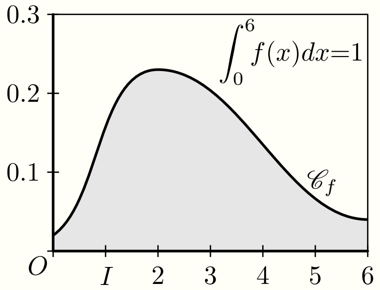
Probabilité et loi continue

A. Loi continue:

Définition :

Soit I un intervalle de $\mathbb{R}$. On appelle **densité de probabilité sur** I toute fonction f définie sur I vérifiant les trois conditions suivantes :

- f continue sur I ;
- f est positive sur I *(pour tout réel de I, $f(x)\geqslant 0$)* ;
- l'aire siutée sous sa courbe $\mathscr{C}$ est égale à une unité d'aire : $\displaystyle\int_I f(x)\,\mathrm{d}x = 1$



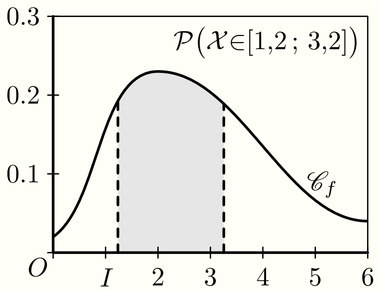
Définition :

On dit qu'une variable aléatoire $\mathcal{X}$ à valeurs dans I **suit une la loi de probabilité $\mathcal{P}$ de densité f sur** I lorsque pour tout entier a et b appartenant à I *($a<b$)* :

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

On définit l'espérance de cette variable aléatoire par :

$$E(\mathcal{X}) = \int_I t\cdot f(t)\,\mathrm{d}t$$



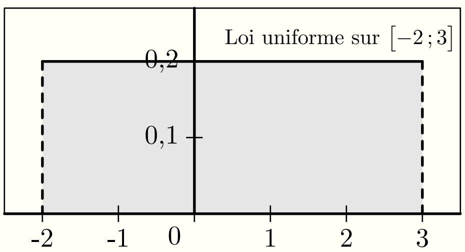
B. Loi uniforme:

Définition :

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

On dit qu'une variable aléatoire $\mathcal{X}$ suit une **loi uniforme sur l'intervalle** $I=\big[a\,;b\big]$ si la variable $\mathcal{X}$ est continue sur I et admet pour densité la fonction constante f définie par :

$$f(x) = \frac{1}{b-a}$$



Proposition :

Soit $\mathcal{X}$ une variable aléatoire suivant une loi de probabilité uniforme sur l'intervalle $I=\big[a\,;b\big]$. Pour α et β deux réels de de l'intervalle I tels que $\alpha<\beta$, on a :

$$\mathcal{P}\big(\alpha\leqslant\mathcal{X}\leqslant\beta\big) = \mathcal{P}\big(\mathcal{X}\in[\alpha\,;\,\beta]\big) = \int_\alpha^\beta \frac{1}{b-a}\,\mathrm{d}x$$

$$= \Big[\frac{x}{b-a}\Big]_\alpha^\beta = \frac{\beta-\alpha}{b-a}$$

L'espérance de la variable aléatoire $\mathcal{X}$ a pour valeur :

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

Exemple :

Dans une population d'étude, on suppose que le poids des individus est équitablement réparti entre $55\,kg$ et $95\,kg$. Déterminons la probabilité de choisir un individu dont le poid est entre $60\,kg$ et $70\,kg$.

On modélise le poid de chaque individu par une variable aléatoire suivant une loi uniforme sur l'intervalle $\big[60\,;95\big]$.

La probabilité recherchée a pour valeur :

$$\mathcal{P}\big(60\leqslant\mathcal{X}\leqslant 70\big) = \int_{60}^{70}\frac{1}{95-55}\,\mathrm{d}x = \frac{70-60}{95-55} = \frac{10}{40} = 0{,}25$$

Preuve :

Soit $\mathcal{X}$ une variable aléatoire $\mathcal{X}$ suivant une loi uniforme sur l'intervalle $\big[a\,;b\big]$. Ainsi, la variable $\mathcal{X}$ admet pour densité la fonction f constante définie par :

$$f(x) = \frac{1}{b-a}$$

- La fonction f admet pour primitive la fonction F dont l'expression est donnée par :
  $$F(x) = \frac{1}{b-a}\cdot x$$
  Ainsi, on a le calcul de probabilité :
  $$\mathcal{P}\big(\alpha\leqslant\mathcal{X}\leqslant b\big) = \int_\alpha^\beta f(t)\,\mathrm{d}t = \big[F(t)\big]_\alpha^\beta$$
  $$= \frac{1}{b-a}\cdot\beta - \frac{1}{b-a}\cdot\alpha = \frac{\beta-\alpha}{b-a}$$
- L'espérance de la variable aléatoire est définie par le calcul intégral :
  $$E(\mathcal{X}) = \int_a^b t\cdot f(t)\,\mathrm{d}t$$
  Notons g la fonction définie par :
  $$g(t) = t\cdot\frac{1}{b-a} = \frac{t}{b-a}$$
  Cette fonction g admet pour primitive la fonction G dont l'expression est :
  $$G(t) = \frac{t^2}{2(\beta-\alpha)}$$
  Ainsi, on a le calcul de l'espérance :
  $$E(\mathcal{X}) = \int_a^b t\cdot f(t)\,\mathrm{d}t = \int_a^b g(t)\,\mathrm{d}t = \big[g(t)\big]_a^b$$
  $$= \frac{b^2}{2\cdot(b-a)} - \frac{a^2}{2\cdot(b-a)} = \frac{(b+a)(b-a)}{2\cdot(b-a)} = \frac{b+a}{2}$$

C. Loi exponentielle:

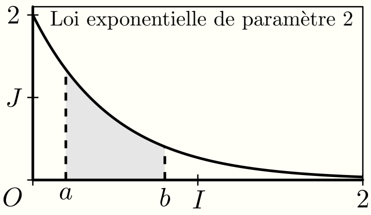
Définition :

