

Biostatistique

Loi binomiale

Des échantillons de sang sont utilisés pour dépister certaines maladies.



Source 

Lorsque la maladie est rare, les travailleurs de la santé combinent ou regroupent parfois les échantillons de sang d'un groupe d'individus en un seul lot, puis le testent.



Source 

Si le résultat du test du lot est

négatif, aucun autre test n'est

nécessaire car aucun des individus

du groupe n'est atteint de la maladie.

Si le résultat du test du lot est positif, chaque individu du groupe doit être testé.



Source

Quelle est la probabilité qu'aucun test supplémentaire ne soit nécessaire pour les individus de l'échantillon ?

La réponse à cette question peut être trouvée en
utilisant ce qu'on appelle la distribution binomiale.

De nombreuses décisions dans les affaires, les assurances et d'autres situations de la vie réelle sont prises en attribuant des probabilités à tous les résultats possibles concernant la situation, puis en évaluant les résultats.

Par exemple, une vendeuse peut calculer la probabilité qu'elle réalise 0, 1, 2 ou 3 ventes ou plus en une seule journée.



Source

Une variable aléatoire est une variable dont les valeurs sont déterminées au hasard.

Les variables discrètes ont un nombre fini de valeurs possibles ou un nombre de valeurs qui peuvent être comptées.

Par exemple, le d'appels téléphoniques reçus après la diffusion d'une publicité télévisée sont des exemples de variables discrètes, car ils peuvent être comptés.

Les variables qui peuvent prendre toutes les valeurs dans l'intervalle entre deux valeurs données sont appelées **variables continues**.

Les variables qui peuvent prendre toutes les valeurs dans l'intervalle entre deux valeurs données sont appelées **variables continues**.

Par exemple, si **la température** passe de **62°** à **78°** en **24 heures**, elle est passée par tous les nombres possibles de 62 à 78.

Les variables qui peuvent prendre toutes les valeurs dans l'intervalle entre deux valeurs données sont appelées variables continues.

Par exemple, si la température passe de 62° à 78° en 24 heures, elle est passée par tous les nombres possibles de 62 à 78.

Les **variables aléatoires continues** peuvent prendre un **nombre infini** de valeurs et peuvent être des valeurs **décimales et fractionnaires**

La procédure présentée ici pour

construire une distribution de

probabilité pour une variable

aléatoire discrète utilise

l'expérience de probabilité

consistant à lancer trois pièces.



Source 



Zéro face	Une	Deux	Trois



Zéro face	Une	Deux	Trois

Une distribution de probabilité discrète se compose des valeurs qu'une variable aléatoire peut prendre et des probabilités correspondantes des valeurs. Les probabilités sont déterminées théoriquement ou par observation.

De nombreux types de problèmes de probabilité n'ont que deux résultats ou peuvent être réduits à deux résultats. Par exemple, lorsqu'une pièce de monnaie est lancée, elle peut tomber pile ou face.



Source 

Dans un match de basket, une équipe gagne ou perd.



Source 

Un traitement médical peut être classé comme efficace ou inefficace, selon les résultats.



Source 

Une personne peut être classée comme

ayant une tension artérielle **normale** ou

anormale, selon la mesure de la jauge de

tension artérielle.



Source

Des situations comme celles-ci sont appelées **expériences binomiales**.

Les probabilités peuvent être calculées à l'aide de formule de probabilité binomiale

$$P(x = k) = \frac{n!}{(n - k)! \times k!} \cdot p^k \cdot q^{n-k}$$

Les probabilités peuvent être calculées à l'aide de formule de probabilité binomiale

$$P(x = k) = \frac{n!}{(n - k)! \times k!} \cdot p^k \cdot q^{n-k}$$

n : nombre d'essais

x : nombre de succès parmi n essais

p : probabilité de succès dans un essai

q : probabilité d'échec dans un essai ($q = 1 - p$)

Une enquête a révélé qu'un Américain sur cinq déclare avoir consulté un médecin au cours d'un mois donné.

Si 10 personnes sont sélectionnées au hasard, trouvez la probabilité qu'exactement 3 aient consulté un médecin le mois dernier.

Dans ce cas: $n = 10$; $k = 3$; $p = \frac{1}{5}$; $q = \frac{4}{5}$

- $P(3) = \frac{10!}{(10-3)! \times 3!} \cdot \left(\frac{1}{5}\right)^3 \cdot \left(\frac{4}{5}\right)^7 \approx 0,201$

Ainsi, il y a une probabilité de 0,201 que dans un échantillon aléatoire de 10 personnes, exactement 3 d'entre elles aient consulté un médecin au cours du dernier mois.

La moyenne, la variance et l'écart type d'une variable qui a la distribution binomiale peuvent être trouvés en utilisant les formules suivantes:

La moyenne, la variance et l'écart type d'une variable qui a la distribution binomiale peuvent être trouvés en utilisant les formules suivantes:

$$\text{Moyenne} = n \cdot p$$

$$\text{Variance} = n \cdot p \cdot q$$

$$\text{Ecart-type} = \sqrt{n \cdot p \cdot q}$$

Le *Sourcebook of Criminal Justice Statistics* indique que 65% des Américains préfèrent condamner les conducteurs ivres à la prison même s'ils n'ont pas causé d'accident.

Si un nombre aléatoire de 1000 individus est sélectionné, trouvez la moyenne, la variance et l'écart type des personnes qui ressentent cela

Le *Sourcebook of Criminal Justice Statistics* indique que 65% des Américains préfèrent condamner les conducteurs ivres à la prison même s'ils n'ont pas causé d'accident.

Si un nombre aléatoire de 1000 individus est sélectionné, trouvez la moyenne, la variance et l'écart type des personnes qui ressentent cela

$$\text{Moyenne} = 1000 \times 0,65 = 650$$

Le *Sourcebook of Criminal Justice Statistics* indique que 65% des Américains préfèrent condamner les conducteurs ivres à la prison même s'ils n'ont pas causé d'accident.

Si un nombre aléatoire de 1000 individus est sélectionné, trouvez la moyenne, la variance et l'écart type des personnes qui ressentent cela

$$\text{Moyenne} = 1000 \times 0,65 = 650$$

$$\text{Variance} = 1000 \times 0,65 \times 0,35 = 227,5$$

Le *Sourcebook of Criminal Justice Statistics* indique que 65% des Américains préfèrent condamner les conducteurs ivres à la prison même s'ils n'ont pas causé d'accident.

Si un nombre aléatoire de 1000 individus est sélectionné, trouvez la moyenne, la variance et l'écart type des personnes qui ressentent cela

$$\text{Moyenne} = 1000 \times 0,65 = 650$$

$$\text{Variance} = 1000 \times 0,65 \times 0,35 = 227,5$$

$$\text{Ecart-type} = \sqrt{1000 \times 0,65 \times 0,35} = 15,08$$

Loi normale

Les chercheurs médicaux ont déterminé des intervalles dits normaux pour la tension artérielle, le cholestérol, les triglycérides, etc. d'une personne.

La plage normale de la pression artérielle systolique est de 110 à 140.



Source



Source

L'intervalle normal pour les triglycérides d'une personne est de 30 à 200 milligrammes par décilitre (mg/dl).

Un médecin peut déterminer si les statistiques vitales d'un patient se situent dans l'intervalle normal ou si un certain type de traitement est nécessaire pour corriger une condition et éviter de futures maladies.

Comment déterminer les intervalles dits normaux ?

Si une variable aléatoire a une distribution de probabilité dont le graphique est continu, en forme de cloche et symétrique, on l'appelle une distribution normale.

L'équation mathématique d'une distribution normale est

$$x \sim N(m; \sigma) = \frac{1}{\sigma \times \sqrt{2\pi}} e^{-0,5(\frac{x-m}{\sigma})^2}$$

$$x \sim N(m; \sigma) = \frac{1}{\sigma \times \sqrt{2\pi}} e^{-0,5(\frac{x-m}{\sigma})^2}$$

Moyenne

Ecart-type

Variable aléatoire réelle

$$m \text{ et } \sigma > 0$$

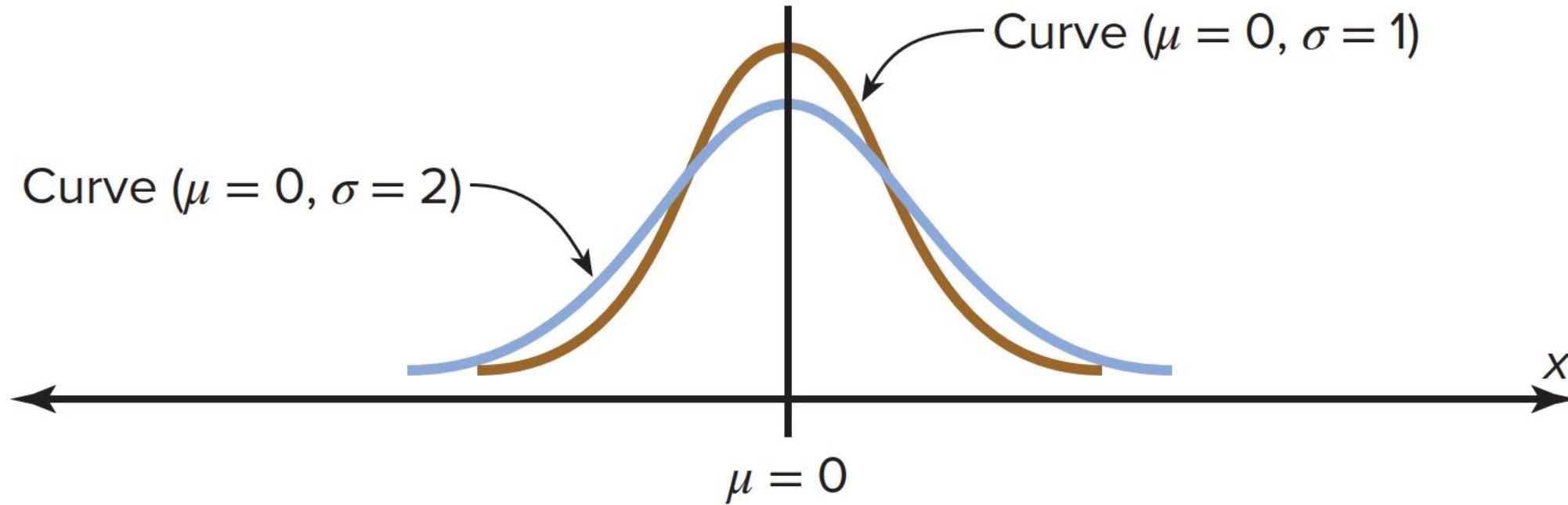
- Cas particulier de la loi normale $N(m; \sigma)$

- Loi normale centrée et réduite $N(0,1)$

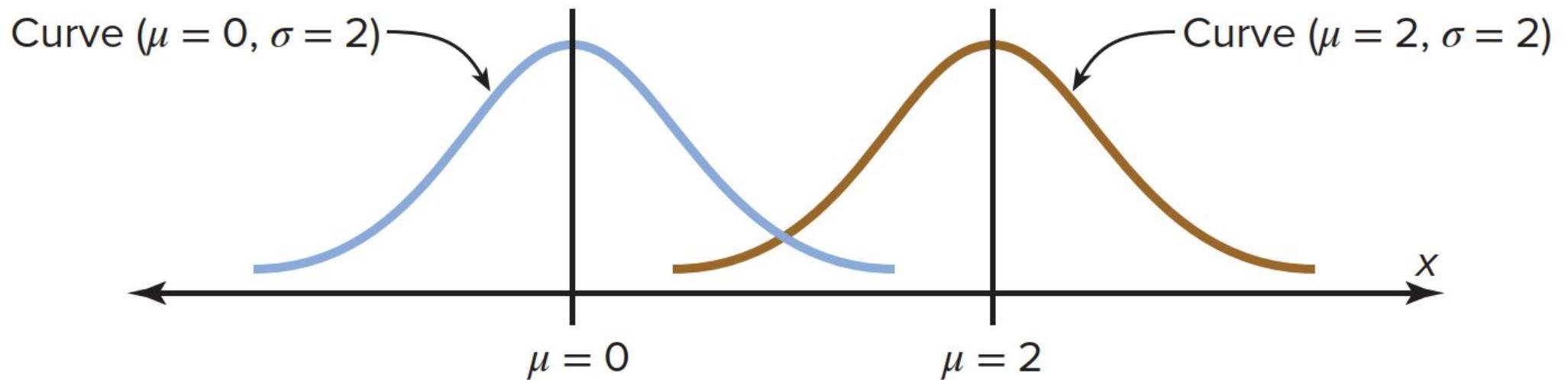
$$N(0; 1) = \frac{1}{\sqrt{2\pi}} e^{-0,5(\frac{\mu - o}{1})^2} = \frac{1}{\sqrt{2\pi}} e^{-0,5 \times \mu^2}$$

La forme et la position d'une courbe de distribution normale dépendent de deux paramètres: la moyenne et l'écart type.

