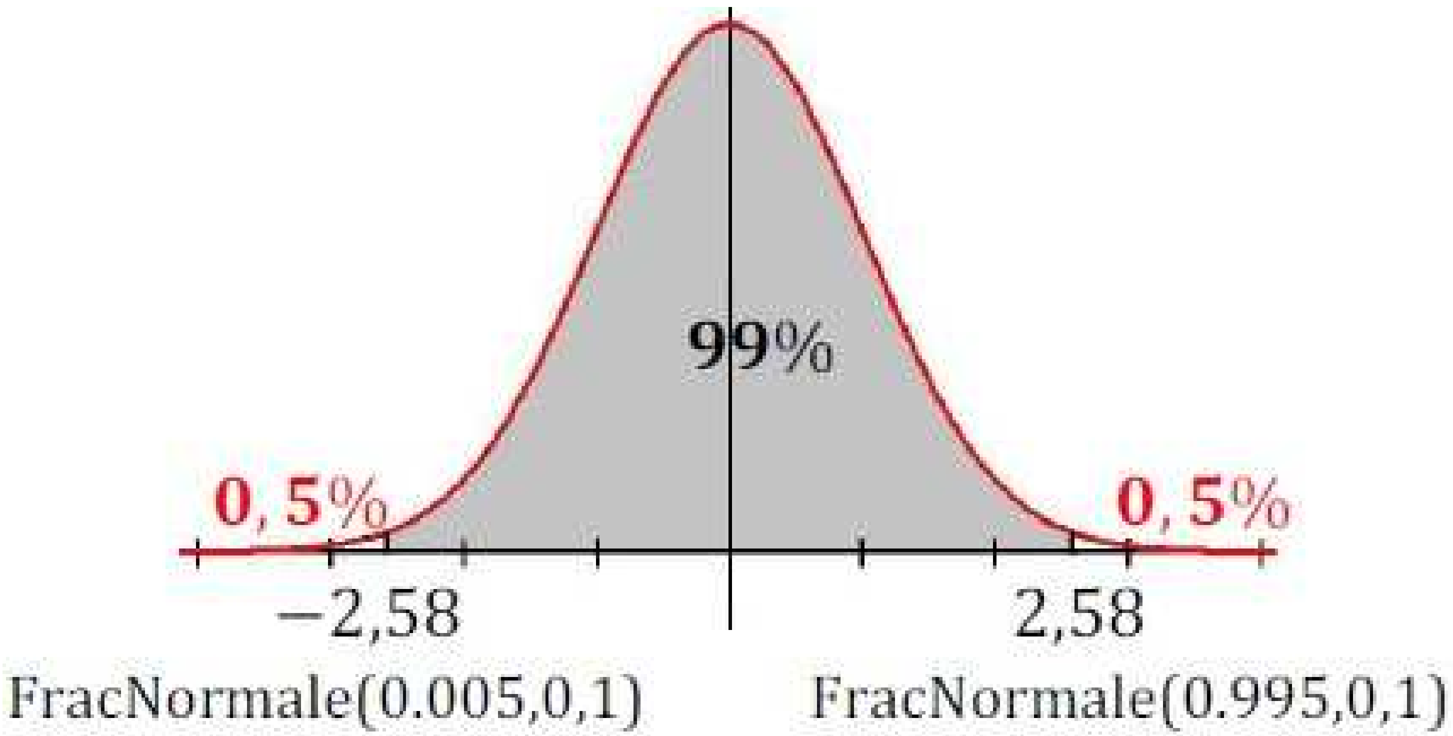


$u_{0,01}$ est le réel tel que :

$$P(-u_{0,01} \leq T \leq u_{0,01}) = 1 - 0{,}01 = 0{,}99$$

$$\boldsymbol{P(-2,58 \leq T \leq 2,58) = 0,99}$$

Environ 99% des réalisations de T se trouvent entre $-2{,}58$ et $2{,}58$.



Propriété :
Soit T une variable aléatoire qui suit la loi normale $\mathcal{N}(0;1)$, on note f sa densité
T admet une espérance

$$\boldsymbol{E(T)=\lim_{x\to-\infty}\int_x^0 t\times f(t)dt+\lim_{x\to+\infty}\int_0^x t\times f(t)dt=0}$$

Propriété (admise):
Soit T une variable aléatoire qui suit la loi normale $\mathcal{N}(0;1)$, on note f sa densité
T admet une variance et un écart type

$$\boldsymbol{V(T)=E\left[\left(T-E(T)\right)^2\right]=1} \text{ et } \boldsymbol{\sigma(T)=\sqrt{V(T)}=1}$$