Soit λ un réel positif. On dit qu'une variable aléatoire suit une **loi exponentielle de paramètre** λ, si la variable $\mathcal{X}$ prend ses valeurs dans $\mathbb{R}_+$ et si elle admet pour densité la

CC BY-NC

fonction f définie par :

$$f(x)=\lambda\cdot e^{-\lambda\cdot x}$$



Proposition :

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

Soit $\mathcal{X}$ une variable aléatoire suivant une loi exponentielle de paramètre λ. On a :

$$\mathcal{P}\big(a\leqslant\mathcal{X}\leqslant b\big)=\mathcal{P}\big(\mathcal{X}\in[a\,;b]\big)$$

$$=\int_a^b\lambda\cdot e^{-\lambda\cdot x}\,\mathrm{d}x=\big[-e^{-\lambda\cdot x}\big]_a^b$$

L'espérance de cette variable aléatoire a pour valeur :

$$E(\mathcal{X})=\lim_{x\mapsto+\infty}\int t\cdot\lambda e^{\lambda t}\,\mathrm{d}t=\frac{1}{\lambda}$$

Preuve :

Soit $\mathcal{X}$ une variable aléatoire suivant une loi exponentielle de paramètre λ. Ainsi, la variable aléatoire $\mathcal{X}$ admet une fonction f de densité définie par :

$$f(x)=\lambda\cdot e^{-\lambda}$$

- La fonction f admet pour primitive la fonction F définie par :

  $$F(x)=-e^{-\lambda}$$

  Ainsi, on a le calcul de probabilité suivant :

  $$\mathcal{P}\big(a\leqslant\mathcal{X}\leqslant b\big)=\int_a^b f(t)\,\mathrm{d}t=\big[F(t)\big]_a^b=\big[-e^{\lambda\cdot t}\big]_a^b$$

  $$=-e^{\lambda\cdot b}-\big[-e^{-\lambda\cdot a}\big]=e^{-\lambda\cdot a}-e^{-\lambda\cdot b}$$

- L'espérance de la variable aléatoire $\mathcal{X}$ est définie par la limite du calcul intégral suivant :

  $$E(\mathcal{X})=\lim_{x\mapsto+\infty}\int_0^x t\cdot f(t)\,\mathrm{d}t$$

  Notons g la fonction définie sur $\mathbb{R}^+$ par :

  $$g(t)=t\cdot f(t)=\lambda\cdot t\cdot e^{-\lambda\cdot t}$$

  Montrons que la fonction G définie ci-dessous est une primitive de g :

  $$G(x)=-\Big(x+\frac{1}{\lambda}\Big)\cdot e^{-\lambda t}$$

  La fonction G est définie par le produit des fonctions u et v définies par :

  $$u(x)=-\Big(x+\frac{1}{\lambda}\Big)\quad;\quad v(x)=e^{-\lambda\cdot x}$$

  qui admettent pour dérivée :

  $$u(x)=-1\quad;\quad v(x)=-\lambda\cdot e^{-\lambda x}$$

  La formule de dérivation d'un produit permet d'obtenir l'expression de la fonction G' :

$$G'(x)=u'(x)\cdot v(x)+u(x)\cdot v'(x)$$

$$=-1\times e^{-\lambda\cdot x}+\Big[-\Big(x+\frac{1}{\lambda}\Big)\Big]\cdot(-\lambda\cdot e^{-\lambda\cdot x})$$

$$=-e^{-\lambda\cdot x}+\Big(x+\frac{1}{\lambda}\Big)\cdot\lambda\cdot e^{-\lambda\cdot x}=-e^{-\lambda\cdot x}+x\cdot\lambda\cdot e^{-\lambda\cdot x}+\frac{1}{\lambda}\cdot\lambda\cdot e^{-\lambda\cdot x}$$

$$=-e^{-\lambda\cdot x}+x\cdot\lambda\cdot e^{-\lambda\cdot x}+e^{-\lambda\cdot x}=x\cdot\lambda\cdot e^{-\lambda\cdot x}=g(t)$$

Ainsi, pour $x\in\mathbb{R}_+$, on a le calcul intégral :

$$\int_0^x t\cdot f(t)\,\mathrm{d}t=\int_0^x g(t)\,\mathrm{d}t=\big[G(x)\big]_0^x=\Big[-\Big(t+\frac{1}{\lambda}\Big)\cdot e^{-\lambda\cdot t}\Big]_0^x$$

$$=-\Big(x+\frac{1}{\lambda}\Big)e^{-\lambda\cdot x}-\Big[-\Big(0+\frac{1}{\lambda}\Big)\cdot e^{-\lambda\cdot 0}\Big]=-\Big(x+\frac{1}{\lambda}\Big)e^{-\lambda\cdot x}+\frac{1}{\lambda}$$

On montre facilement la limite suivante :

$$\lim_{x\mapsto+\infty}-\Big(x+\frac{1}{\lambda}\Big)e^{-\lambda\cdot x}=0$$

On en déduit la valeur de l'espérance de la variable aléatoire $\mathcal{X}$:

$$E(\mathcal{X})=\lim_{x\mapsto+\infty}\int_0^x t\cdot f(t)\,\mathrm{d}t=\frac{1}{\lambda}$$

Proposition : *(durée de vie sans vieillissement)*

Soit $\mathcal{X}$ une variable aléatoire suivant une loi exponentielle de paramètre λ.

Pour tout réels t et h positifs, on a l'égalité :

$$\mathcal{P}_{(\mathcal{X}\geqslant t)}\big(\mathcal{X}\geqslant t+h\big)=\mathcal{P}\big(\mathcal{X}\geqslant h\big)$$

Preuve :

Considérons t et h deux nombres positifs quelconques.
D'après la définition d'une probabilité conditionnelle, on a :

$$\mathcal{P}_{(\mathcal{X}\geqslant t)}\big(\mathcal{X}\geqslant t+h\big)=\frac{\mathcal{P}\Big(\big\{\mathcal{X}\geqslant t+h\big\}\cap\big\{\mathcal{X}\geqslant t\big\}\Big)}{\mathcal{P}\big(\mathcal{X}\geqslant t}$$

Comme $\big\{\mathcal{X}\geqslant t+h\big\}\subset\big\{\mathcal{X}\geqslant t\big\}$, on a en déduit : $\big\{\mathcal{X}\geqslant t+h\big\}\cap\big\{\mathcal{X}\geqslant t\big\}=\big\{\mathcal{X}\geqslant t+h\big\}$.