Mêmes moyennes mais écarts-types différents

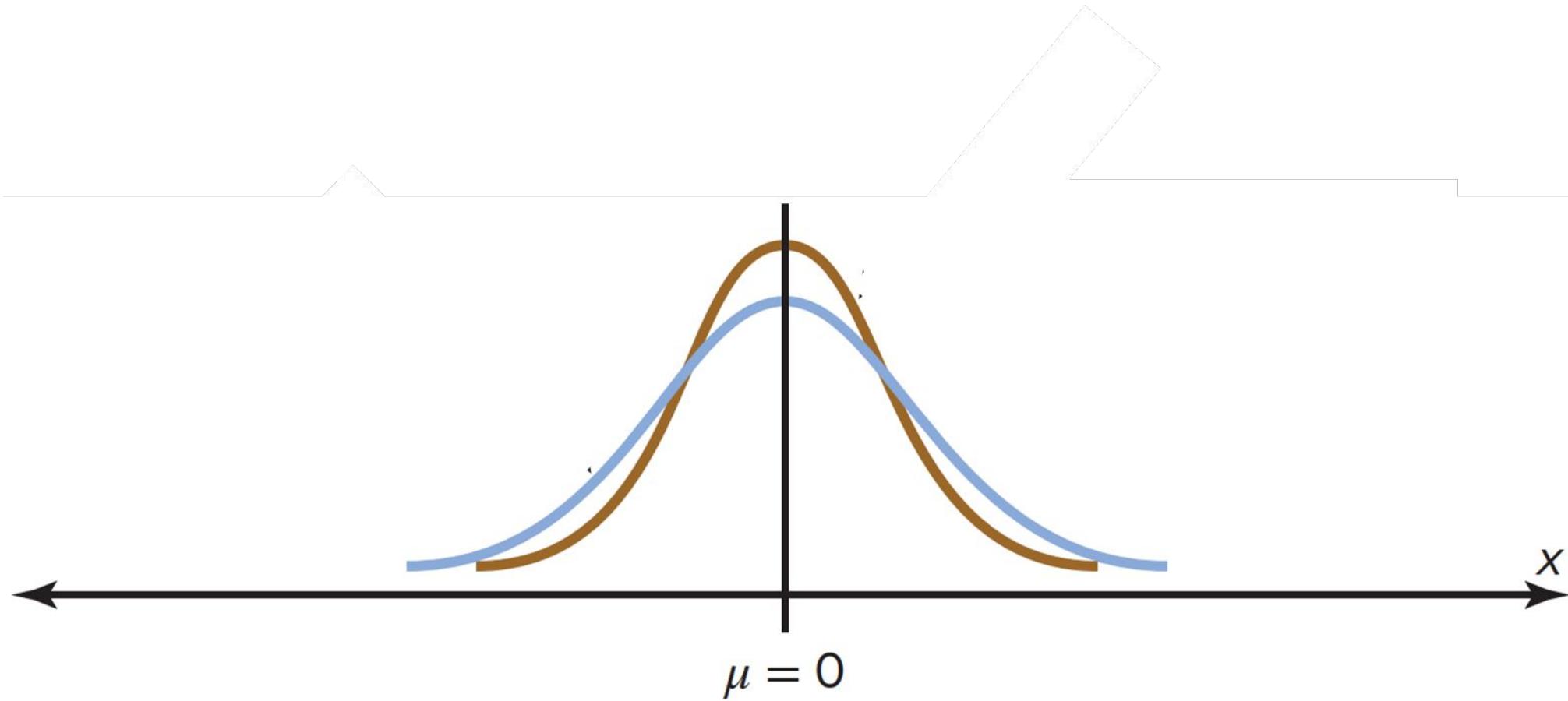


Moyennes différentes mais mêmes écarts-types

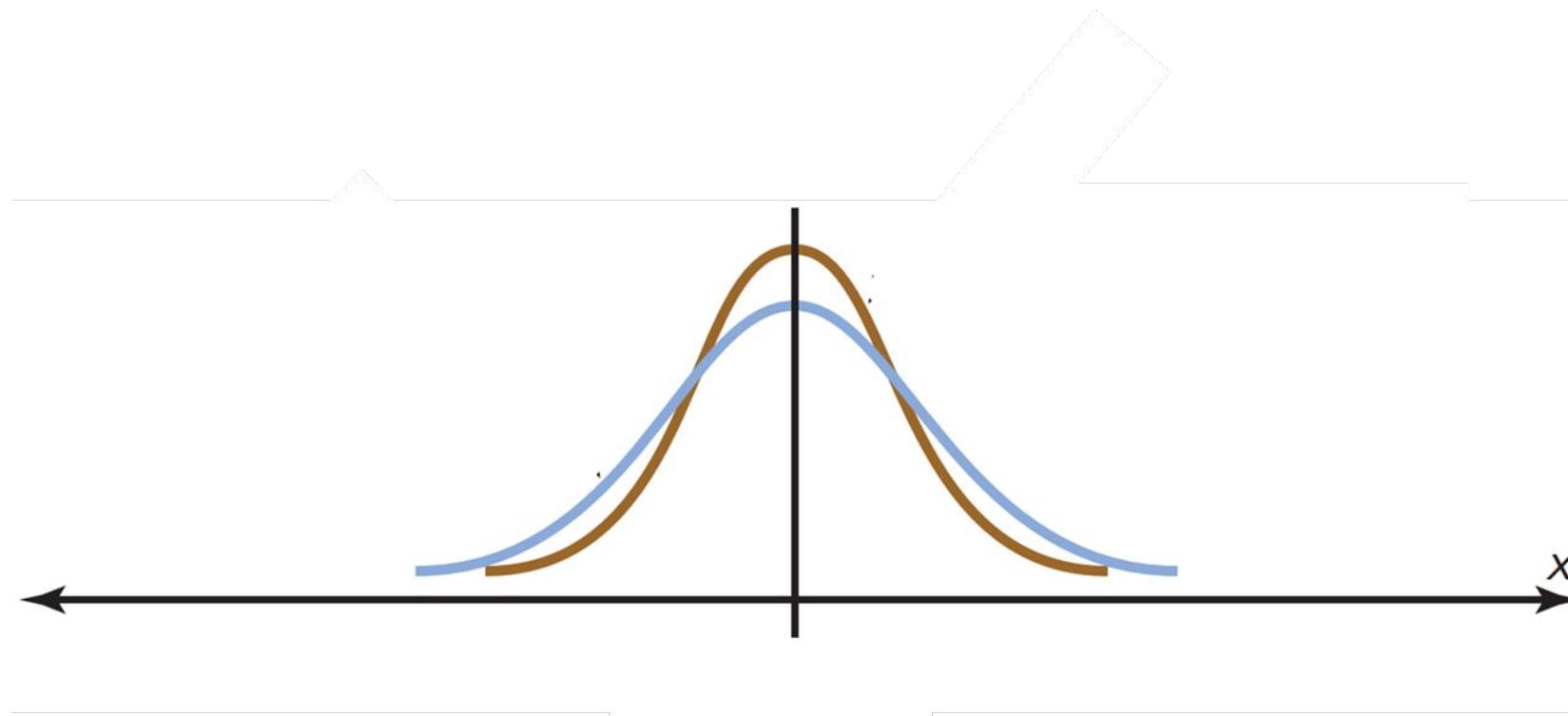


Propriétés de la distribution normale théorique

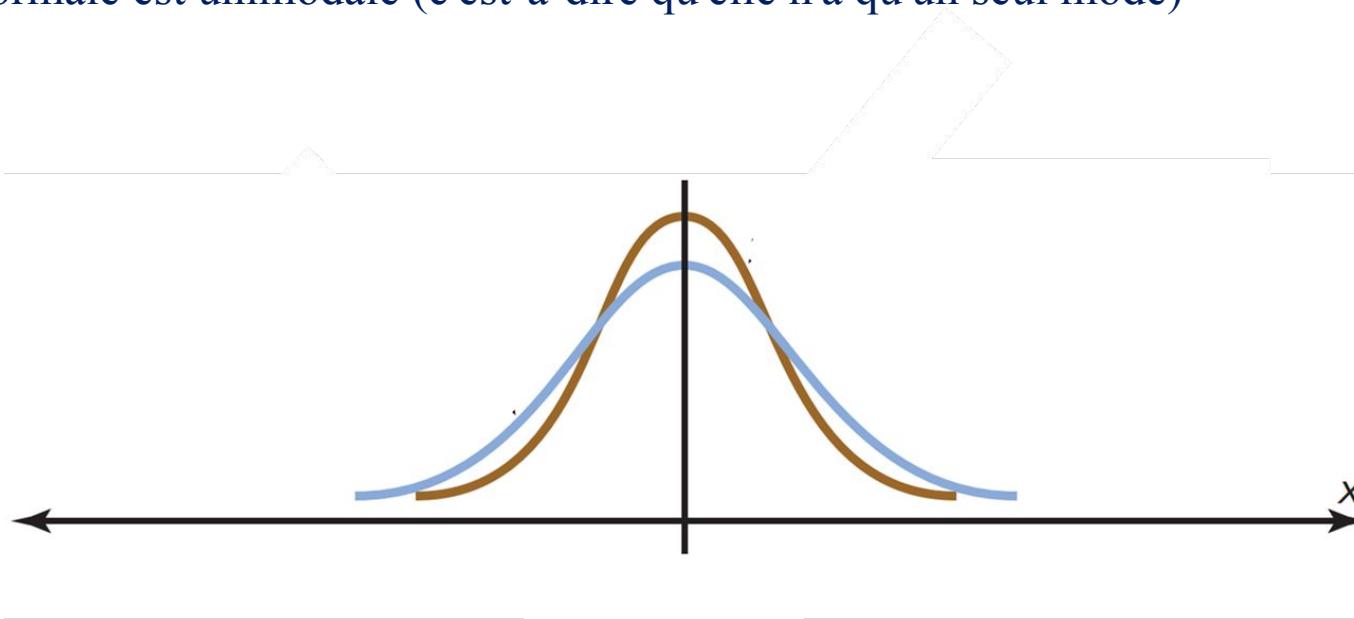
1. Une courbe de distribution normale est en forme de cloche

