

Démonstration :

Soit \mathcal{X} une variable aléatoire \mathcal{X} suivant une loi uniforme sur l'intervalle $[a ; b]$. 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 :

$$\mathcal{P}(a \leq \mathcal{X} \leq b) = \int_a^b f(t) dt = [F(t)]_a^b = \frac{1}{b-a} \cdot b - \frac{1}{b-a} \cdot a = \frac{b-a}{b-a}$$

- L'espérance de la variable aléatoire est définie par le calcul intégral :

$$E(\mathcal{X}) = \int_a^b t \cdot f(t) dt$$

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 :

$$G(t) = \frac{t^2}{2(b-a)}$$

Ainsi, on a le calcul de l'espérance :

$$E(\mathcal{X}) = \int_a^b t \cdot f(t) dt = \int_a^b g(t) dt = [g(t)]_a^b = \frac{b^2}{2 \cdot (b-a)} - \frac{a^2}{2 \cdot (b-a)} = \frac{b^2 - a^2}{2 \cdot (b-a)} = \frac{(b+a)(b-a)}{2 \cdot (b-a)} = \frac{b+a}{2}$$

Démonstration :

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 :

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

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

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

Ainsi, on a le calcul de probabilité suivant :

$$\mathcal{P}(a \leq \mathcal{X} \leq b) = \int_a^b f(t) dt = [F(t)]_a^b = [-e^{-\lambda t}]_a^b = -e^{-\lambda b} - [-e^{-\lambda a}] = e^{-\lambda a} - e^{-\lambda b}$$

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

$$\lim_{x \rightarrow +\infty} \int_0^x t \cdot f(t) dt$$

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

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

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

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

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

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

qui admettent pour dérivée :

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

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

$$\begin{aligned} G'(x) &= u'(x) \cdot v(x) + u(x) \cdot v'(x) = -1 \times e^{-\lambda x} + \left[-\left(x + \frac{1}{\lambda}\right)\right] \cdot (-\lambda \cdot e^{-\lambda x}) = -e^{-\lambda x} + \left(x + \frac{1}{\lambda}\right) \cdot \lambda \cdot e^{-\lambda x} \\ &= -e^{-\lambda x} + x \cdot \lambda \cdot e^{-\lambda x} + \frac{1}{\lambda} \cdot \lambda \cdot e^{-\lambda x} = -e^{-\lambda x} + x \cdot \lambda \cdot e^{-\lambda x} + e^{-\lambda x} = x \cdot \lambda \cdot e^{-\lambda x} = g(t) \end{aligned}$$

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

$$\begin{aligned} \int_0^x t \cdot f(t) dt &= \int_0^x g(t) dt = [G(x)]_0^x = \left[-\left(t + \frac{1}{\lambda}\right) \cdot e^{-\lambda t}\right]_0^x \\ &= -\left(x + \frac{1}{\lambda}\right) e^{-\lambda x} - \left[-\left(0 + \frac{1}{\lambda}\right) \cdot e^{-\lambda \cdot 0}\right] = -\left(x + \frac{1}{\lambda}\right) e^{-\lambda x} + \frac{1}{\lambda} \end{aligned}$$

On montre facilement la limite suivante :

$$\lim_{x \rightarrow +\infty} -\left(x + \frac{1}{\lambda}\right) e^{-\lambda x} = 0$$

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

$$E(\mathcal{X}) = \lim_{x \rightarrow +\infty} \int_0^x t \cdot f(t) dt = \frac{1}{\lambda}$$

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 :

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

Démonstration :

Considérons t et h deux nombres positifs quelconques.

D'après la définition d'une probabilité conditionnelle, on a :

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

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

Ainsi, on a l'égalité :

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

Etant sur une variable aléatoire continue, on a :

$$\begin{aligned} &= \frac{1 - \mathcal{P}(\mathcal{X} \leq t+h)}{1 - \mathcal{P}(\mathcal{X} \leq t)} = \frac{1 - [-e^{-\lambda \cdot x}]_0^{t+h}}{1 - [-e^{-\lambda \cdot x}]_0^t} = \frac{1 - (-e^{-\lambda \cdot (t+h)} + 1)}{1 - (-e^{-\lambda \cdot t} + 1)} = \frac{e^{-\lambda \cdot (t+h)}}{e^{-\lambda \cdot t}} = e^{-\lambda \cdot (t+h) - (-\lambda \cdot t)} \\ &= e^{-\lambda \cdot (t+h) + \lambda \cdot t} = e^{-\lambda \cdot h} \end{aligned}$$

Remarque :

Une loi exponentielle permet de modéliser des phénomènes dont la durée de vie est un phénomène sans mémoire : la probabilité $\mathcal{P}_{(\mathcal{X} \geq t)}(\mathcal{X} \geq t+h)$ que le phénomène dure au moins $t+h$ heures sachant qu'il a déjà duré t heures est égale à la probabilité $\mathcal{P}(\mathcal{X} \geq h)$ 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.