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

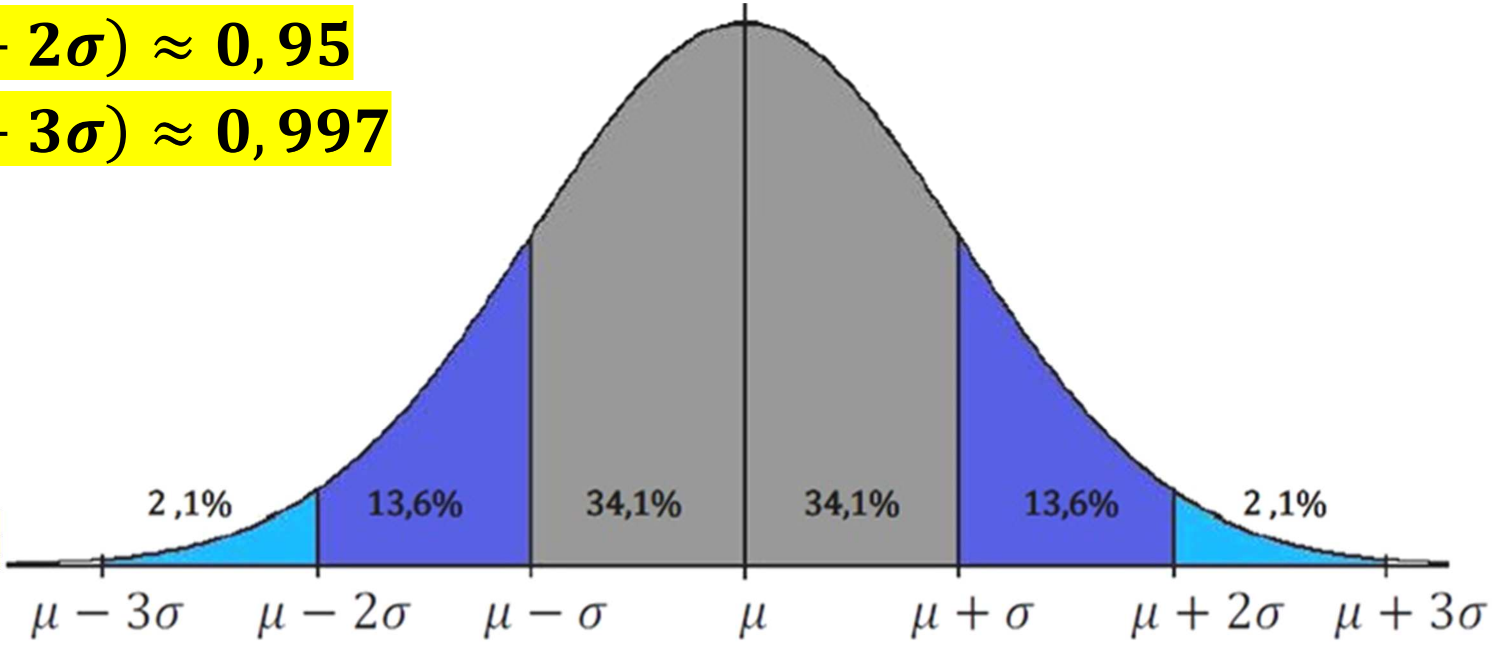
Définition : Une variable aléatoire X suit la loi normale $\mathcal{N}(\mu; \sigma^2)$ lorsque $\frac{X-\mu}{\sigma}$ suit la loi normale centrée réduite $\mathcal{N}(0; 1)$.

Propriété : Si $X \hookrightarrow \mathcal{N}(\mu; \sigma^2)$ alors $E(T) = \mu$ et $V(T) = \sigma^2$.

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = P\left(-1 \leq \frac{X-\mu}{\sigma} \leq 1\right) \approx 0,68$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0,95$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0,997$$



gaelle.buffet@ac-montpellie

V. Intervalle de fluctuation

Définition :

Soit $X \hookrightarrow \mathcal{B}(n;p)$, $\alpha \in]0;1[$, $a \in \mathbb{R}$ et $b \in \mathbb{R}$

L'intervalle $[a,b]$ est un **intervalle de fluctuation** de X **au seuil de $\boldsymbol{1-\alpha}$** signifie que $\boldsymbol{P(a \leq X \leq b) \geq 1-\alpha}$.

Propriété : $p \in]0;1[$

Si X_n suit la loi $\mathcal{B}(n;p)$, alors pour tout α dans $]0;1[$, on a

$$\lim_{n \to +\infty} P\left(\frac{X_n}{n} \in I_n\right) = 1 - \alpha$$

$$\text{où } \boldsymbol{I_n = \left[p - u_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}, p + u_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]}$$

et u_α désigne le nombre réel tel que $P(-u_\alpha \leq T \leq u_\alpha) = 1 - \alpha$ lorsque T suit la loi $\mathcal{N}(0;1)$.

Remarque : Dès que $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$, on peut approcher $P\left(\frac{X_n}{n} \in I_n\right)$ par $1 - \alpha$.