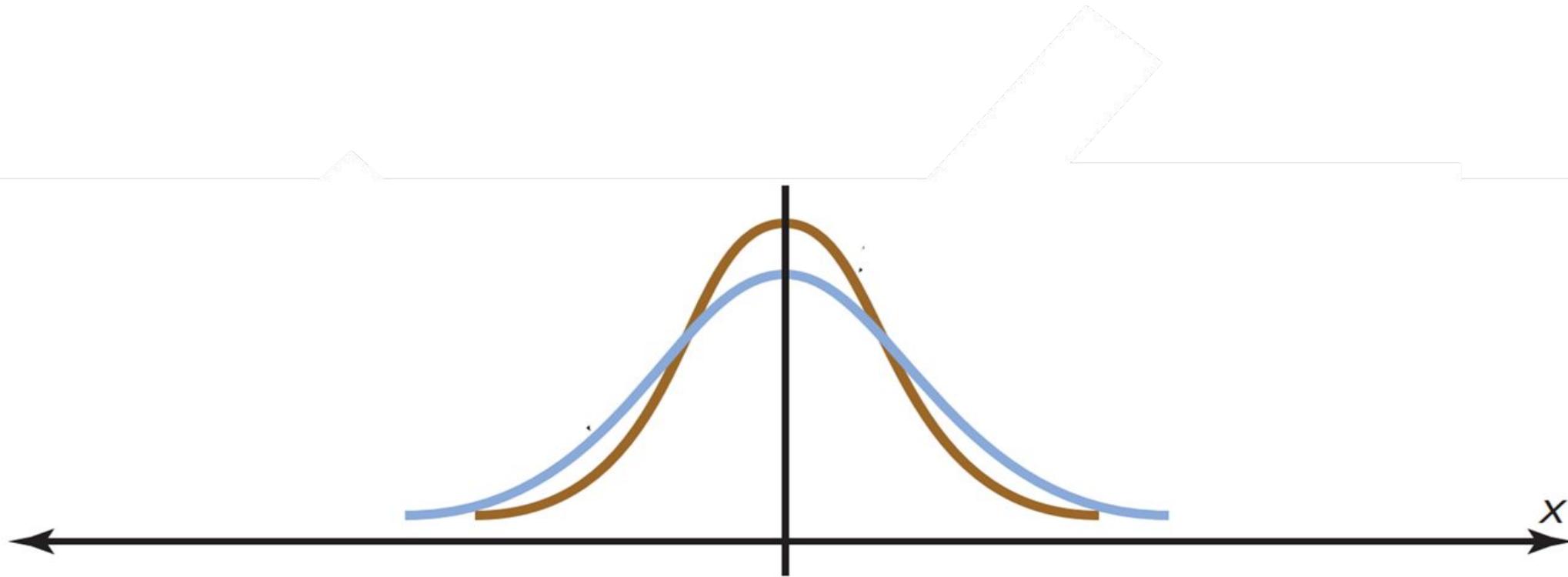


1. Une courbe de distribution normale est en forme de cloche
2. La moyenne, la médiane et le mode sont égaux et se situent au centre de la distribution



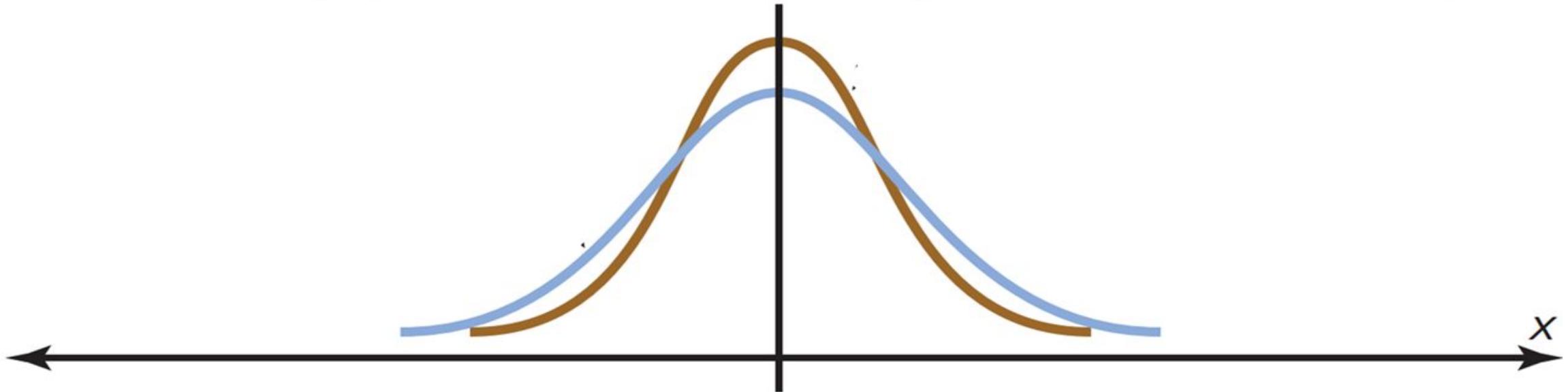
1. Une courbe de distribution normale est en forme de cloche
2. La moyenne, la médiane et le mode sont égaux et se situent au centre de la distribution
3. Une courbe de distribution normale est unimodale (c'est-à-dire qu'elle n'a qu'un seul mode)





4. La courbe est symétrique par rapport à la moyenne

1. Une courbe de distribution normale est en forme de cloche



5. La courbe est continue ; c'est-à-dire qu'il n'y a pas d'espaces ou de trous.

Calcul des probabilités à partir de la loi normale centrée et réduite

Calcul des probabilités à partir de la loi normale centrée et réduite

$$x \sim N(0; 1)$$

Il est important de savoir qu'un **calcul de probabilité** dans le cadre de **lois continues** est un calcul d'aire et donc d'intégrale.

$$\bullet \ P(u \leq a) = \int_{-\infty}^a N(0; 1)$$

- $P(u \leq a) = \int_{-\infty}^a N(0; 1)$
- $P(u \leq a) = \frac{1}{\sqrt{2\pi}} \times \int_{2\pi}^a e^{-0,5u^2} du$

- $P(u \leq a) = \int_{-\infty}^a N(0; 1)$
- $P(u \leq a) = \frac{1}{\sqrt{2\pi}} \times \int_{2\pi}^a e^{-0,5u^2} du$
- $\int_{2\pi}^a e^{-0,5u^2} du$ La primitive n'existe pas

Solution

Intégral de Gauss !!

Maintenant, si $x \in [0, \sqrt{n}]$, on a $\frac{x^2}{n} \leq 1$. Il est clair que pour $u \in [1, +\infty]$, on a $\ln(1+u) \leq u$ (la fonction $u \mapsto \ln(1+u)$ est concave sur $[1, +\infty]$ et y admet une dérivée seconde négative de sorte que son graphe est au dessous de sa tangente en $(0,0)$ sur $[1, +\infty]$). On en déduit que

$$g_n(x) = e^{n \ln(1-\frac{x^2}{n})} \leq e^{n(-\frac{x^2}{n})} = e^{-x^2} = g(x).$$

Ainsi, $\forall n \in \mathbb{N}^*$, $|g_n| \leq g$ avec g continue et intégrable sur $[0, +\infty]$.

En résumé : La suite de fonctions $(g_n)_{n \in \mathbb{N}^*}$ converge simplement vers la fonction g sur $[0, +\infty]$ et la fonction est continue sur $[0, +\infty]$.

Chaque fonction g_n est intégrable sur $[0, +\infty]$.

Il existe une fonction φ continue et intégrable sur $[0, +\infty]$ telle que $\forall n \in \mathbb{N}^*$, $|g_n| \leq \varphi$ à savoir $\varphi = g$.

D'après le théorème de convergence dominée, on a alors

$$\int_0^{+\infty} e^{-x^2} dx = \int_0^{+\infty} g(x) dx = \lim_{n \rightarrow +\infty} \int_0^{+\infty} g_n(x) dx = \lim_{n \rightarrow +\infty} \int_0^{+\infty} g_n(x) dx = \lim_{n \rightarrow +\infty} \int_0^{\sqrt{n}} \left(1 - \frac{x^2}{n}\right)^n dx.$$

ii) Démonstration de $\lim_{n \rightarrow +\infty} \int_0^{\sqrt{n}} \left(1 - \frac{x^2}{n}\right)^n dx$. Pour $n \in \mathbb{N}$, posons $I_n = \int_0^{\sqrt{n}} \left(1 - \frac{x^2}{n}\right)^n dx$. Le changement de variables $x = \sqrt{n}t$ et donc $dx = \sqrt{n}dt$ nous fournit

$$I_n = \int_0^{\sqrt{n}} \left(1 - \frac{x^2}{n}\right)^n dx = \int_{-\sqrt{n}/2}^0 (1 - \cos^2 t)^n = \sqrt{n} \int_0^{n/2} \sin^{2n+1} t dt = \sqrt{n} W_{2n+1},$$

où W_n est la n -ème intégrale de Wallis. L'étude de ces intégrales montre que

$$I_n \underset{n \rightarrow +\infty}{\sim} \sqrt{n} \sqrt{\frac{\pi}{2(2n+1)}} = \frac{\sqrt{\pi}}{2},$$

et on retrouve encore $\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$.

iii) Bonne. Montrons que la suite de fonctions $(f_n)_{n \in \mathbb{N}^*}$ converge uniformément vers la fonction f sur $[0, +\infty]$.

Soit $n \in \mathbb{N}^*$. On a vu précédemment que $\forall x \in [0, +\infty]$, $f(x) - f_n(x) \geq 0$.

Posons $h_n = f - f_n$ et étudions la fonction h_n . Il est déjà clair que h_n est continue, positive sur $[0, +\infty]$ et décroissante sur $[0, \infty]$.

h_n est dérivable sur $[0, n]$ et pour $x \in [0, n]$,

$$h'_n(x) = -e^{-x} + \left(1 - \frac{x}{n}\right)^{n-1}.$$

Par croissance de la fonction exponentielle sur \mathbb{R} , on a pour $x \in [0, n]$

$$\operatorname{sgn}(h'_n(x)) = \operatorname{sgn}(e^{-(n-1)\ln(1-\frac{x}{n})} - e^{-x}) = \operatorname{sgn}((n-1)\ln(1-\frac{x}{n}) - (-x)) = \operatorname{sgn}((n-1)\ln(1-\frac{x}{n}) + x).$$

Pour $x \in [0, n]$, posons $k_n(x) = (n-1)\ln(1-\frac{x}{n}) + x$. k_n est dérivable sur $[0, n]$ et pour $x \in [0, n]$,

$$k'_n(x) = (n-1)\frac{1}{1-\frac{x}{n}} + 1 - \frac{n-1}{n-x} + 1 - \frac{1-x}{n-x}.$$

k_n est donc strictement croissante sur $[0, 1]$ et strictement décroissante sur $[1, n]$. Comme $k_n(0) = 0$, on a $k_n(1) > 0$ et comme $\lim_{x \rightarrow n} k_n(x) = -\infty$, on en déduit qu'il existe $a_n \in [1, n]$ tel que $k_n(a_n) = 0$ ou encore $k'_n(a_n) = 0$. De plus, k_n est positive sur $[0, a_n]$ et négative sur $[a_n, n]$ et il en est de même de k'_n .