A. Loi normale :

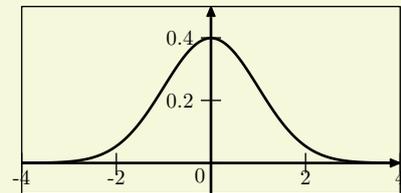
1. Introduction :

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 :

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

On note alors $\mathcal{X} \sim \mathcal{N}(0; 1)$



Remarque :

On admet que $\lim_{x \rightarrow +\infty} \int_{-x}^x \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt = 1$.

On en déduit que la fonction f de la définition précédente représente bien 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 :

$$E(\mathcal{X}) = \lim_{x \rightarrow +\infty} \int_{-x}^x t \cdot f(t) dt = \lim_{x \rightarrow -\infty} \int_x^0 t \cdot f(t) dt + \lim_{x \rightarrow +\infty} \int_0^x t \cdot f(t) dt = 0$$

Démonstration :

Notons \mathcal{X} une variable aléatoire suivant une 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 l'espérance $E(\mathcal{X})$, nous nous intéressons aux intégrales :

$$\int_x^0 t \cdot f(t) dt \quad ; \quad \int_0^x t \cdot f(t) dt$$

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 :

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

La fonction g admet pour expression :

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

Ainsi, la fonction g admet pour primitive la fonction G définie par :

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

Ainsi, on a les expressions des deux intervalles recherchées :

$$\bullet \int_0^x t \cdot f(t) dt = \int_0^x g(t) dt = [G(x)]_0^x = -\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} + \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{0^2}{2}} = \frac{1}{\sqrt{2\pi}} - \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

$$\bullet \int_x^0 t \cdot f(t) dt = [G(x)]_x^0 = -\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{0^2}{2}} + \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} - \frac{1}{\sqrt{2\pi}}$$

On obtient les deux limites suivantes :

$$\bullet \lim_{x \rightarrow -\infty} \int_x^0 t \cdot f(t) dt = \lim_{x \rightarrow -\infty} \frac{1}{\sqrt{2\pi}} - \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}}$$

$$\bullet \lim_{x \rightarrow +\infty} \int_0^x t \cdot f(t) dt = \lim_{x \rightarrow +\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} - \frac{1}{\sqrt{2\pi}} = -\frac{1}{\sqrt{2\pi}}$$

$$\text{On en déduit la limite : } E(x) = \lim_{x \rightarrow -\infty} \int_x^0 t \cdot f(t) dt + \lim_{x \rightarrow +\infty} \int_0^x t \cdot f(t) dt = \frac{1}{\sqrt{2\pi}} + \left(-\frac{1}{\sqrt{2\pi}} \right) = 0$$

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 :

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

2. Intervalle centré et probabilité :

Proposition :

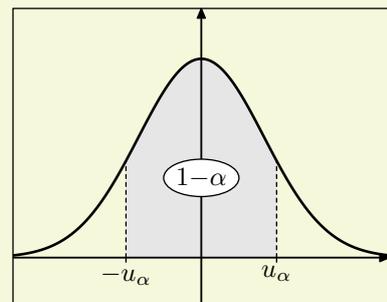
Soit \mathcal{X} une variable aléatoire suivant une loi de probabilité $\mathcal{N}(0;1)$. Pour tout nombre réel $\alpha \in]0;1[$, il existe un unique réel positif u_α tel que :

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

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 $1 - \alpha$.



Démonstration :

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 $]0; 1[$. On souhaite montrer que l'équation en x ci-dessous admet une unique solution :

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

On remarque que l'équation impose à l'inconnue x à prendre des valeurs positive.

En considérant la fonction G définie par :

$$G(x) = \mathcal{P}(-x \leq \mathcal{X} \leq x) \quad \text{sur } \mathbb{R}_+,$$

l'équation devient : $G(x) = 1 - \alpha$.

