

Introduction aux Statistiques Variables aléatoires

L1 STE

- **Variables aléatoires discrètes**
 - Loi de probabilité
 - Fonction de répartition
 - Espérance
 - Variance

- **Variables aléatoires continues**
 - Fonction de répartition
 - Fonction de densité
 - Espérance
 - Variance

Variable aléatoire discrète

Un petit jeu de dés

$$P(X = 1) = 1/6$$

$$P(X = -10) = 5/6$$

pas très favorable à vue de nez...

Loi de probabilité

La **loi de probabilité** (ou distribution ou fonction de densité) décrit les répartitions des fréquences d'apparition des résultats d'une expérience aléatoire.

$$f(x_i) = P(X = x_i)$$

$$\sum_{i=1}^n f(x_i) = 1$$

Fonction de répartition

La fonction de répartition correspond à la distribution cumulée que nous avons dans le cours sur les représentation graphique des données.

$$F(x_i) = P(X \leq x_i)$$

Espérance

L'espérance, notée $E(x)$ correspond à une moyenne pondérée:

$$E(X) = \sum_{i=1}^n x_i f(x_i)$$

Variance, $V(X)$

$$V(X) = E \left\{ [X - E(X)]^2 \right\} = E(X^2) - [E(X)]^2$$

Il s'agit de rappels. Si cela n'est pas clair, les TD vont préciser les choses.

Fonction de répartition

C'est la même chose que pour les variables aléatoires discrètes, excepté que x appartient à \mathbb{R} cette fois.

$$\forall x \in \mathbb{R}, F_X(x) = P(X \leq x)$$

Propriétés intéressantes...

Propriété 1 : $F_X(x)$ est croissante

Propriété 2 :

$$\lim_{x \rightarrow +\infty} F_X(x) = 1$$

Propriété 3 :

$$\lim_{x \rightarrow -\infty} F_X(x) = 0$$

$$\forall a \in \mathbb{R}, P(X \leq a) = F_X(a) = \int_{-\infty}^a f(t) dt$$

Fonction de densité

C'est la même chose que pour les variables aléatoires discrètes, excepté que x appartient à \mathbb{R} . Il s'agit de $f(x)$ dans la formule précédente.

Propriétés intéressantes...

$$\forall x \in \mathbb{R}, f(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f(t) dt = 1$$

Espérance

En suivant exactement la même construction que pour les variables aléatoires discrètes, on définit l'espérance:

$$E(X) = \int_{-\infty}^{+\infty} t f(t) dt$$

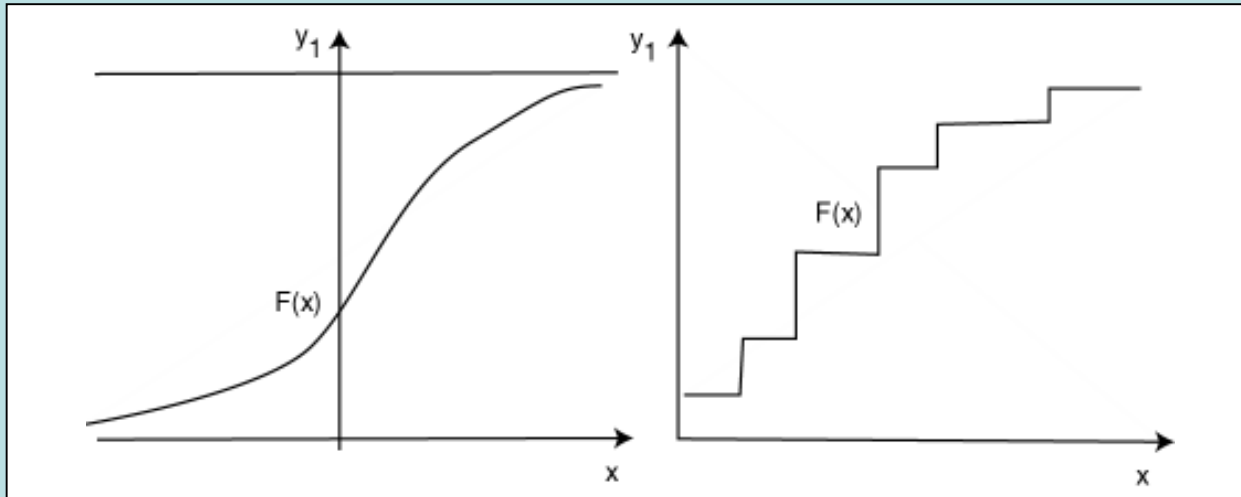
$$E(X^2) = \int_{-\infty}^{+\infty} t^2 f(t) dt$$