Mais alors, h_n est croissante sur $[0, a_n]$ et décroissante sur $[a_n, n]$. Comme de plus h_n est continue sur $[0, +\infty]$ et décroissante sur $[0, a_n]$, h_n est décroissante sur $[a_n, +\infty]$. En résumé, h_n est positive sur $[0, +\infty]$, croissante sur $[0, a_n]$ et décroissante sur $[a_n, +\infty]$.

$$\forall x \in [0, +\infty], 0 \leq h_n(x) \leq h_n(a_n).$$

Maintenant, l'égalité $h'_n(a_n) = 0$ fournit $e^{-a_n} = \left(1 - \frac{a_n}{n}\right)^{n-1}$ et donc

$$h_n(a_n) = e^{-a_n} - \left(1 - \frac{a_n}{n}\right)^n = e^{-a_n} - \left(1 - \frac{a_n}{n}\right)^{n-1} \left(1 - \frac{a_n}{n}\right) = e^{-a_n} - \left(1 - \frac{a_n}{n}\right) e^{-a_n} = \frac{a_n e^{-a_n}}{n}.$$

Rafin, la fonction $u : x \mapsto x e^{-x}$ est dérivable sur $[0, +\infty]$ de dérivée $u' : x \mapsto (1-x)e^{-x}$. La fonction u admet donc un maximum en 1 égal à $1 \times e^{-1} = \frac{1}{e}$. On en déduit que $h_n(a_n) = \frac{a_n e^{-a_n}}{n} \leq \frac{1}{e n}$. On a montré que

$$\forall x \in [0, +\infty], 0 \leq f(x) - f_n(x) \leq \frac{1}{e n}.$$

Mais alors $\sup\{|f(x) - f_n(x)|, x \in [0, +\infty]\} \leq \frac{1}{e n}$ et puisque $\lim_{n \rightarrow +\infty} \frac{1}{e n} = 0$, on a montré que

la suite de fonctions $(f_n)_{n \in \mathbb{N}^*}$ converge uniformément vers la fonction f sur $[0, +\infty]$.

3) Calcul de $\Gamma\left(\frac{1}{2}\right) = \int_0^{+\infty} \frac{e^{-x}}{\sqrt{x}} dx$.

a) Existence de l'intégrale. La fonction $x \mapsto \frac{e^{-x}}{\sqrt{x}}$ est continue sur $[0, +\infty]$ et donc localement intégrable sur $[0, +\infty]$, positive et équivalente en 0 à $\frac{1}{\sqrt{x}}$ et donc intégrable sur un voisinage de 0, négligeable devant $\frac{1}{x^2}$ en $+ \infty$ et donc intégrable sur un voisinage de $+\infty$. Finalement, la fonction $x \mapsto \frac{e^{-x}}{\sqrt{x}}$ est intégrable sur $[0, +\infty]$.

b) Calcul de l'intégrale. En posant $z = x^2$ et donc $dx = 2z dz$, on obtient

$$\begin{aligned} \int_0^{+\infty} \frac{e^{-x}}{\sqrt{x}} dx &= \int_0^{+\infty} \frac{e^{-\sqrt{z}}}{\sqrt{z}} 2z dz = 2 \int_0^{+\infty} e^{-\sqrt{z}} dz = \sqrt{\pi}. \\ \Gamma\left(\frac{1}{2}\right) &= \int_0^{+\infty} \frac{e^{-x}}{\sqrt{x}} dx = \sqrt{\pi}. \end{aligned}$$

4) Calcul de $I_n = \int_0^{+\infty} x^n e^{-x^2} dx$.

a) Existence. Soit n un entier naturel. La fonction $x \mapsto x^n e^{-x^2}$ est continue sur $[0, +\infty]$ et négligeable devant $\frac{1}{x^2}$ en $+ \infty$. Donc la fonction $x \mapsto x^n e^{-x^2}$ est intégrable sur $[0, +\infty]$ et l'intégrale proposée existe.

b) Calcul.

i) Relation de récurrence. Pour n entier naturel donné, posons $I_n = \int_0^{+\infty} x^n e^{-x^2} dx$.

Soit A un réel positif. Une intégration par parties fournit

$$\begin{aligned} \int_0^A x^{n+2} e^{-x^2} dx &= \int_0^A x^{n+1} x e^{-x^2} dx = \left[-\frac{1}{2} e^{-x^2} x^{n+1} \right]_0^A + \frac{n+1}{2} \int_0^A x^n e^{-x^2} dx \\ &= \frac{A^{n+1} e^{-A^2}}{2} + \frac{n+1}{2} \int_0^A x^n e^{-x^2} dx. \end{aligned}$$

Quand A tend vers $+\infty$, on obtient $\int_0^{+\infty} x^{n+2} e^{-x^2} dx = \frac{n+1}{2} \int_0^{+\infty} x^n e^{-x^2} dx$. D'où la relation de récurrence

$$\forall n \in \mathbb{N}, I_{n+2} = \frac{n+1}{2} I_n.$$

ii) Calcul de I_{2n} et I_{2n+1} . D'après 2), on a déjà $I_0 = \frac{\sqrt{\pi}}{2}$. D'autre part, $I_1 = \int_0^{+\infty} x e^{-x^2} dx = \left[-\frac{1}{2} e^{-x^2} \right]_0^{+\infty} = \frac{1}{2}$. Soit alors n un entier naturel non nul.

$$I_{2n} = \frac{(2n-1)2n-3}{2} \dots \frac{1}{2} I_0 = \frac{(2n)(2n-1)(2n-2) \dots 2\sqrt{\pi}}{2^{2n}(2n)(2n-2) \dots 2} = \frac{(2n)!}{2^{2n} n!} \frac{\sqrt{\pi}}{2}.$$

et

$$I_{2n+1} = \frac{(2n)2n-2}{2} \dots \frac{2}{2} I_1 = \frac{n!}{2}.$$

Ces égalités restent vraies pour $n = 0$, on a montré que

$$\forall n \in \mathbb{N}, \int_0^{+\infty} x^{2n} e^{-x^2} dx = \frac{(2n)!}{2^{2n} n!} \frac{\sqrt{\pi}}{2} \text{ et } \int_0^{+\infty} x^{2n+1} e^{-x^2} dx = \frac{n!}{2}.$$

Remarque. Par le changement de variables $u = x^2$, les intégrales précédentes s'écrivent respectivement

$$\int_0^{+\infty} x^{2n+1} e^{-x^2} dx = \frac{1}{2} \int_0^{+\infty} u^n e^{-u} du = \frac{1}{2} \Gamma(n+1) \text{ et } \int_0^{+\infty} x^{2n} e^{-x^2} dx = \frac{1}{2} \int_0^{+\infty} u^{n-1} e^{-u} du = \frac{1}{2} \Gamma(n+\frac{1}{2}).$$

et on a donc montré que

$$\forall n \in \mathbb{N}, \Gamma(n+1) = n! \text{ et } \Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!}{2^{2n} n!} \sqrt{\pi}.$$

5) Calcul de $\int_{-\infty}^{+\infty} e^{-(x+iy)^2} dx$. Pour y réel on pose $F(y) = \int_{-\infty}^{+\infty} e^{-(x+iy)^2} dx$.

a) Existence. Soit y un réel fixé. La fonction $x \mapsto e^{-(x+iy)^2}$ est continue sur \mathbb{R} . De plus, pour tout réel x,

$$|e^{-(x+iy)^2}| = |e^{-x^2+2ixy}| = e^{-x^2+y^2} < e^{-x^2+y^2}.$$

Cette dernière expression est négligeable devant $\frac{1}{x^2}$ quand x tend vers $+\infty$ ou vers $-\infty$.
Donc, pour tout réel y, la fonction $x \mapsto e^{-(x+iy)^2}$ est intégrable sur \mathbb{R} .

F est définie sur \mathbb{R} .

b) Calcul. Soit u un réel strictement positif. Soit $f : \mathbb{R} \times [-u, u] \rightarrow \frac{C}{(x,y) \mapsto e^{-(x+iy)^2}}$.

* Pour chaque y $\in [-u, u]$, la fonction $x \mapsto f(x,y)$ est continue et intégrable sur \mathbb{R} .

* f est pourvu sur $\mathbb{R} \times [-u, u]$ d'une dérivée partielle par rapport à sa deuxième variable y et pour $(x,y) \in \mathbb{R} \times [-u, u]$,

$$\frac{df}{dy}(x,y) = -2i(x+iy)e^{-(x+iy)^2}.$$

* Pour chaque z $\in \mathbb{R}$, la fonction $y \mapsto \frac{df}{dy}(x,y)$ est continue sur $[-u, u]$,

$$\frac{d}{dy} \left(\frac{df}{dy}(x,y) \right) = \frac{d}{dy} \left(-2i(x+iy)e^{-(x+iy)^2} \right) = 2\sqrt{x^2+y^2} e^{-x^2+y^2} < 2\sqrt{x^2+y^2} e^{-x^2+y^2} = q(x)$$

soit R et intégrable sur R car négligeable devant $\frac{1}{x^2}$ en $+ \infty$ et $- \infty$.

D'après le théorème de dérivation sous le signe somme, F est de classe C^1 sur tout segment de \mathbb{R} et donc sur \mathbb{R} et pour tout réel u, on a

$$F(y) = \int_{-\infty}^{+\infty} \frac{df}{dy}(x,y) dx = \int_{-\infty}^{+\infty} -2i(x+iy)e^{-(x+iy)^2} dx = \left[ie^{-(x+iy)^2} \right]_{-\infty}^{+\infty} = 0,$$

car $|e^{-(x+iy)^2}| = e^{-x^2+y^2} \xrightarrow[x \rightarrow \pm \infty]{} 0$.

F est donc constante sur \mathbb{R} et pour tout réel y, $F(y) = F(0) = \int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$.

$$\forall y \in \mathbb{R}, \int_{-\infty}^{+\infty} e^{-(x+iy)^2} dx = \sqrt{\pi}.$$

Rafin, puisque $\int_{-\infty}^{+\infty} e^{-(x+iy)^2} dx = C^1 \int_{-\infty}^{+\infty} e^{-x^2+2ixy} dx$, on a aussi montré que :

$$\forall y \in \mathbb{R}, \int_{-\infty}^{+\infty} e^{-x^2+2ixy} dx = \sqrt{\pi} e^{iy^2}.$$

Ordinateurs puissants

$$P(u \leq a) = N(0; 1)du = F(a)$$

$F(a)$: sur la Table de la loi normale

$F(a)$: sur la Table de la loi normale Centrée et Réduite

Table de la loi normale C.R

Unités centième de 0 à 0,09

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389

↑

Unités décimale de 0 à 3,9

$a = 1,25$

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441

$a = 1,25$

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441

$a = 1,2 + 0,05$

$$P(u < a) = P(u < 1,25) =$$

u	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.892	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441

$$P(u < a) = P(u \leq 1,25) = 0,8944$$

u	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.531
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.571
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.610
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.64
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.684
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.71
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.751
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.782
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.810
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.836

1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.859	0.88
------------	--------	--------	---------------	---------------	---------------	---------------	---------------	-------	------

1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.899
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.916
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.930
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.942