Démonstration (ROC)

Définition : $p \in]0;1[$

X_n suit la loi $\mathcal{B}(n;p)$ et $\alpha \in]0;1[$, l'intervalle I_n ci-dessus est un **intervalle de fluctuation asymptotique au seuil $1-\alpha$** de la variable aléatoire fréquence $F_n = \frac{X_n}{n}$ qui à tout échantillon de taille n associe la fréquence obtenue f.

En particulier : l'**intervalle de fluctuation asymptotique au seuil de 95%** est

$$\left[\boldsymbol{p - 1,96\frac{\sqrt{p(1-p)}}{\sqrt{n}}, p + 1,96\frac{\sqrt{p(1-p)}}{\sqrt{n}}}\right]$$

où p désigne la proportion dans la population.

Exemple 3. D'après les lois génétiques de Mendel, certains croisements de différentes variétés de pois devraient donner des pois jaunes et verts dans une proportion égale à 3 pour 1. Lors d'une expérience, on a obtenu un échantillon présentant 176 pois jaunes et 48 pois verts. ***Ces résultats sont-ils cohérents avec la théorie de Mendel ?***

Exemple 4. Un grossiste a acheté 50000 clés USB à un fabriquant qui lui a certifié que 60% avait une capacité de 4 Go et 40% une capacité de 2 Go. Un technicien prélève au hasard 50 clés USB parmi lesquelles 23 ont une capacité de 4 Go. ***Le technicien doit-il alerter son patron ?***

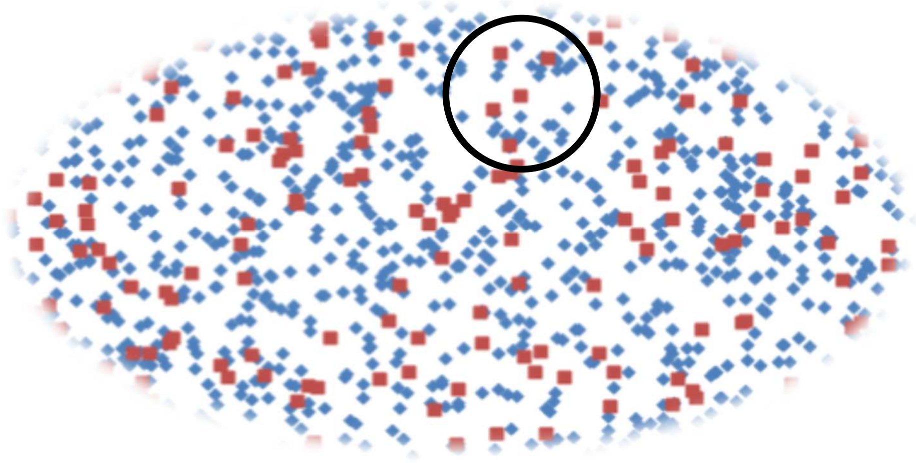
Remarque :

L'intervalle de fluctuation asymptotique au seuil de 95% vérifie

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

où $\left[p-\frac{1}{\sqrt{n}} ; p+\frac{1}{\sqrt{n}}\right]$ est l'intervalle de fluctuation utilisé en 2^{nde}.

VI. Estimation



On considère une expérience aléatoire à deux issues possibles dont on ne connait pas la probabilité p d'un succès. On cherche à estimer p à partir d'un échantillon composé de n réalisations indépendantes.

En notant X_n le nombre de succès parmi n, une estimation de p est donnée par la fréquence $F_n = \frac{X_n}{n}$.

La fréquence F_n variant d'un échantillon à l'autre, l'estimation sera donnée par un intervalle selon un niveau de confiance.

Propriété :

Pour une valeur de p fixée, l'intervalle

$$\left[\boldsymbol{F_n - \frac{1}{\sqrt{n}}, F_n + \frac{1}{\sqrt{n}}}\right]$$

contient, pour n assez grand, la proportion p avec une probabilité au moins égale à 0,95.

Conséquence : Si f est la fréquence observée sur un échantillon de taille n tel que $n \geq 30$, $nf \geq 5$ et $n(1-f) \geq 5$ alors p est élément de l'intervalle $\left[f - \frac{1}{\sqrt{n}}, f + \frac{1}{\sqrt{n}}\right]$ avec un niveau de confiance de plus de 95 %.

Démonstration (ROC) :

Exemple 5.
Lors d'une épreuve, on corrige un échantillon de copies afin de finaliser le barème pour que 80% des notes soient supérieures à 10.
Sur un échantillon de 45 copies, 25 ont plus de 10.
Pourquoi le jury décide-t-il de modifier le barème ?

Avec le nouveau barème, sur un échantillon de 36 copies, 25 ont plus de 10. Ce barème est-il plus acceptable ?