La fonction G admet pour expression :

$$G(x) = \mathcal{P}(-x \leq \mathcal{X} \leq x) = \int_{-x}^x f(t) dt$$

où f représente la densité de la loi normale centrée réduite. Notons F la primitive de la fonction f (même si on ne peut déterminer son expression, on connaît son existence) :

$$= F(x) - F(-x)$$

La fonction G admet pour dérivée :

$$G'(x) = F'(x) - [-1 \times F'(-x)] = F'(x) + F'(-x)$$

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

$$= \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} + \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} = \frac{2}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

On en déduit que la fonction G' est positive sur \mathbb{R}_+ . Ainsi, la fonction G est croissante sur \mathbb{R}_+ .

On a les deux valeurs particulières de la fonction G :

- $G(0) = \mathcal{P}(0 \leq \mathcal{X} \leq 0) = \mathcal{P}(\mathcal{X}=0) = 0$
car \mathcal{X} suit une loi continue.
- $\lim_{x \rightarrow +\infty} G(x) = \lim_{x \rightarrow +\infty} \mathcal{P}(-x \leq \mathcal{X} \leq x) = \mathcal{P}(\mathcal{X} \in \mathbb{R}) = 1$
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 appartenant à l'intervalle $]0; 1[$, on en déduit que $1 - \alpha$ appartient également à l'intervalle $]0; 1[$. Ainsi, pour toute valeur de α , le nombre $1 - \alpha$ appartient à l'image de l'intervalle \mathbb{R}_+ par la fonction G .

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

$$G(x_0) = 1 - \alpha$$

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

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 $\frac{\mathcal{X} - \mu}{\sigma}$ 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 établi les égalités suivantes pour une variable aléatoire \mathcal{Y} discrète :

$$E(a \cdot \mathcal{Y} + b) = a \cdot E(\mathcal{Y}) + b \quad ; \quad V(a \cdot \mathcal{Y} + b) = a^2 \cdot V(\mathcal{Y}) \quad \text{où } a \text{ et } b \text{ sont deux réels}$$

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

Considérons \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 :

$$V(\mathcal{X}) = V(\sigma \cdot \mathcal{Z} + \mu) = \sigma^2 \cdot V(\mathcal{Z})$$

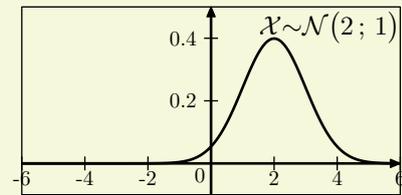
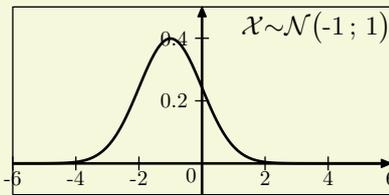
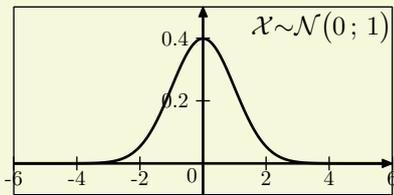
Or, la variable \mathcal{Z} suit la loi normale centrée réduite : $V(\mathcal{Z})=1$:

$$= \sigma^2 \times 1 = \sigma^2$$

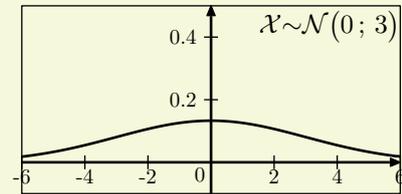
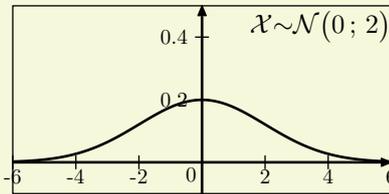
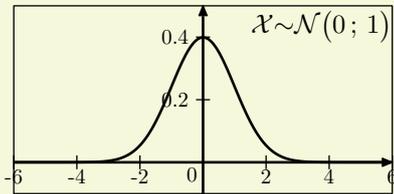
Remarque :

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

- Le paramètre μ est un paramètre de position :



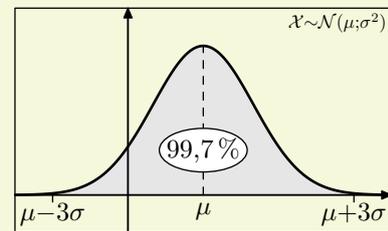
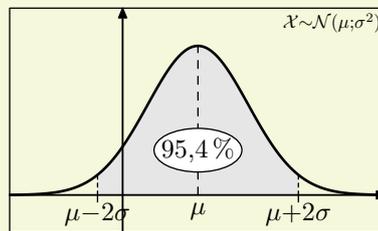
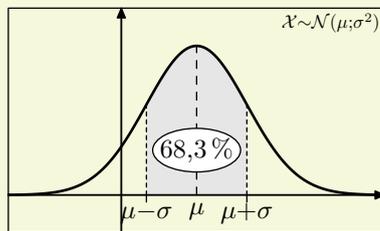
- Le paramètre σ^2 est un paramètre de dispersion :