Lecture de la table de la loi normale centrée et réduite

$$P(u \leq 0) = 0,50$$

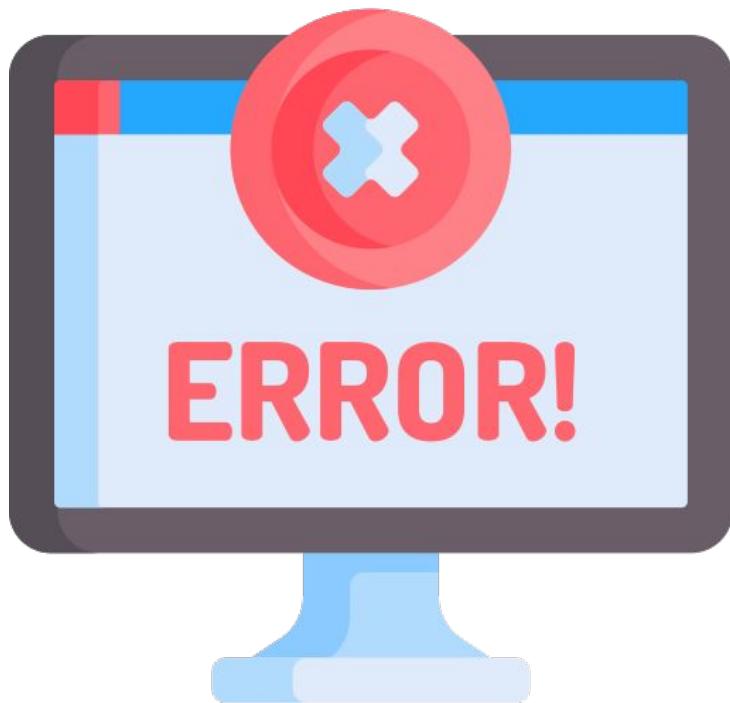
$$P(u \leq 1,12) = 0,8686$$

$$P(u \leq 1,48) = 0,9306$$

$$P(u \leq 1,96) = 0,975$$

$$P(u \leq 3,9) = 0,9999$$

$$P(u \geq 0,85) = 0,8023$$



$$P(u \geq 0,85) = 0,8023$$

Pour utiliser la table :
inférieur ou égal

Solution

$$P(u \geq a) + P(u \leq a) = 1$$

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

$$P(u \geq a) + P(u \leq a) = 1$$

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

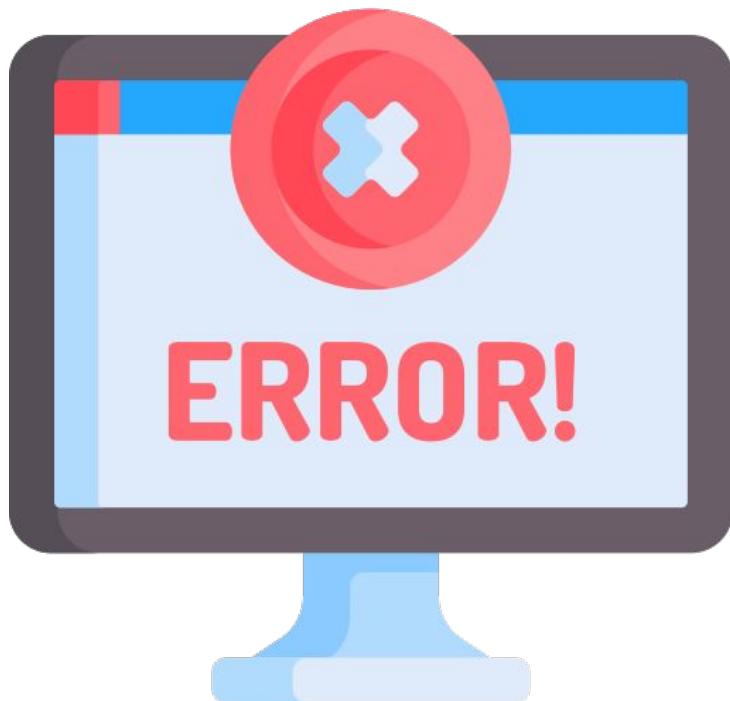
$$P(u \geq 0,85) = 1 - P(u \leq 0,85)$$

$$P(u \geq 0,85) = 1 - 0,8023$$

$$P(u \geq 0,85) = 1 - 0,8023$$

$$P(u \geq 0,85) = 0,1977$$

$$P(u \leq -0,81) = 0,791$$



$$P(u \leq -0,81) = 0,791$$

Pour utiliser la table

La valeur de a doit être
toujours positive

Solution

$$P(u \leq -a) = P(u \geq +a)$$

$$P(u \leq -0,81) = P(u \geq +0,81)$$

$$P(u \geq +0,81) = 1 - P(u \leq +0,81)$$

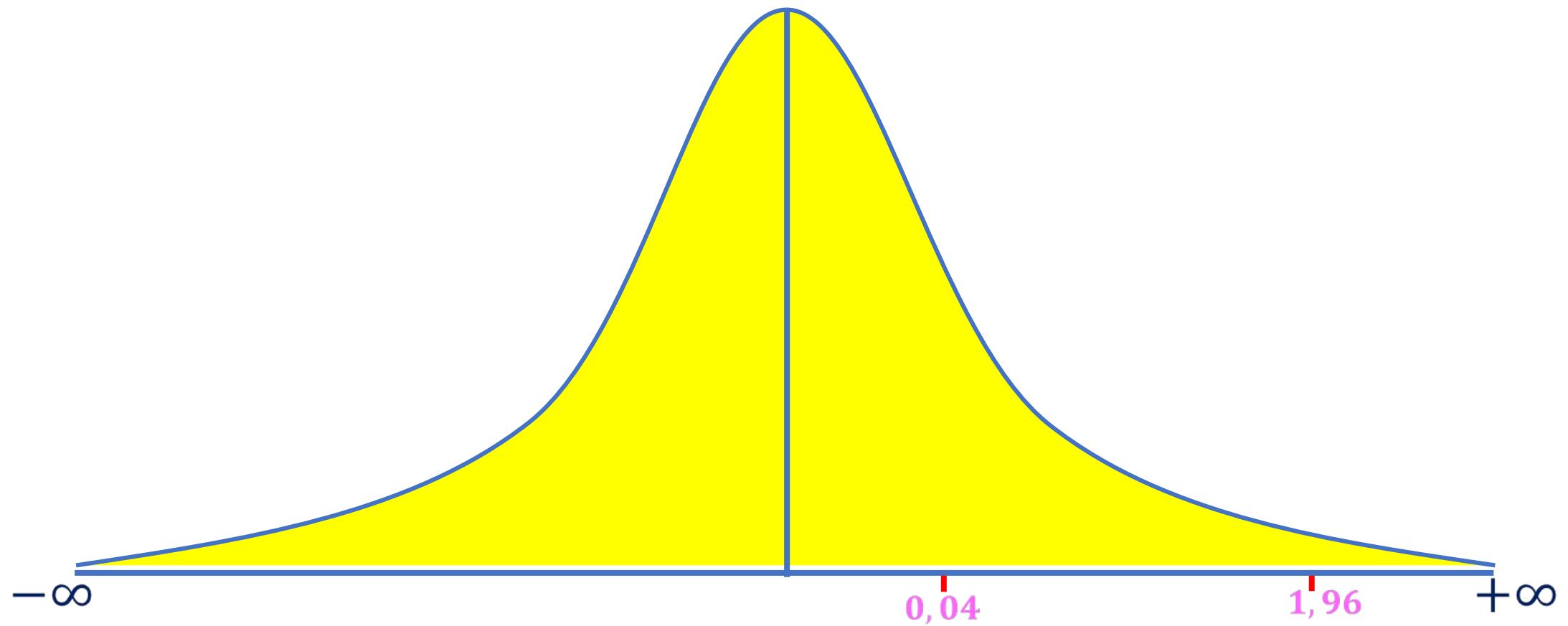
$$P(u \geq +0,81) = 1 - P(u \leq +0,81)$$