Variance, $V(X)$:

Idem pour la variance

$$V(X) = E \left\{ [X - E(X)]^2 \right\} = E(X^2) - [E(X)]^2$$

Une fonction de répartition d'une variable aléatoire x est une fonction $F(x)$ qui, pour tout x , indique la probabilité pour que x soit inférieur ou égal à x_1 .



Variable continue (sigmoïde
pour La loi normale)

Variable discrète (escalier)

Introduction aux Statistiques Lois classiques

L1 STE

- **Lois discrètes**
 - Loi binomiale
 - Loi de Poisson
- **Loi continue**
 - Loi normale

La loi binomiale:

C'est une distribution discontinue qui donne les probabilités de voir apparaître un événement de probabilité p respectivement $0, 1, 2, 3, \dots, n$ fois au cours de n épreuves.

Seulement deux évènements peuvent apparaître.

$B(n,p)$: n est le nombre d'épreuves, p est la probabilité d'un des deux évènements (p.e. succès), q est la probabilité complémentaire (p.e. échec).

La probabilité de voir apparaître x fois le même événement de probabilité p au cours de n épreuves indépendantes peut s'écrire:

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

Exemple:

Une famille de n enfants, quelle est la probabilité d'avoir x garçons?

$$B(2,0.5)?$$

$$B(7,0.5)?$$

Avec $n=2$: $B(2,0.5)$

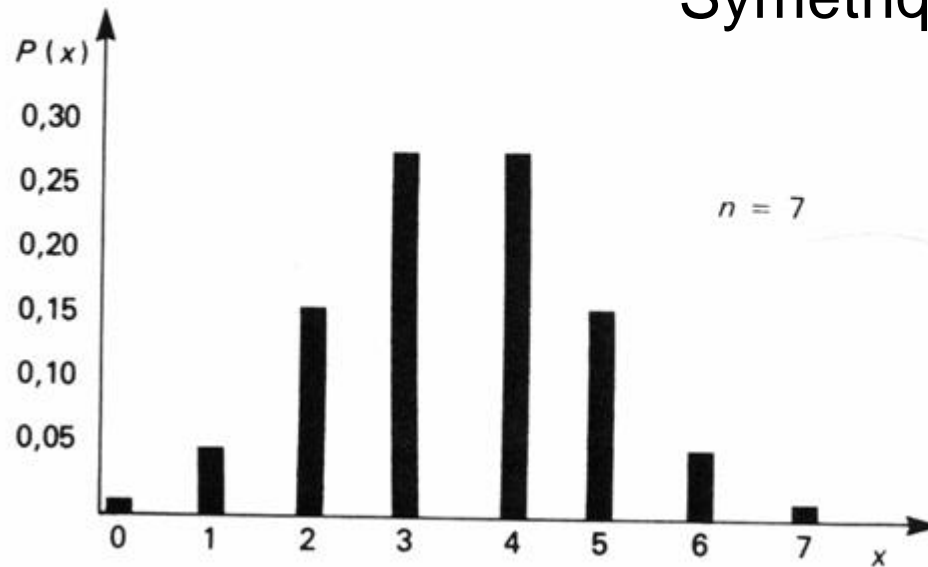
x	P(x)
0	0.25
1	0.5
2	0.25

Distribution de probabilité : $B(7, 0,5)$

x	P(x)
0	$\frac{7!}{0!(7-0)!} 0,5^7 \cdot 0,5^0 = 0,0078$
1	$\frac{7!}{1!(7-1)!} 0,5^6 \cdot 0,5^1 = 0,0546$