Proposition :

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 :

- $\mathcal{P}(\mu - \sigma \leq \mathcal{X} \leq \mu + \sigma) \simeq 0,683$
- $\mathcal{P}(\mu - 2\sigma \leq \mathcal{X} \leq \mu + 2\sigma) \simeq 0,954$
- $\mathcal{P}(\mu - 3\sigma \leq \mathcal{X} \leq \mu + 3\sigma) \simeq 0,997$



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 $\mathcal{Z} = \frac{\mathcal{X} - \mu}{\sigma}$, la définition de la loi normale montre que la variable aléatoire suit une loi normale centrée réduite ($\mathcal{Z} \sim \mathcal{N}(\mu; \sigma^2)$)

Etudions chaque cas :

- $\mathcal{P}(\mu - \sigma \leq \mathcal{X} \leq \mu + \sigma) = \mathcal{P}(-\sigma \leq \mathcal{X} - \mu \leq \sigma) = \mathcal{P}(-1 \leq \frac{\mathcal{X} - \mu}{\sigma} \leq 1) = \mathcal{P}(-1 \leq \mathcal{Z} \leq 1) = \int_{-1}^1 \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt$

A l'aide de la calculatrice, on obtient :

$$\simeq 0,683$$

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

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

B. 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 :

$$\lim_{n \rightarrow +\infty} \mathcal{P}\left(a \leq \frac{\mathcal{X}_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \leq b\right) = \mathcal{P}(a \leq \mathcal{Z} \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt$$

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 :

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

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

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

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

$$\mathcal{P}\left(a \leq \frac{\mathcal{X}_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \leq b\right) \simeq \mathcal{P}(a \leq \mathcal{Z} \leq b)$$

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

Corollaire : (Intervalle de fluctuation 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 : $\mathcal{P}(-u_\alpha \leq \mathcal{Z} \leq u_\alpha) = 1 - \alpha$

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

$$\lim_{n \rightarrow +\infty} \mathcal{P}\left(p - u_\alpha \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \leq F_n \leq p + u_\alpha \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}\right) = 1 - \alpha$$

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 :

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

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

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

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

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

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

Ainsi, on a l'égalité des probabilités :

$$\mathcal{P}\left(p + a \cdot \frac{\sqrt{p \cdot (1 - p)}}{\sqrt{n}} \leq F_n \leq p + b \cdot \frac{\sqrt{p \cdot (1 - p)}}{\sqrt{n}}\right) = \mathcal{P}\left(a \leq \frac{\mathcal{X}_n - n \cdot p}{\sqrt{n \cdot p \cdot (1 - p)}} \leq b\right)$$

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

$$\lim_{n \rightarrow +\infty} \mathcal{P}\left(p + a \cdot \frac{\sqrt{p \cdot (1 - p)}}{\sqrt{n}} \leq F_n \leq p + b \cdot \frac{\sqrt{p \cdot (1 - p)}}{\sqrt{n}}\right) = \lim_{n \rightarrow +\infty} \mathcal{P}\left(a \leq \frac{\mathcal{X}_n - n \cdot p}{\sqrt{n \cdot p \cdot (1 - p)}} \leq b\right)$$

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

$$= \lim_{n \rightarrow +\infty} \mathcal{P}\left(p + a \cdot \frac{\sqrt{p \cdot (1 - p)}}{\sqrt{n}} \leq F_n \leq p + b \cdot \frac{\sqrt{p \cdot (1 - p)}}{\sqrt{n}}\right) = \mathcal{P}(a \leq \mathcal{Z} \leq b) \quad \text{où } \mathcal{Z} \sim \mathcal{N}(0; 1)$$

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

On en déduit :

$$\lim_{n \rightarrow +\infty} \mathcal{P}\left(p - u_\alpha \cdot \frac{\sqrt{p \cdot (1 - p)}}{\sqrt{n}} \leq F_n \leq p + u_\alpha \cdot \frac{\sqrt{p \cdot (1 - p)}}{\sqrt{n}}\right) = \mathcal{P}(-u_\alpha \leq \mathcal{Z} \leq u_\alpha) = 1 - \alpha$$

Remarque :

En supposant que $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$:

- Pour obtenir un intervalle de fluctuation à 95 %, il faut choisir : $\alpha = 0,05$; $u_\alpha = 1,96$:

La fréquence des succès F_n appartient à l'intervalle $\left[p - 1,96 \cdot \frac{\sqrt{p(1-p)}}{\sqrt{n}} ; p + 1,96 \cdot \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$ 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 succès F_n appartient à l'intervalle $\left[p - 2,58 \cdot \frac{\sqrt{p(1-p)}}{\sqrt{n}} ; p + 2,58 \cdot \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$ avec une probabilité de 99 %.