Ainsi, on a l'égalité :

$$\mathcal{P}_{(\mathcal{X}\geqslant t)}\big(\mathcal{X}\geqslant t+h\big)=\frac{\mathcal{P}\big(\mathcal{X}\geqslant t+h\big)}{\mathcal{P}\big(\mathcal{X}\geqslant t\big)}=\frac{1-\mathcal{P}\big(\mathcal{X}<t+h\big)}{1-\mathcal{P}\big(\mathcal{X}<t\big)}$$

Etant sur une variable aléatoire continue, on a :

$$=\frac{1-\mathcal{P}\big(\mathcal{X}\leqslant t+h\big)}{1-\mathcal{P}\big(\mathcal{X}\leqslant t\big)}=\frac{1-\big[-e^{-\lambda\cdot x}\big]_0^{t+h}}{1-\big[-e^{-\lambda\cdot x}\big]_0^{h}}$$

$$=\frac{1-\big(-e^{-\lambda\cdot(t+h)}+1\big)}{1-\big(-e^{-\lambda\cdot t}+1\big)}=\frac{e^{-\lambda\cdot(t+h)}}{e^{-\lambda\cdot t}}=e^{-\lambda\cdot(t+h)-(-\lambda\cdot t)}$$

$$=e^{-\lambda\cdot(t+h)+\lambda\cdot t}=e^{-\lambda\cdot h}$$

Remarque :

Une loi exponentielle permet de modéliser des phénomènes dont la durée de vie est un phénomène sans mémoire : la probabilité $\mathcal{P}_{(\mathcal{X}\geqslant t)}\big(\mathcal{X}\geqslant t+h\big)$ que le phénomène dure au moins $t+h$ heures sachant qu'il a déjà duré t heures est égale à la probabilité $\mathcal{P}\big(\mathcal{X}\geqslant h\big)$ que le phénomène dure au moins h heures.
On peut dire que l'espérance de vie d'un tel phénomène ne dépend pas de l'âge de l'objet.

Cette loi de probabilité est utilisé notamment pour modéliser la durée de vie d'un élément radioactif.

D. Loi normale:

1. Centrée et réduite :

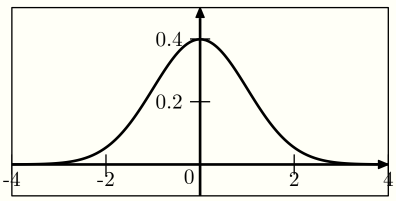
1. *Introduction :*

Définition :

On dit qu'une variable aléatoire $\mathcal{X}$ à valeurs dans $\mathbb{R}$ suit une **loi normale centrée réduite**, si la variable aléatoire $\mathcal{X}$ admet pour densité la fonction f :

CC BY-NC

$$f(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{\frac{-x^2}{2}}.\qquad \text{On note } \mathcal{X}\sim\mathcal{N}(0\,;1)$$



Remarque :

On admet que : $\displaystyle\lim_{x\mapsto+\infty}\int_{-x}^{x}\frac{1}{\sqrt{2\pi}}\cdot e^{\frac{-t^2}{2}}\,\mathrm{d}t=1.$

Ainsi, la fonction f vérifie les trois conditions pour être une loi de probabilité.

Proposition :

Soit $\mathcal{X}$ suivant une loi de probabilité $\mathcal{N}(0\,;1)$.

L'espérance de la variable aléatoire $\mathcal{X}$ a pour valeur :

$$E(\mathcal{X})=\lim_{x\mapsto+\infty}\int_{-x}^{x}t\cdot f(t)\,\mathrm{d}t$$
$$=\lim_{x\mapsto-\infty}\int_{x}^{0}t\cdot f(t)\,\mathrm{d}t+\lim_{x\mapsto+\infty}\int_{0}^{x}t\cdot f(t)\,\mathrm{d}t=0$$

Preuve :

Notons $\mathcal{X}$ une variable aléatoire suivant la loi normale centrée réduite. En notant f la densité de la variable $\mathcal{X}$, on a :

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

Pour déterminer une expression de ces intégrales, considérons la fonction g définie par :

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

Considérons la fonction u définie par :

$$u(x)=-\frac{x^2}{2}\quad;\quad u'(x)=-x$$

Ainsi, la fonction g admet pour expression :

$$g(x)=-\frac{1}{\sqrt{2\pi}}\times u'(x)\cdot e^{u(x)}$$

qui admet la fonction G pour primitive :

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

Pour déterminer l'espérance $E(\mathcal{X})$, nous nous intéressons aux deux intégrales :

- $\displaystyle\int_0^x t\cdot f(t)\,\mathrm{d}t=\int_0^x g(t)\,\mathrm{d}t=\big[G(x)\big]_0^x$
  $\displaystyle=-\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}+\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{0^2}{2}}=\frac{1}{\sqrt{2\pi}}-\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}$
- $\displaystyle\int_x^0 t\cdot f(t)\,\mathrm{d}t=\big[G(x)\big]_x^0=-\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{0^2}{2}}+\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}$
  $\displaystyle=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}-\frac{1}{\sqrt{2\pi}}$

On obtient les deux limites suivantes :

- $\displaystyle\lim_{x\mapsto-\infty}\int_x^0 t\cdot f(t)\,\mathrm{d}t=\lim_{x\mapsto-\infty}\frac{1}{\sqrt{2\pi}}-\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}=\frac{1}{\sqrt{2\pi}}$
- $\displaystyle\lim_{x\mapsto+\infty}\int_0^x t\cdot f(t)\,\mathrm{d}t=\lim_{x\mapsto+\infty}\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}-\frac{1}{\sqrt{2\pi}}=-\frac{1}{\sqrt{2\pi}}$