2	$\frac{7!}{2!(7-2)!} 0,5^5 \cdot 0,5^2 = 0,1640$
3	$\frac{7!}{3!(7-3)!} 0,5^4 \cdot 0,5^3 = 0,2734$
4	$\frac{7!}{4!(7-4)!} 0,5^3 \cdot 0,5^4 = 0,2734$
5	$\frac{7!}{5!(7-5)!} 0,5^2 \cdot 0,5^5 = 0,1640$
6	$\frac{7!}{6!(7-6)!} 0,5^1 \cdot 0,5^6 = 0,0546$
7	$\frac{7!}{7!(7-7)!} 0,5^0 \cdot 0,5^7 = 0,0078$

Distribution de probabilité du nombre de garçons (x) dans une famille de sept enfants – $B(7, 0,5)$

Symétrique!!



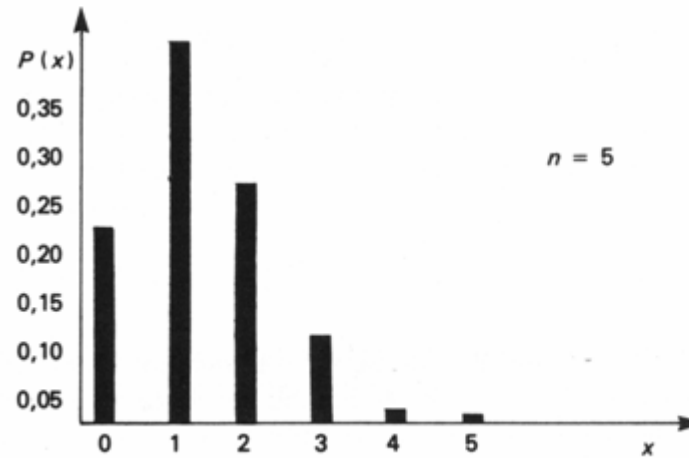
Des rats sont conditionnés. Un passage a 25% d'être emprunté. 5 essais...

$B(5,0.25)???$

Distribution de probabilité : $B(5, 0,25)$

x	$P(x)$
0	$\frac{5!}{(5-0)!0!} 0,75^5 = 0,2373$
1	$\frac{5!}{(5-1)!1!} 0,75^4 \cdot 0,25 = 0,3955$
2	$\frac{5!}{(5-2)!2!} 0,75^3 \cdot 0,25^2 = 0,2636$
3	$\frac{5!}{(5-3)!3!} 0,75^2 \cdot 0,25^3 = 0,0879$
4	$\frac{5!}{(5-4)!4!} 0,75 \cdot 0,25^4 = 0,0146$
5	$\frac{5!}{(5-5)!5!} 0,25^5 = 0,0009$

Distribution de probabilité du nombre de rats (x) empruntant la branche supérieure droite du labyrinthe en H $B(5, 0,25)$



Dissymétrique!!!!

Espérance mathématique

$$E(x) = \sum_{i=1}^n P(x_i) \cdot x_i$$

Pour la loi binomiale:

$$E(x) = np$$

Exemple:

Quelle est l'espérance mathématique du nombre de garçons dans une famille de 7 enfants?

$$E(x) = np = 7 \cdot 0.5 = 3.5$$

Pour une distribution binomiale:

$$\sigma^2 = npq$$

Exemple:

Quelle est la variance du nombre de garçons dans une famille de 7 enfants?