$$P(u \geq +0,81) = 1 - 0,791$$

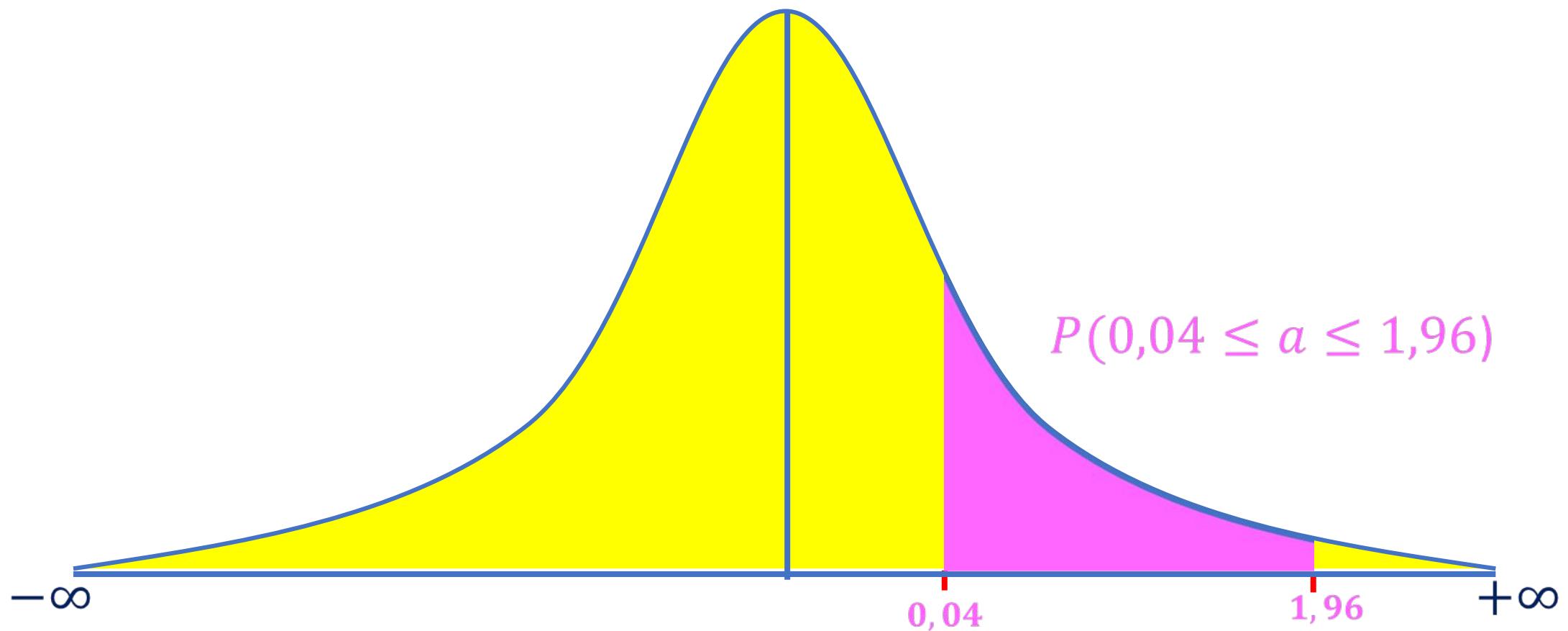
$$P(u \geq +0,81) = 0,209$$

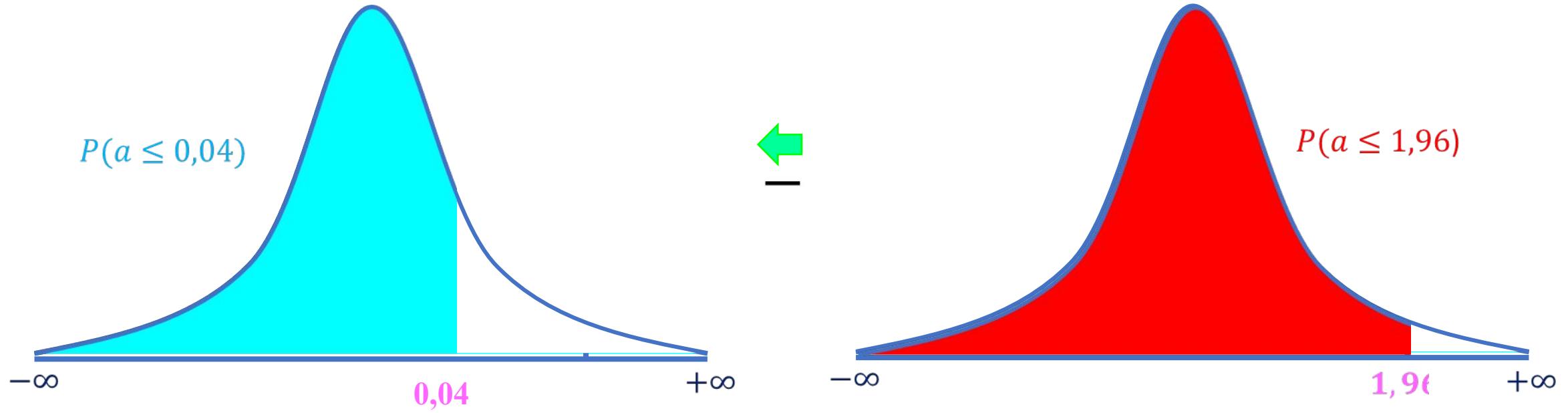
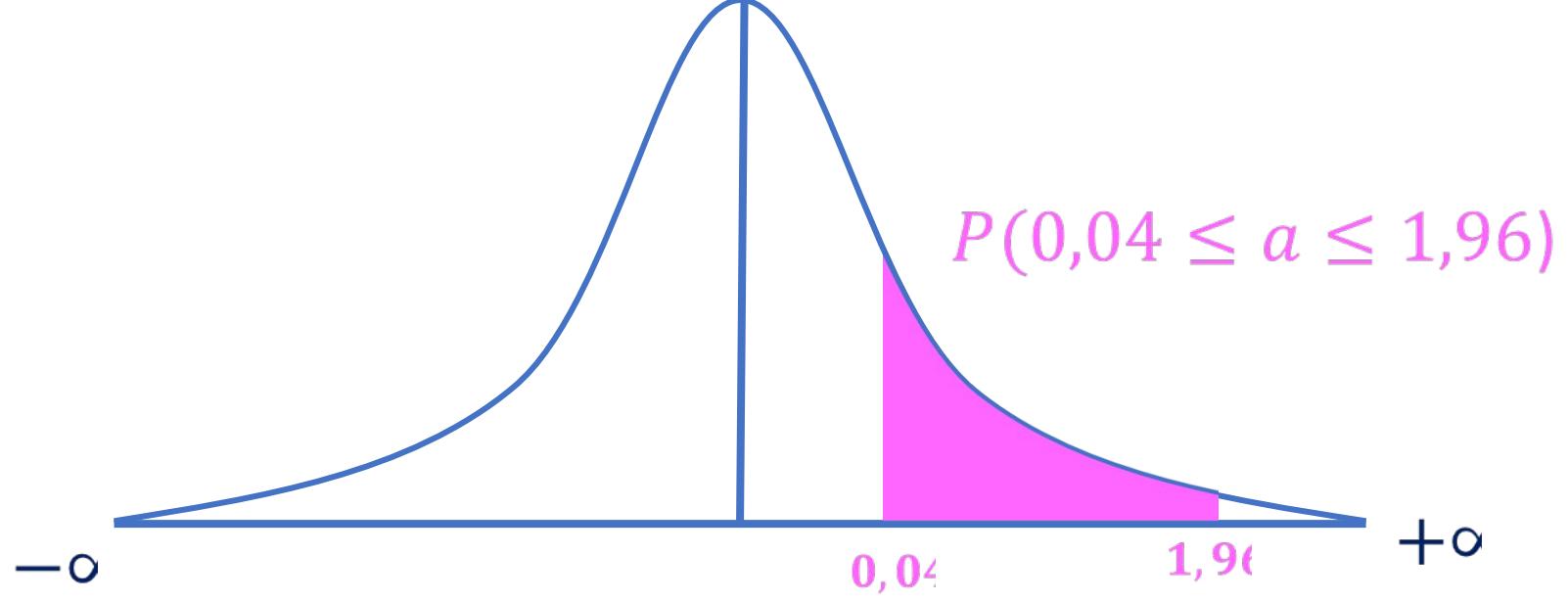
$$P(u \leq -0,81) = 0,209$$

$$P(0,04 \leq a \leq 1,96) =$$

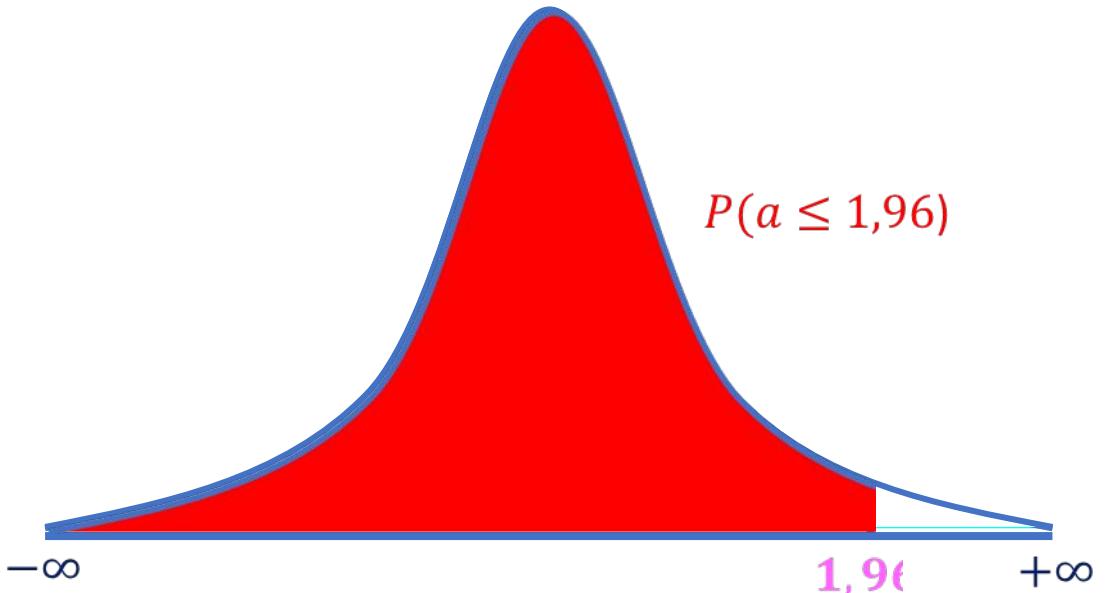
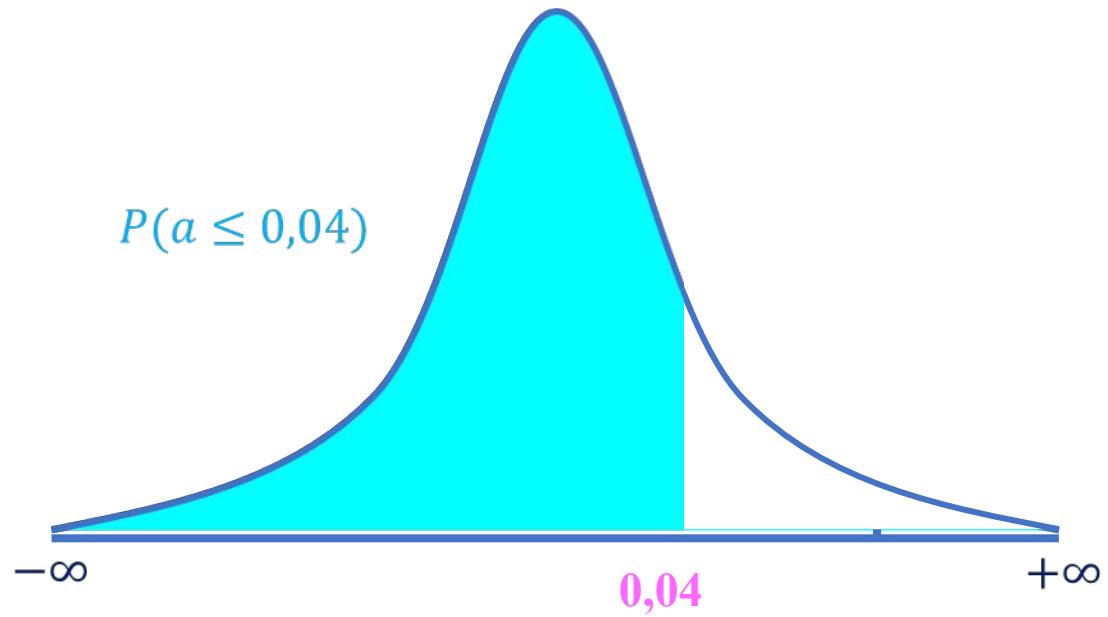


$$0,04 \leq a \leq 1,96$$





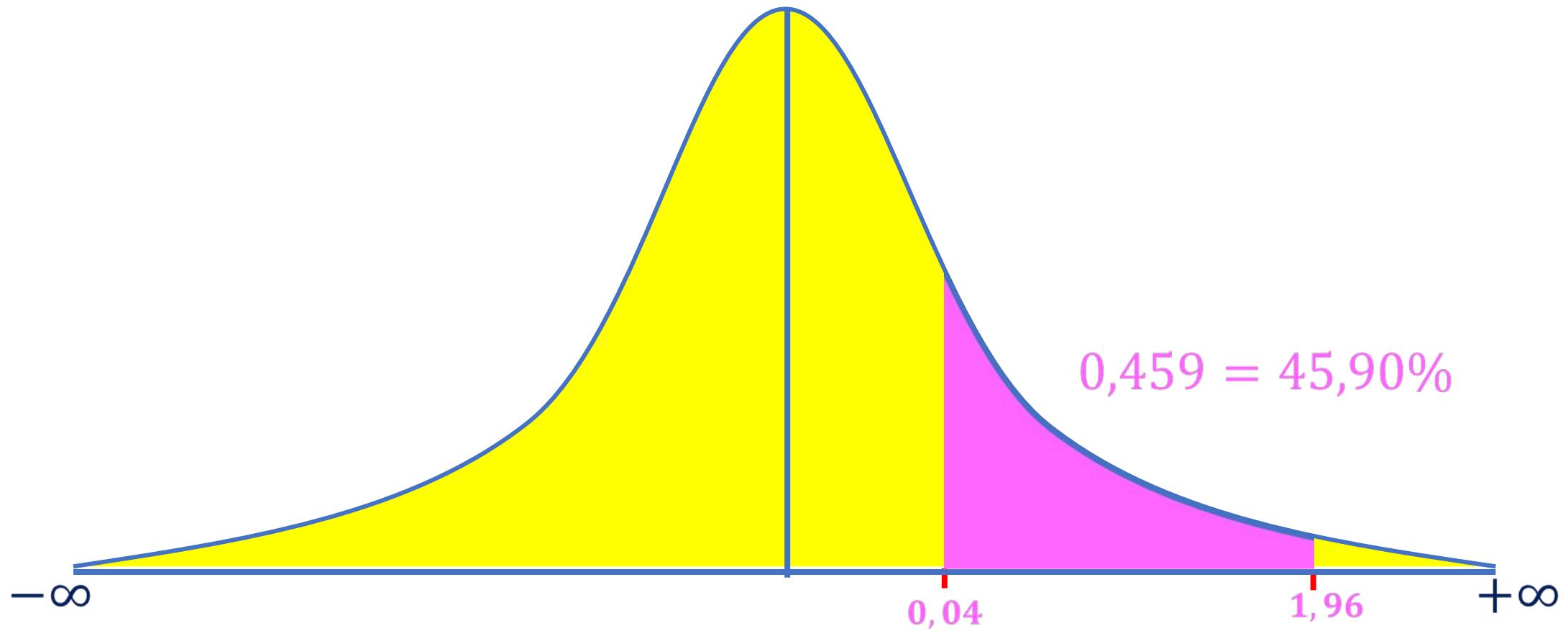
$$P(0,04 \leq a \leq 1,96) = P(u \leq 1,96) - P(u \leq 0,04)$$



$$P(0,04 \leq a \leq 1,96) = P(u \leq 1,96) - P(u \leq 0,04)$$

$$P(0,04 \leq a \leq 1,96) = 0,975 - 0,516$$

$$P(0,04 \leq a \leq 1,96) = 0,459$$



$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - P(u \leq -1,14)$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - P(u \geq +1,14)$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - P(u \leq -1,14)$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - P(u \geq +1,14)$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - [1 - P(u \leq +1,14)]$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - [1 - P(u \leq +1,14)]$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - P(u \leq -1,14)$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - P(u \geq +1,14)$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - [1 - P(u \leq +1,14)]$$

$$P(-1,14 \leq u \leq 2,58) = P(u \leq 2,58) - [1 - P(u \leq +1,14)]$$

$$P(-1,14 \leq u \leq 2,58) = 0,9951 - [1 - 0,8729]$$

$$P(-1,14 \leq u \leq 2,58) = 0,9951 - [1 - 0,8729]$$

$$P(-1,14 \leq u \leq 2,58) = 0,8823$$

Comment se fait le calcul d'une probabilité pour une variable qui suit une loi normale