On en déduit la limite :

$$E(x)=\lim_{x\mapsto-\infty}\int_x^0 t\cdot f(t)\,\mathrm{d}t+\lim_{x\mapsto+\infty}\int_0^x t\cdot f(t)\,\mathrm{d}t$$
$$=\frac{1}{\sqrt{2\pi}}+\Big(-\frac{1}{\sqrt{2\pi}}\Big)=0$$

Proposition : *(admis)*

On considère une variable aléatoire $\mathcal{X}$ suivant une loi de normale centrée réduite $(\mathcal{X}\sim\mathcal{N}(0\,;1))$. La variance de la variable aléatoire $\mathcal{X}$ a pour valeur :

$$V(\mathcal{X})=E\big[\mathcal{X}-E(\mathcal{X})^2\big]=1$$

2. *Intervalle centré et probabilité :*

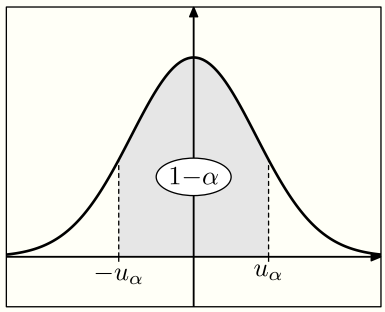
Proposition :

Soit $\mathcal{X}$ une variable aléatoire suivant une loi de probabilité $\mathcal{N}(0\,;1)$. Pour tout nombre réel $\alpha\in\,]0\,;1[$, il existe un unique réel positif u_α tel que :

$$\mathcal{P}\big(-u_\alpha\leqslant\mathcal{X}\leqslant u_\alpha\big)=1-\alpha$$

Remarque :

Il n'existe pas de formule algébrique donnant u_α directement en fonction de α. Des tables de correspondances existent ou l'usage de la calculatrice est nécessaire.



Pour une valeur de α appartenant à l'intervalle $]0\,;1[$ et fixé, la proposition précédente d'un nombre positif u_α tel que la surface grisée ci-dessous a une aire de $1-\alpha$.

Preuve :

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

Soit α un nombre réel appartenant à l'intervalle $]0\,;1[$. Nous allons montrer qu'il existe qu'une valeur de $x\in\mathbb{R}_+$ réalisant l'égalité :

$$\mathcal{P}\big(-x\leqslant\mathcal{X}\leqslant x\big)=1-\alpha$$

On note G la fonction définie sur $\mathbb{R}_+$ par :

$$G(x)=\mathcal{P}\big(-x\leqslant\mathcal{X}\leqslant x\big)$$

l'équation devient : $G(x)=1-\alpha$.

Notons F une primitive de la densité f. La fonction G admet pour expression :

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

La fonction G admet pour dérivée :

$$G'(x)=F'(x)-\big[-1\times F'(-x)\big]=F'(x)+F'(-x)$$
$$=f(x)+f(-x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}+\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(-x)^2}{2}}$$
$$=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}+\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}=\frac{2}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}$$

G' est positive sur $\mathbb{R}_+$: G est croissante sur $\mathbb{R}_+$.

On a les deux valeurs particulières de la fonction G :

BY-NC

- $G(0)=\mathcal{P}(0\leqslant\mathcal{X}\leqslant 0)=\mathcal{P}(\mathcal{X}=0)=0$
car $\mathcal{X}$ suit une loi continue.
- $\lim\limits_{x\mapsto+\infty} G(x)=\lim\limits_{x\mapsto+\infty}\mathcal{P}(-x\leqslant\mathcal{X}\leqslant x)=\mathcal{P}(\mathcal{X}\in\mathbb{R})=1$
qui est la probabilité de l'univers.

On obtient le tableau de variation de la fonction G :

x	0	$+\infty$
Variation de G	0 ↗	1

Le nombre α, représentant une probabilité, appartient à l'intervalle $]0\,;1[$: on a aussi : $1-\alpha\in]0\,;1[$.
Ainsi, $1-\alpha$ est compris entre les limites de la fonction G aux bornes de l'intervalle $\mathbb{R}_+$.

La fonction g est continue et strictement monotone sur l'intervalle $\mathbb{R}_+$. D'après le corollaire du théorème des valeurs intermédiaires, il existe un unique nombre x_0 réalisation l'égalité :

$$G(x_0)=1-\alpha$$
$$\mathcal{P}(-x_0\leqslant\mathcal{X}\leqslant x_0)=1-\alpha$$

Le nombre u_α recherché est ce nombre x_0.

3. *Loi normale :*

Définition :

Soit $\mu\in\mathbb{R}$ et $\sigma>0$. On dit qu'une variable aléatoire $\mathcal{X}$ suit une loi normale $\mathcal{N}(\mu\,;\sigma^2)$ si $\dfrac{\mathcal{X}-\mu}{\sigma}$ suit une loi normale $\mathcal{N}(0\,;1)$.

Proposition :

Soit $\mathcal{X}$ une variable aléatoire suivant une loi normale de paramètre μ et σ^2 *($\mathcal{X}\sim\mathcal{N}(\mu\,;\sigma^2)$)*. On a :

$$E(\mathcal{X})=\mu \quad ; \quad V(\mathcal{X})=\sigma^2$$

Ainsi, la variable aléatoire $\mathcal{X}$ a un écart-type égal à σ.

Démonstration :

En classe de première, on a établit les égalités suivantes pour une variable aléatoire $\mathcal{Y}$ discrète :

$$E(a\cdot\mathcal{Y}+b)=a\cdot E(\mathcal{Y})+b \quad ; \quad V(a\cdot\mathcal{Y}+b)=a^2\cdot\mathcal{Y}$$

où a et b sont deux réels

On admet que ces égalités restent vraies pour une variable aléatoire $\mathcal{Y}$ continue.