Cet intervalle est appelé **intervalle de fluctuation asymptotique** à 99 % de F_n

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 $p \in]0; 1[$, il existe un entier n_0 tel que pour tout $n \geq n_0$:

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

C. Estimation de p :

Définition :

Pour tout nombre réel α appartenant à l'intervalle $]0; 1[$, 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 :

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

Remarque :

En pratique, n assez grand signifie :

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

Démonstration :

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

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

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

On a les égalités suivantes de probabilités :

$$\begin{aligned} \mathcal{P}\left(F_n \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]\right) &= \mathcal{P}\left(p - \frac{1}{\sqrt{n}} \leq F_n \leq p + \frac{1}{\sqrt{n}}\right) = \mathcal{P}\left(-\frac{1}{\sqrt{n}} \leq F_n - p \leq \frac{1}{\sqrt{n}}\right) \\ &= \mathcal{P}\left(-\frac{1}{\sqrt{n}} \leq p - F_n \leq \frac{1}{\sqrt{n}}\right) = \mathcal{P}\left(F_n - \frac{1}{\sqrt{n}} \leq p \leq F_n + \frac{1}{\sqrt{n}}\right) = \mathcal{P}\left(p \in \left[F_n - \frac{1}{\sqrt{n}} ; F_n + \frac{1}{\sqrt{n}}\right]\right) \end{aligned}$$

On en déduit la inégalité :

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

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 $[0 ; 1]$ 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	$\frac{1}{4}$	1

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

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

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

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

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

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

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

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

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

On en déduit les comparaisons suivantes :

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

Ajoutons p à chaque membre de l'inégalité :

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

On vient d'établir l'inclusion :

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

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

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

A l'aide de la calculatrice, on obtient :

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

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

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathcal{P}\left(p - 2 \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq F_n \leq p + 2 \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right) \\ = \mathcal{P}(-2 \leq \mathcal{Z} \leq 2) = 1 - \alpha \simeq 0,954 \end{aligned}$$

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 $[0,951 ; 0,957]$ centré en 0,954.

$$n \geq n_0 \implies \mathcal{P}\left(p - 2 \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq F_n \leq p + 2 \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right) > 0,95$$

De l'inclusion des intervalles :

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

On en déduit la comparaison des probabilités :

$$\begin{aligned} \mathcal{P}\left(F_n \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]\right) \\ \geq \mathcal{P}\left(F_n \in \left[p - 2 \frac{\sqrt{p(1-p)}}{\sqrt{n}} ; p + 2 \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]\right) > 0,95 \end{aligned}$$

Remarque :

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

Pour tout $\alpha \in]0; 1[$, il existe un nombre réel u_α positif tel que :

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

Pour $p \in]0; 1[$, 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 $F_n = \frac{\mathcal{X}_n}{n}$ la fréquence de succès liée à l'échantillon de taille n .

Voici quelques valeurs particulières :

- Pour $\alpha = 0,05$, on a $u_{0,05} \simeq 1,96$:

$$\Rightarrow \mathcal{P}(-1,96 \leq \mathcal{Z} \leq 1,96) = 0,95$$

$$\Rightarrow \lim_{n \rightarrow +\infty} \mathcal{P}\left(-1,96 \leq \frac{\mathcal{X}_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \leq 1,96\right) = 0,95$$

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

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

- Pour $\alpha = 0,01$, on a $u_{0,01} \simeq 2,58$:

$$\Rightarrow \mathcal{P}(-2,58 \leq \mathcal{Z} \leq 2,58) = 0,99$$

$$\Rightarrow \lim_{n \rightarrow +\infty} \mathcal{P}\left(-2,58 \leq \frac{\mathcal{X}_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \leq 2,58\right) = 0,99$$

$$\Rightarrow \lim_{n \rightarrow +\infty} \mathcal{P}\left(n \cdot p - 2,58 \cdot \sqrt{n \cdot p \cdot (1-p)} \leq \mathcal{X}_n \leq n \cdot p + 2,58 \cdot \sqrt{n \cdot p \cdot (1-p)}\right) = 0,99$$

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