Exemple

La taille des arbres dans la forêt suit la loi normale tel que $T \sim N(320, 9)$

La densité de probabilité pour la variable T

$$T \sim N(m, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0,5(\frac{t-m}{\sigma})^2}$$

La densité de probabilité pour la variable T

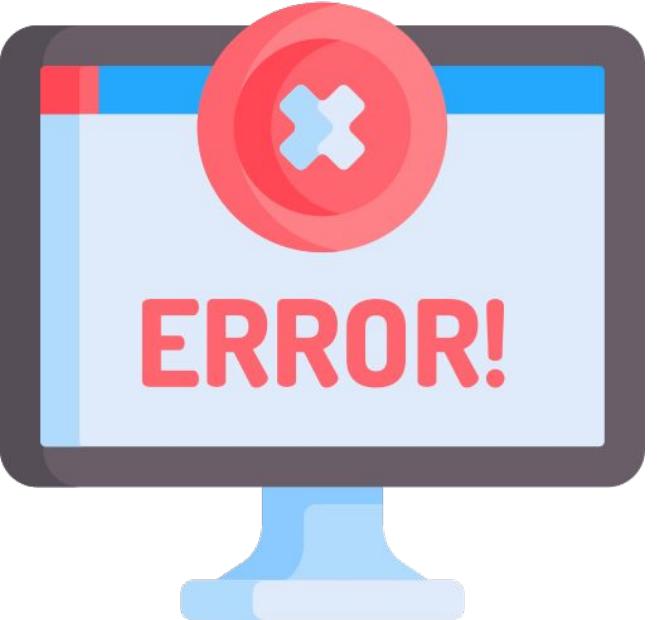
$$T \sim N(m, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0,5(\frac{t-m}{\sigma})^2}$$

$$T \sim N(320, 9) = \frac{1}{9\sqrt{2\pi}} e^{-0,5(\frac{t-320}{\sigma})^2}$$

Calculer la probabilité

$$P(t \leq 335) =$$

$$T \sim N(320, 9)$$



Calculer la probabilité

$$P(t \leq 335) =$$

Elle doit suivre une loi
normale centrée et
réduite $\sim N(0, 1)$

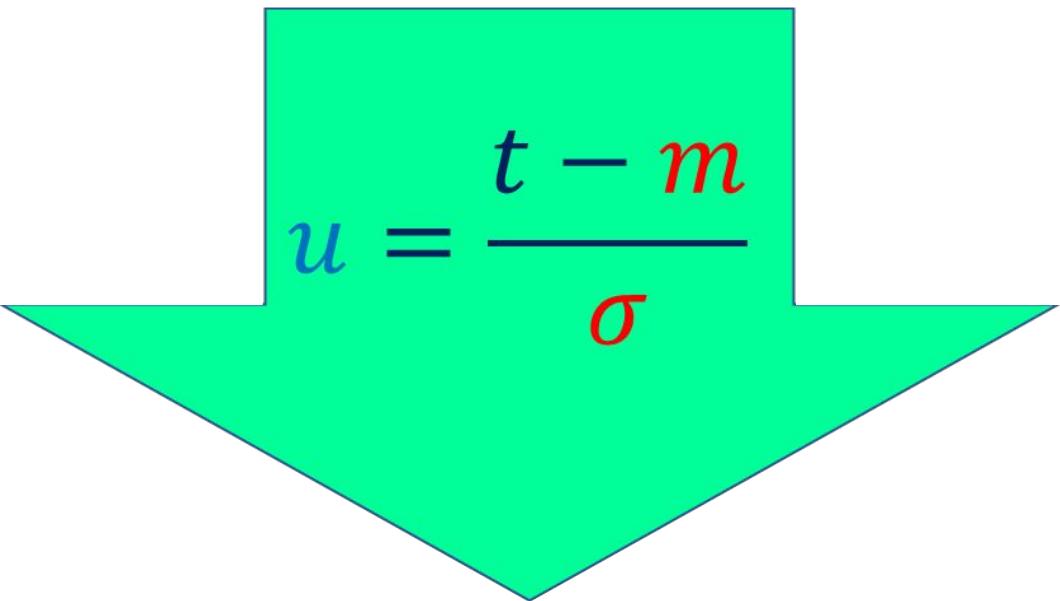
Solution

Calculer la probabilité

$$P(t \leq 335)$$

et

$$T \sim N(m, \sigma)$$


$$u = \frac{t - m}{\sigma}$$

$$P(u \leq \frac{335 - 320}{9})$$

et

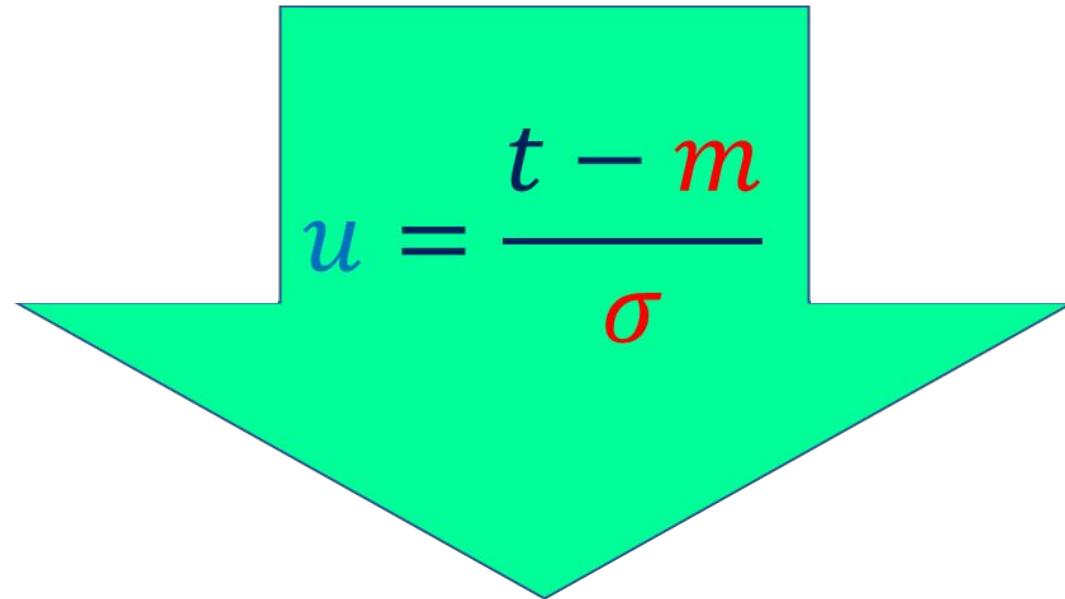
$$u \sim N(0, 1)$$

Calculer la probabilité

$$P(t \leq 335)$$

et

$$T \sim N(m, \sigma)$$


$$u = \frac{t - m}{\sigma}$$

$$P(u \leq 1,66)$$

et

$$u \sim N(0, 1)$$

$$P(u \leq 1,66) = 0,9515$$

Comment se fait la lecture inverse

Lecture inverse

Si nous connaissons le résultat de la probabilité, peut-on connaître la valeur a ?

Dans le cas où vous ignorez l'abscisse et vous connaissez la probabilité

$$P(x \leq a) = \text{connu} \quad x \sim N(0,1)$$

$$P(x \leq a) = b$$

$$x \sim N(0,1)$$

$$b \geq 0,5$$



$$a = F^{-1}(b)$$

$$P(x \leq a) = b$$

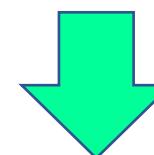
$$x \sim N(0,1)$$

$$b \geq 0,5$$

$$b \leq 0,5$$

$$a = F^{-1}(b)$$

$$a = -F^{-1}(1 - b)$$



Exemple de la lecture inverse b est supérieur à 0,5

$$P(u \leq a) = 0,9115$$

$$P(u \leq a) = 0,9115$$

$$u \sim N(0,1)$$

$$P(u \leq a) = 0,9115$$

$$u \sim N(0,1)$$

- On cherche le a

$$P(u \leq a) = 0,9115$$

$$a = F^{-1}(0,9115)$$

$$a = 0,05 + 1,3$$

$$a = 1,35$$

$$P(u \leq a) = 0,9493$$

$$a = F^{-1}(0,9493)$$

$$P(u \leq a) = 0,9493$$

$$a = F^{-1}(0,9493)$$

0,9493 n'est pas dans la table. On prend la valeur la plus proche à gauche

$$a = 0,03 + 1,6$$

$$a = 1,63$$

$$P(u \leq a) = 0,9493$$

$$a = F^{-1}(0,9493)$$

0,9493 n'est pas dans la table. On prend la valeur la plus proche à gauche

$$a = 0,03 + 1,6$$

$$a = 1,63$$

Exemple de la lecture inverse b est inférieur à 0,5

$$P(u \leq a) = 0,025$$

Exemple de la lecture inverse b est inférieur à 0,5

$$P(u \leq a) = 0,025$$

$$a = -F^{-1}(1 - 0,025)$$

$$a = -F^{-1}(0,975)$$

Exemple de la lecture inverse b est inférieur à 0,5

$$P(u \leq a) = 0,025$$

$$a = -F^{-1}(1 - 0,025)$$

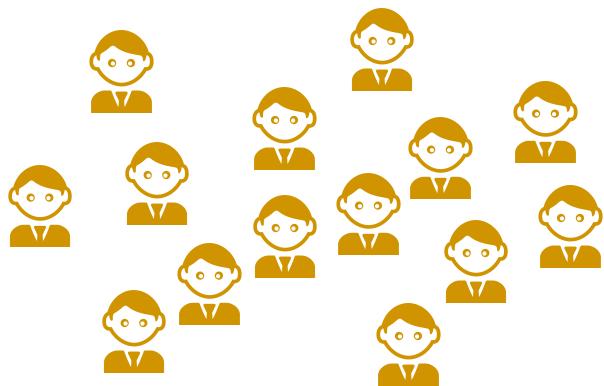
$$a = -F^{-1}(0,975)$$

$$a = -(0,06 + 1,9)$$

$$a = -1,96$$

Le théorème central limite

Supposons qu'un chercheur sélectionne un échantillon de 30 hommes adultes et trouve que la moyenne de la mesure des niveaux de triglycérides pour les sujets de l'échantillon est de 187 milligrammes/déclitre.

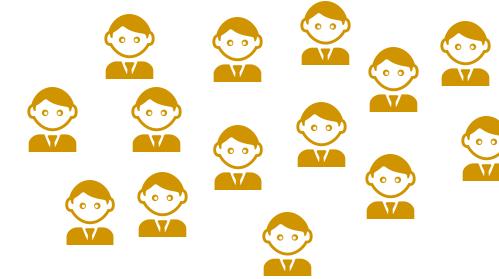


Supposons ensuite qu'un deuxième échantillon de 50 hommes soit sélectionné et que la moyenne de cet échantillon soit de 192 milligrammes/décilitre.

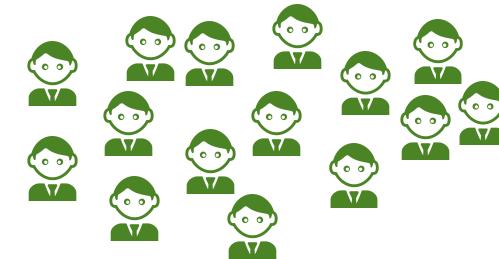
Continuer le processus pour 100 échantillons.