Condidérons $\mathcal{X}$ une variable aléatoire suivant une loi normale de paramètre μ et σ^2 *($\mathcal{X}\sim\mathcal{N}(\mu\,;\sigma^2)$)*. En notant $\mathcal{Z}$ une variable aléatoire normale centrée réduite *($\mathcal{Z}\sim\mathcal{N}(0\,;1)$)*, on a la relation :

$$\mathcal{Z}=\frac{\mathcal{X}-\mu}{\sigma}$$
$$\sigma\cdot\mathcal{Z}=\mathcal{X}-\mu$$
$$\mathcal{X}=\sigma\cdot\mathcal{Z}+\mu$$

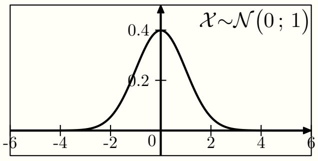
Etablissons les deux égalités recherchées :

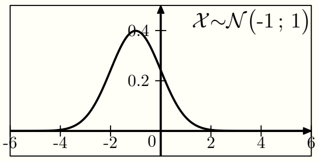
- D'après la propriété citée ci-dessus, on a :
$$E(\mathcal{X})=E(\sigma\cdot\mathcal{Z}+\mu)=\sigma\cdot E(\mathcal{Z})+\mu$$
$\mathcal{Z}$ suivant la loi normale centrée réduite, on a $E(\mathcal{Z})=0$:
$$=\sigma\times 0+\mu=\mu$$
- D'après la propriété citée ci-dessus, on a :
$$V(\mathcal{X})=V(\sigma\cdot\mathcal{Z}+\mu)=\sigma^2\cdot V(\mathcal{Z})$$
Or, $\mathcal{Z}$ suit la loi normale centrée réduite : $V(\mathcal{Z})=1$:
$$=\sigma^2\times 1=\sigma^2$$

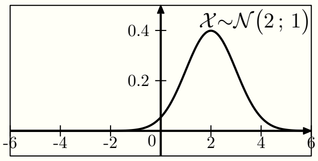
Remarque :

Voici une interprétation des deux paramètres d'une loi normale :

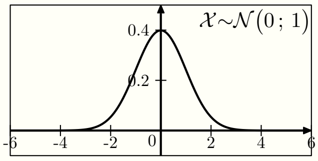
Le paramètre μ est un paramètre de position :

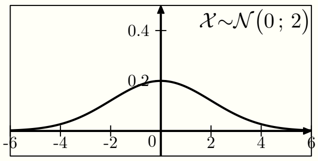


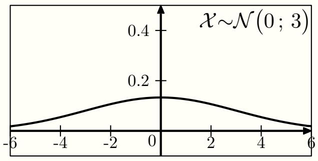




Le paramètre σ^2 est un paramètre de dispersion :



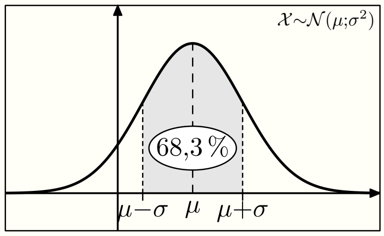




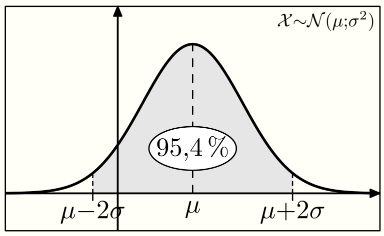
Proposition : *(Intervalle de normalité)*

Soit $\mathcal{X}$ une variable aléatoire suivant une loi normale de paramètre μ et σ^2 *($\mathcal{X}\sim\mathcal{N}(\mu\,;\sigma^2)$)*. On a les égalités suivantes :

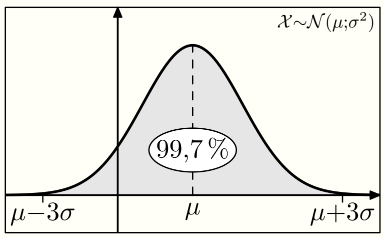
- $\mathcal{P}(\mu-\sigma\leqslant\mathcal{X}\leqslant\mu+\sigma)\simeq 0{,}683$



- $\mathcal{P}(\mu-2\sigma\leqslant\mathcal{X}\leqslant\mu+2\sigma)\simeq 0{,}954$



- $\mathcal{P}(\mu-3\sigma\leqslant\mathcal{X}\leqslant\mu+3\sigma)\simeq 0{,}997$



Démonstration :

Soit $\mathcal{X}$ une variable aléatoire suivant une loi normale de paramètre μ et σ^2 *($\mathcal{X}\sim\mathcal{N}(\mu\,;\sigma^2)$)*.

En posant $\mathcal{Z}=\dfrac{\mathcal{X}-\mu}{\sigma}$, la définition de la loi normale montre que la variable aléatoire suit une loi normale centrée réduite

 BY-NC

$(\mathcal{Z}\sim\mathcal{N}(0;1))$

Etudions chaque cas :

- $\mathcal{P}\big(\mu-\sigma\leqslant\mathcal{X}\leqslant\mu+\sigma\big) = \mathcal{P}\big(-\sigma\leqslant\mathcal{X}-\mu\leqslant\sigma\big)$

$= \mathcal{P}\Big(-1\leqslant\dfrac{\mathcal{X}-\mu}{\sigma}\leqslant1\Big) = \mathcal{P}\big(-1\leqslant\mathcal{Z}\leqslant1\big) = \displaystyle\int_{-1}^{1}\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{t^2}{2}}\,\mathrm{d}t$

A l'aide de la calculatrice, on obtient : $\simeq 0{,}683$

- $\mathcal{P}\big(\mu-2\sigma\leqslant\mathcal{X}\leqslant\mu+2\sigma\big) = \mathcal{P}\big(-2\leqslant\mathcal{Z}\leqslant2\big) \simeq 0{,}954$
- $\mathcal{P}\big(\mu-3\sigma\leqslant\mathcal{X}\leqslant\mu+3\sigma\big) = \mathcal{P}\big(-3\leqslant\mathcal{Z}\leqslant3\big) \simeq 0{,}997$

E. Théorème de Moivre-Laplace :

Théorème : *(de Moivre-Laplace - admis)*

Soit n un entier naturel non-nul et p un nombre réel dans $[0;1]$. On considère la variable aléatoire $\mathcal{X}_n$ suivant une loi binomiale de paramètre n et p $(\mathcal{X}_n\sim\mathcal{B}(n;p))$.

On note $\mathcal{Z}$ une variable aléatoire suivant une loi normale centrée réduite $(\mathcal{Z}\sim\mathcal{N}(0;1))$. On a l'égalité :

$$\lim_{n\mapsto+\infty}\mathcal{P}\Big(a\leqslant\frac{\mathcal{X}_n-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\leqslant b\Big)$$

$$= \mathcal{P}\big(a\leqslant\mathcal{Z}\leqslant b\big) = \int_a^b\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{t^2}{2}}\,\mathrm{d}t$$

Remarque :

- On rappelle qu'une variable aléatoire $\mathcal{X}$ suivant une loi binomiale de paramètre n et p $(\mathcal{X}\sim\mathcal{B}(n;p))$ admet les indicateurs de position et de dispersion suivants :

$E(\mathcal{X})=n\cdot p$; $V(\mathcal{X})=n\cdot p\cdot(1-p)$; $\sigma(\mathcal{X})=\sqrt{n\cdot p\cdot(1-p)}$

- Dans les études pratiques, on admet que si les trois conditions suivantes sont admises :

$n\geqslant 30$; $n\cdot p\geqslant 5$; $n\cdot(1-p)\geqslant 5$

Alors lesvaleurs des termes de la suite sont assez proches de la limite pour accepter la relation suivante :

$$\mathcal{P}\Big(a\leqslant\frac{\mathcal{X}_\backslash-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\leqslant b\Big)\simeq\mathcal{P}\big(a\leqslant\mathcal{Z}\leqslant b\big)$$

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

Corollaire : *(Intervalle de fluctuation asymptotique de F_n)*

Soit α un nombre réel appartenant à $]0;1[$ et $\mathcal{Z}$ une variable aléatoire suivant une loi normale centrée réduite $(\mathcal{Z}\sim\mathcal{N}(0;1))$.

On note u_α l'unique réel positif vérifiant l'égalité :

$$\mathcal{P}\big(-u_\alpha\leqslant\mathcal{Z}\leqslant u_\alpha\big) = 1-\alpha$$

Si $\mathcal{X}_n$ suit une loi binomiale $\mathcal{B}(n;p)$ alors la fréquence de succès $F_n=\dfrac{\mathcal{X}_n}{n}$ vérifie :

$$\lim_{n\mapsto+\infty}\mathcal{P}\Big(p-u_\alpha\frac{\sqrt{p(1-p)}}{\sqrt{n}}\leqslant F_n\leqslant p+u_\alpha\frac{\sqrt{p(1-p)}}{\sqrt{n}}\Big)=1-\alpha$$

Démonstration :

On considère une variable aléatoire $\mathcal{X}_n$ suivant une loi binomiale de paramètre n et p $(\mathcal{X}_n\sim\mathcal{B}(n;p))$.

Soit a et b deux nombres réels tels que $a<b$. On a les transformations suivantes :

$$a\leqslant\frac{\mathcal{X}_n-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\leqslant b$$

$$a\cdot\sqrt{n\cdot p\cdot(1-p)}\leqslant\mathcal{X}_n-n\cdot p\leqslant b\cdot\sqrt{n\cdot p\cdot(1-p)}$$

$$a\cdot\sqrt{n\cdot p\cdot(1-p)}+n\cdot p\leqslant\mathcal{X}_n\leqslant b\cdot\sqrt{n\cdot p\cdot(1-p)}+n\cdot p$$

$$\frac{a\cdot\sqrt{n\cdot p\cdot(1-p)}+n\cdot p}{n}\leqslant\frac{\mathcal{X}_n}{n}\leqslant\frac{b\cdot\sqrt{n\cdot p\cdot(1-p)}+n\cdot p}{n}$$

$$a\cdot\frac{\sqrt{n\cdot p\cdot(1-p)}}{n}+p\leqslant F_n\leqslant b\cdot\frac{\sqrt{n\cdot p\cdot(1-p)}}{n}+p$$

$$p+a\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\leqslant F_n\leqslant p+b\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}$$

Ainsi, on a l'égalité des probabilités :

$$\mathcal{P}\Big(p+a\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\leqslant F_n\leqslant p+b\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\Big)$$

$$=\mathcal{P}\Big(a\leqslant\frac{\mathcal{X}_n-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\leqslant b\Big)$$

Ces deux membres étant égaux pour tout n, leurs limites sont égales :

$$\lim_{n\mapsto+\infty}\mathcal{P}\Big(p+a\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\leqslant F_n\leqslant p+b\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\Big)$$

$$=\lim_{n\mapsto+\infty}\mathcal{P}\Big(a\leqslant\frac{\mathcal{X}_n-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\leqslant b\Big)$$

D'après le théorème de Moivre-Laplace, on a

$$=\lim_{n\mapsto+\infty}\mathcal{P}\Big(p+a\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\leqslant F_n\leqslant p+b\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\Big)$$

$$=\mathcal{P}\big(a\leqslant\mathcal{Z}\leqslant b\big)\qquad\text{où } \mathcal{Z}\sim\mathcal{N}(0;1)$$

Or, pour tout réel $\alpha\in]0;1[$, il existe un unique réel u_α positif réalisant l'égalité : $\mathcal{P}\big(-u_\alpha\leqslant\mathcal{Z}\leqslant u_\alpha\big)=1-\alpha$

On en déduit :

$$\lim_{n\mapsto+\infty}\mathcal{P}\Big(p-u_\alpha\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\leqslant F_n\leqslant p+u_\alpha\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\Big)$$

$$=\mathcal{P}\big(-u_\alpha\leqslant\mathcal{Z}\leqslant u_\alpha\big)=1-\alpha$$

Remarque :