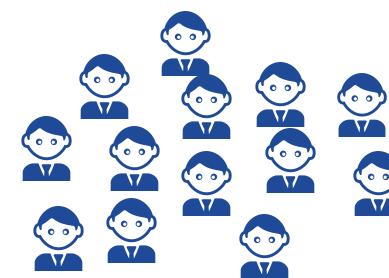
Continuer le processus pour
100 échantillons.



Ech 1

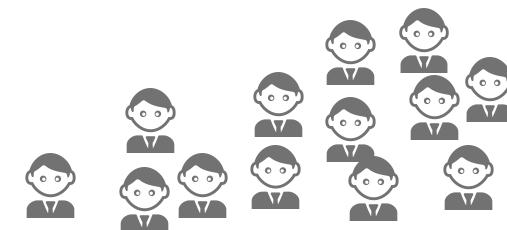


Ech 2



Ech 3

..
..
..



Ech 100

Ce qui se passe alors, c'est que la moyenne devient une variable aléatoire, et l'échantillon signifie 187, 192, 184, . . . , 196 constituent *une distribution d'échantillonnage des moyennes.*

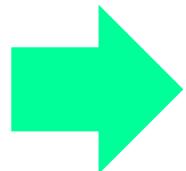
Propriétés de la distribution des moyennes d'échantillon

- La moyenne des moyennes de l'échantillon sera la même que la moyenne de la population.
- L'écart type des moyennes de l'échantillon sera égal à l'écart type de la population divisé par la racine carrée de la taille de l'échantillon.

Supposons que les quatre étudiants constituent la population. La moyenne de la population (2,6,4 et 8) est de $m = 5$ et un $\sigma = 2,23$

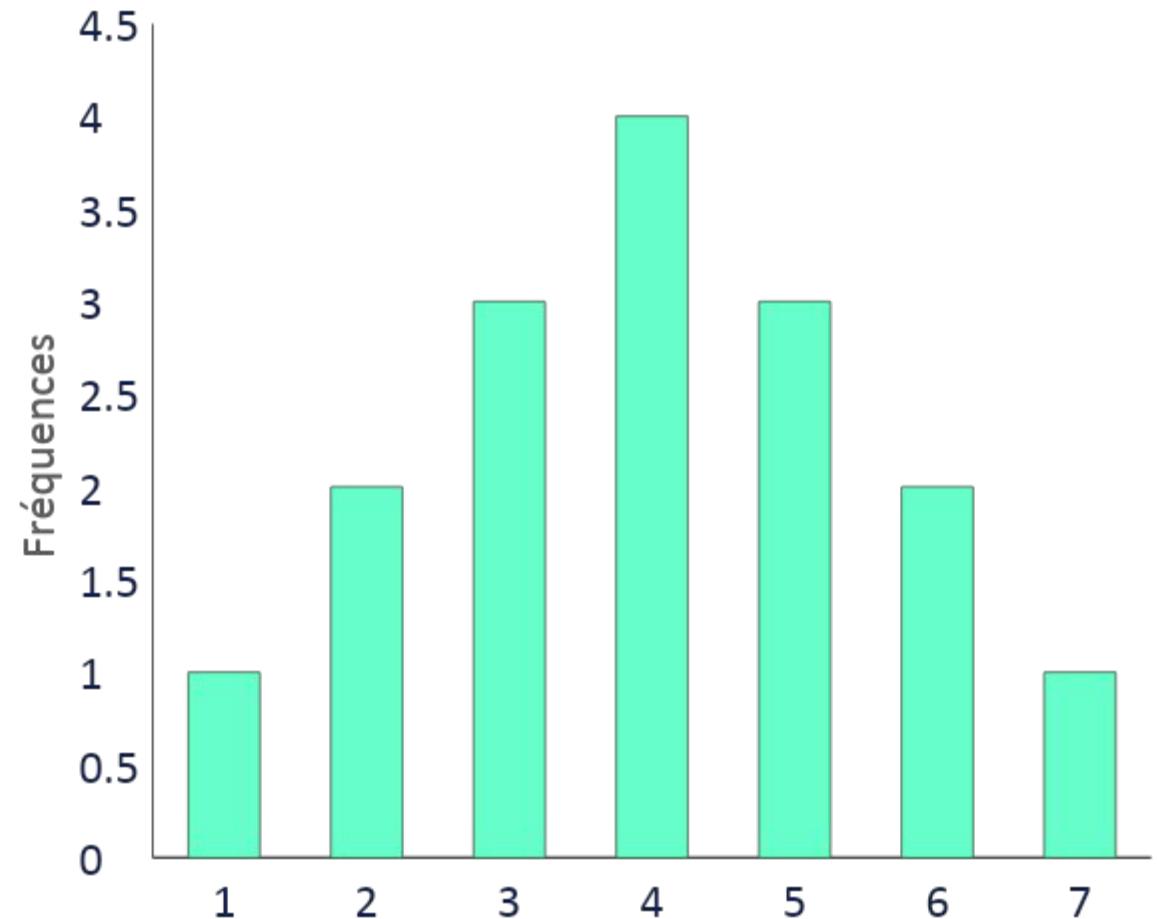
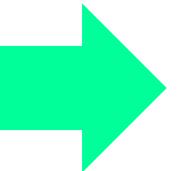
Echantillons	2;2	2	6;2	4
	2;4	3	6;4	5
	2;6	4	6;6	6
	2;8	5	6;8	7
	4;2	3	8;2	5
	4;4	4	8;4	6
	4;6	5	8;6	7
	4;8	6	8;8	8

Echantillons	2;2	2	6;2	4
	2;4	3	6;4	5
	2;6	4	6;6	6
	2;8	5	6;8	7
	4;2	3	8;2	5
	4;4	4	8;4	6
	4;6	5	8;6	7
	4;8	6	8;8	8



Moyenne	Fréquence
8	1
7	2
6	3
5	4
4	3
3	2
2	1

Moyenne	Fréquence
8	1
7	2
6	3
5	4
4	3
3	2
2	1



Echantillons	2;2	2	6;2	4
	2;4	3	6;4	5
	2;6	4	6;6	6
	2;8	5	6;8	7
	4;2	3	8;2	5
	4;4	4	8;4	6
	4;6	5	8;6	7
	4;8	6	8;8	8

$$m_{\bar{x}} = \frac{2 + 3 + \dots + 8}{16} = 5$$

$$m_{\bar{x}} = \frac{2 + 3 + \dots + 8}{16} = 5$$

$$m_{\bar{x}} = m$$

Echantillons	2;2	2	6;2	4
	2;4	3	6;4	5
	2;6	4	6;6	6
	2;8	5	6;8	7
	4;2	3	8;2	5
	4;4	4	8;4	6
	4;6	5	8;6	7
	4;8	6	8;8	8

$$\sigma_{\bar{x}} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(2-5)^2 + \dots + (8-5)^2}{15}} = 1,58$$

$$\sigma_{\bar{x}} = 1,58$$

$$\sigma = 2,23 = \sigma_{\bar{x}} \times 2 = 1,58 \times 2$$

$$\sigma_{\bar{x}} = \frac{\sigma}{2}$$

Le théorème central limite

Au fur et à mesure que la taille de l'échantillon n augmente sans limite, la forme de la distribution des moyennes d'échantillon prises avec remise à partir d'une population de moyenne m et d'écart type σ se rapprochera d'une distribution normale.

Le théorème central limite

Au fur et à mesure que la taille de l'échantillon n augmente sans limite, la forme de la distribution des moyennes d'échantillon prises avec remise à partir d'une population de moyenne m et d'écart type σ se rapprochera d'une distribution normale.

$$n \geq 30$$

Le théorème central limite

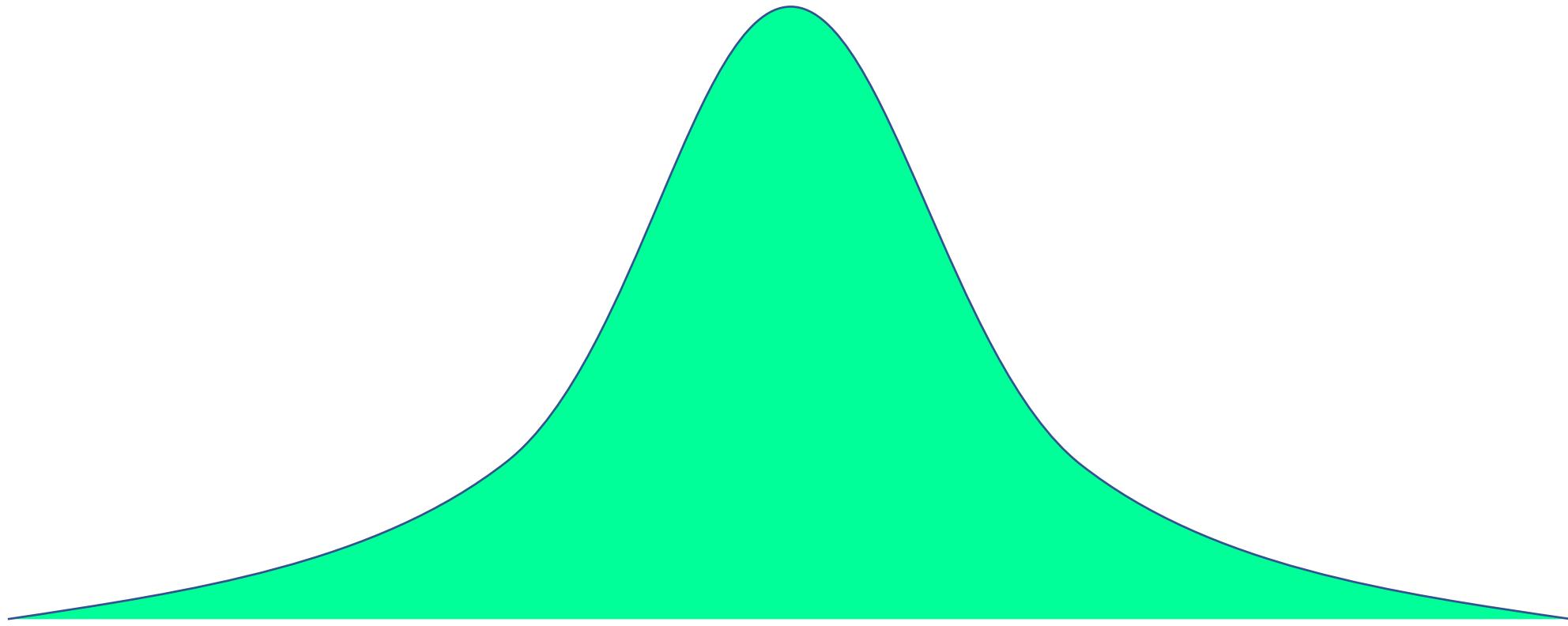
Au fur et à mesure que la taille de l'échantillon n augmente sans limite, la forme de la distribution des moyennes d'échantillon prises avec remise à partir d'une population de moyenne m et d'écart type σ se rapprochera d'une distribution normale.

$$n \geq 30$$
$$\bar{x} \sim N\left(m; \frac{\sigma}{\sqrt{n}}\right)$$

Un médecin peut déterminer si les statistiques vitales d'un patient se situent dans l'intervalle normal ou si un certain type de traitement est nécessaire pour corriger une condition et éviter de futures maladies.

Le théorème central limite

De nombreuses **variables mesurées** dans les tests médicaux - **tension artérielle**, taux de **triglycérides**, etc. - sont à peu près **normalement distribuées** pour la majorité de la population.



Ainsi, les chercheurs peuvent trouver la **moyenne** et l'**écart type** de ces variables.

Ensuite, en utilisant ces deux mesures **avec les valeurs z**, ils peuvent trouver des **intervalles normaux** pour les individus en bonne santé.

Par exemple, 95 % des pressions artérielles systoliques d'individus en bonne santé se situent à moins de 2 écarts-types de la moyenne.