- La construction de cet intervalle permet d'affirmer que la probabilité qu'une issue de l'expérience aléatoire appartient à cet intervalle est d'**au moins** 0,95.
Lors d'arrondie, pour conserver cette propriété, on applique la règle suivante : on prend la valeur par défaut de la borne inférieure et la valeur par excès de la borne supérieure. Voir exemple suivant.
- En supposant que : $n\geqslant30$; $n\cdot p\geqslant5$; $n\cdot(1-p)\geqslant5$
  - Pour obtenir un intervalle de fluctuation à 95 %, il faut choisir : $\alpha=0{,}05$; $u_\alpha=1{,}96$:
  La fréquence des succés F_n appartient à l'intervalle $\Big[p-1{,}96\cdot\dfrac{\sqrt{p(1-p)}}{\sqrt{n}}\,;\,p+1{,}96\dfrac{\sqrt{p(1-p)}}{\sqrt{n}}\Big]$ avec une probabilité de 95 %.
  Cet intervalle est appelé **intervalle de fluctuation asymptotique** à 95 % de F_n
  - Pour obtenir un intervalle de fluctuation à 99 %, il faut choisir : $\alpha=0{,}01$; $u_\alpha=2{,}58$
  La fréquence des succés F_n appartient à l'intervalle $\Big[p-2{,}58\cdot\dfrac{\sqrt{p(1-p)}}{\sqrt{n}}\,;\,p+2{,}58\dfrac{\sqrt{p(1-p)}}{\sqrt{n}}\Big]$ avec une probabilité de 99 %.
  Cet intervalle est appelé **intervalle de fluctuation asymptotique** à 99 % de F_n

Exemple :

(cc) BY-NC

Les conditions d'utilisation du théorème de Moivre-Laplace sont vérifiées :

$n=1000\geqslant 30$; $n\cdot p=150\geqslant 5$; $n\cdot(1-p)=850\geqslant 5$

Ainsi, l'intervalle de fluctuation asymptotique de la fréquence au seuil de 95 % s'écrit :

$$I=\left[p-1{,}96\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\ ;\ p-1{,}96\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\right]$$

$$I=\left[0{,}15-1{,}96\cdot\frac{\sqrt{0{,}15\times 0{,}85}}{\sqrt{1000}}\ ;\ 0{,}15-1{,}96\cdot\frac{\sqrt{0{,}15\times 0{,}85}}{\sqrt{1000}}\right]$$

$$I\approx\big[0{,}127868\ ;\ 0{,}17213\big]$$

On utilise la valeur par défaut de la borne inférieure et la valeur par excès de la borne supèrieure :

$$I\approx\big[0{,}127\ ;\ 0{,}173\big]$$

Corollaire : *(Intervalle de fluctuation de F_n simplifié)*

Considérons une variable aléatoire $\mathcal{X}_n$ suivant une loi binomiale de paramètre n et p *($\mathcal{X}_n\sim\mathcal{B}(n\,;p)$)*.

Pour tout réel $p\in\left]0\,;1\right[$, il existe un entier n_0 tel que pour tout $n\geqslant n_0$:

$$\mathcal{P}\Big(p-\frac{1}{\sqrt{n}}\leqslant F_n\leqslant p+\frac{1}{\sqrt{n}}\Big)>0{,}95$$

Démonstration :

La démonstration de ce corollaire s'effectue en deux étapes :

- Montrons que :

$$\left[p-2\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\ ;\ p+2\cdot\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\right]\subset\left[p-\frac{1}{\sqrt{n}}\ ;\ p+\frac{1}{\sqrt{n}}\right]:$$

Considérons la fonction g définie sur $\big[0\,;1\big]$ par la relation : $g(x)=x\cdot(1-x)$

On montre facilement que $g'(x)=1-2x$ et on obtient le tableau de variation :

x	0		$\frac{1}{2}$		1
Variation de g	0	$\nearrow$	$\frac{1}{4}$	$\searrow$	1

Ainsi, pour toute valeur de p *(où $p\in\big[0\,;1\big]$)* :

$$g(p)\leqslant\frac{1}{4}$$

$$p(1-p)\leqslant\frac{1}{4}$$

La fonction racine carrée est croissante sur $\mathbb{R}_+$:

$$\sqrt{p(1-p)}\leqslant\sqrt{\frac{1}{4}}$$

$$\sqrt{p(1-p)}\leqslant\frac{1}{2}$$

$$2\cdot\sqrt{p(1-p)}\leqslant 2\times\frac{1}{2}$$

$$2\cdot\sqrt{p(1-p)}\leqslant 1$$

$$2\cdot\frac{\sqrt{p(1-p)}}{\sqrt{n}}\leqslant\frac{1}{\sqrt{n}}$$

On en déduit les comparaisons suivantes :

$$-\frac{1}{\sqrt{n}}\leqslant -2\cdot\frac{\sqrt{p(1-p)}}{\sqrt{n}}\leqslant 2\cdot\frac{\sqrt{p(1-p)}}{\sqrt{n}}\leqslant\frac{1}{\sqrt{n}}$$

Ajoutons p à chaque membre de l'inégalité :

$$p-\frac{1}{\sqrt{n}}\leqslant p-2\cdot\frac{\sqrt{p(1-p)}}{\sqrt{n}}\leqslant p+2\cdot\frac{\sqrt{p(1-p)}}{\sqrt{n}}\leqslant p+\frac{1}{\sqrt{n}}$$

On vient d'établir l'inclusion :

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

- Soit α le nombre réalisant l'égalité ci-dessous où $\mathcal{Z}$ est une variable aléatoire suivant une loi normale réduite et centrée *($\mathcal{Z}\sim\mathcal{N}(0\,;1)$)* :

$\mathcal{P}\big(-2\leqslant\mathcal{Z}\leqslant 2\big)=1-\alpha$

A l'aide de la calculatrice, on obtient :

$\mathcal{P}\big(-2\leqslant\mathcal{Z}\leqslant 2\big)\simeq 0{,}954$ à 10^{-3} près.

En appliquant le corollaire du théorème de Moivre-Laplace, on a la valeur de la limite :

$$\lim_{n\mapsto+\infty}\mathcal{P}\Big(p-2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\leqslant\mathcal{F}_n\leqslant p+2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\Big)$$

$$=\mathcal{P}\big(-2\leqslant\mathcal{Z}\leqslant 2\big)=1-\alpha\simeq 0{,}954$$

Or, la convergence de cette limite implique l'existence d'un entier n_0 afin que tous les termes de rang supérieur à n_0 appartient à l'intervalle $\big[0{,}951\,;0{,}957\big]$ centré en 0,954.

$$n\geqslant n_0\implies\mathcal{P}\Big(p-2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\leqslant F_n\leqslant p+2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\Big)>0{,}95$$

De l'inclusion des intervalles :

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

On en déduit la comparaison des probabilités :

$$\mathcal{P}\Big(F_n\in\Big[p-\frac{1}{\sqrt{n}}\ ;\ p+\frac{1}{\sqrt{n}}\Big]\Big)$$

$$\geqslant\mathcal{P}\Big(F_n\in\Big[p-2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\ ;p+2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\Big]\Big)>0{,}95$$

F. Estimation de p:

Définition :

Pour tout nombre réel α appartenant à l'intervalle $\left]0\,;1\right[$, on appelle **intervalle de confiance de la proportion p au niveau de confiance** $1-\alpha$ est un intervalle, obtenu à partir d'un échantillon de la population, contenant le nombre p avec une probabilité supérieure à $1-\alpha$

Proposition :

Pour n est assez grand, l'intervalle $I=\Big[F_n-\frac{1}{\sqrt{n}}\,;F_n+\frac{1}{\sqrt{n}}\Big]$ est un intervalle de confiance au niveau de 95 % pour l'estimation de la proportion p

Remarque :

En pratique, n assez grand signifie :

$n\geqslant 30$; $n\cdot p\geqslant 5$; $n\cdot(1-p)\geqslant 5$

Démonstration :

On considère la répétition de manière indépendante d'une épreuve de Bernoulli de paramètre p. On note F_n la fréquence des succés réalisés sur un échantillon de taille n.

D'après le second théorème de Moivre-Laplace, on a la prob-

CC BY-NC

abilité :

$$\mathcal{P}\left(F_n \in \left[p - \frac{1}{\sqrt{n}}\,;\,p + \frac{1}{\sqrt{n}}\right]\right) \geqslant 0{,}95$$

On a les égalités suivantes de probabilités :

$$\mathcal{P}\left(F_n \in \left[p - \frac{1}{\sqrt{n}}\,;\,p + \frac{1}{\sqrt{n}}\right]\right) = \mathcal{P}\left(p - \frac{1}{\sqrt{n}} \leqslant F_n \leqslant p + \frac{1}{\sqrt{n}}\right)$$

$$= \mathcal{P}\left(-\frac{1}{\sqrt{n}} \leqslant F_n - p \leqslant \frac{1}{\sqrt{n}}\right) = \mathcal{P}\left(-\frac{1}{\sqrt{n}} \leqslant p - F_n \leqslant \frac{1}{\sqrt{n}}\right)$$

$$= \mathcal{P}\left(F_n - \frac{1}{\sqrt{n}} \leqslant p \leqslant F_n + \frac{1}{\sqrt{n}}\right)$$

$$= \mathcal{P}\left(p \in \left[F_n - \frac{1}{\sqrt{n}}\,;\,F_n + \frac{1}{\sqrt{n}}\right]\right)$$

On en déduit l'inégalité :

$$\mathcal{P}\left(p \in \left[F_n - \frac{1}{\sqrt{n}}\,;\,F_n + \frac{1}{\sqrt{n}}\right]\right) \geqslant 0{,}95$$

Remarque :

On considère la variable aléatoire $\mathcal{Z}$ suivant une loi normale centrée réduite *($\mathcal{Z} \sim \mathcal{N}(0\,;1)$)*.

Pour tout $\alpha \in \left]0\,;1\right[$, il existe un nombre réel u_α positif tel que : $\mathcal{P}\big(u_\alpha \leqslant \mathcal{Z} \leqslant u_\alpha\big) = 1 - \alpha$

Pour $p \in \left]0\,;1\right[$, soit $\mathcal{X}_n$ une variable aléatoire suivant une loi binomiale de paramètre n et p *($\mathcal{X} \sim \mathcal{B}(n\,;p)$)*. On note $F_n = \dfrac{\mathcal{X}_n}{n}$ la fréquence de succés liée à l'échantillon de taille n.

Voici quelques valeurs particulières :

- Pour $\alpha = 0{,}05$, on a : $u_{0{,}05} \simeq 1{,}96$:
  - $\mathcal{P}\big(-1{,}96 \leqslant \mathcal{Z} \leqslant 1{,}96\big) = 0{,}95$
  - $\displaystyle\lim_{n \mapsto +\infty} \mathcal{P}\left(-1{,}96 \leqslant \frac{\mathcal{X}_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \leqslant 1{,}96\right) = 0{,}95$

$$\lim_{n \mapsto +\infty} \mathcal{P}\left(n \cdot p - 1{,}96\sqrt{np(1-p)} \leqslant \mathcal{X}_n \leqslant np + 1{,}96\sqrt{np(1-p)}\right) = 0{,}95$$

  - $\displaystyle\lim_{n \mapsto +\infty} \mathcal{P}\left(p - 1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}} \leqslant F_n \leqslant p + 1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right) = 0{,}95$

- Pour $\alpha = 0{,}01$, on a : $u_{0{,}01} \simeq 2{,}58$:
  - $\mathcal{P}\big(-2{,}58 \leqslant \mathcal{Z} \leqslant 2{,}58\big) = 0{,}99$
  - $\displaystyle\lim_{n \mapsto +\infty} \mathcal{P}\left(-2{,}58 \leqslant \frac{\mathcal{X}_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \leqslant 2{,}58\right) = 0{,}99$

$$\lim_{n \mapsto +\infty} \mathcal{P}\left(np - 2{,}58\sqrt{np(1-p)} \leqslant \mathcal{X}_n \leqslant np + 2{,}58\sqrt{np(1-p)}\right) = 0{,}99$$

  - $\displaystyle\lim_{n \mapsto +\infty} \mathcal{P}\left(p - 2{,}58\frac{\sqrt{p(1-p)}}{\sqrt{n}} \leqslant F_n \leqslant p + 2{,}58\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right) = 0{,}99$

https://chingatome.fr (cc) BY-NC