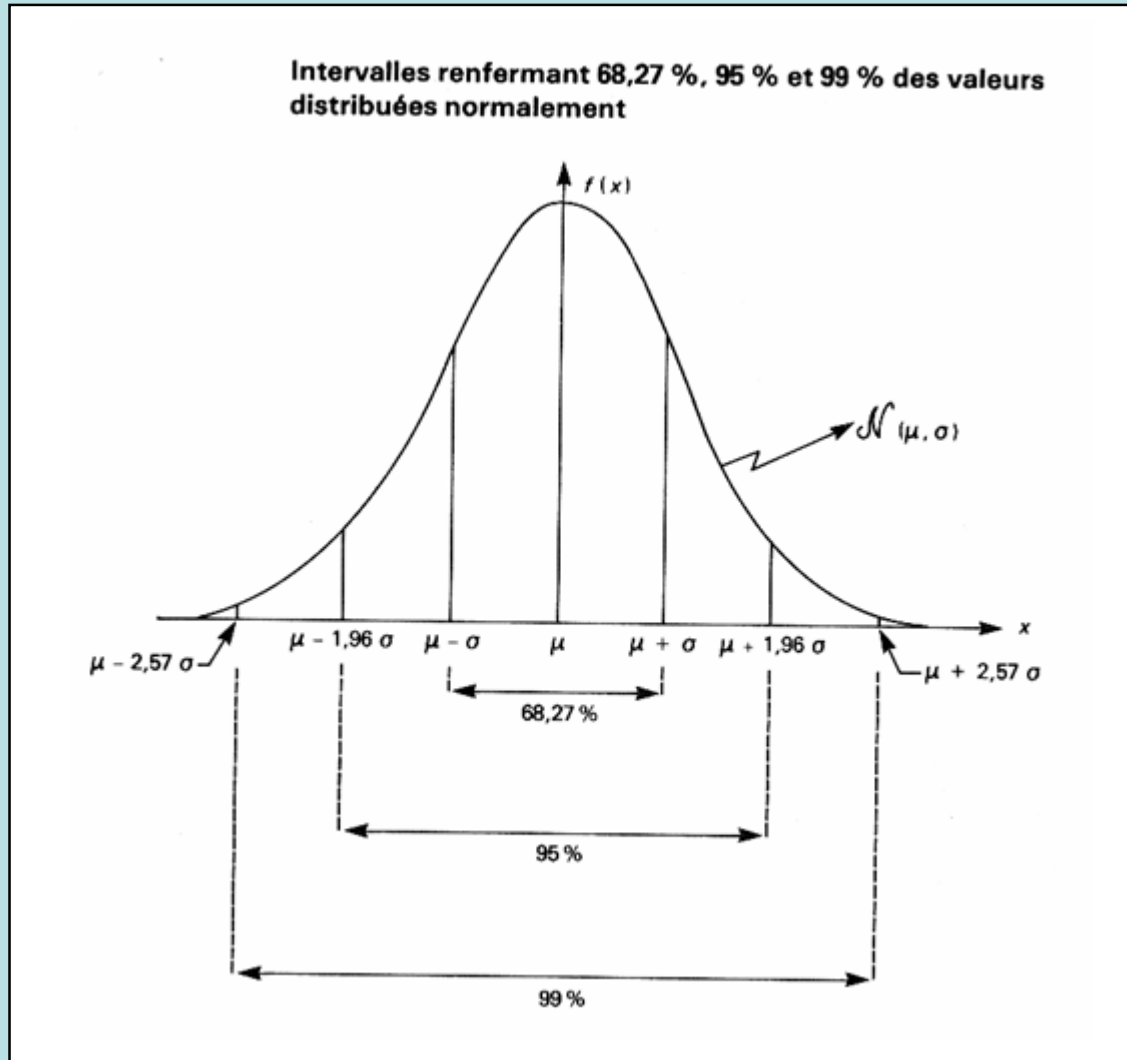
$$\sigma^2 = 7 \cdot 0.5 \cdot 0.5 = 1.75$$

Précurseurs : de Moivre (1733), forme actuelle :
Laplace (1712), Gauss (1809)

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

La variable x peut être continue, sans borne inf ou sup.

La loi normale: $N(\mu, \sigma)$



Espérance : $E(x) = \mu$

Variance : $\text{Var}(x) = \sigma^2$

Symétrique par rapport à μ

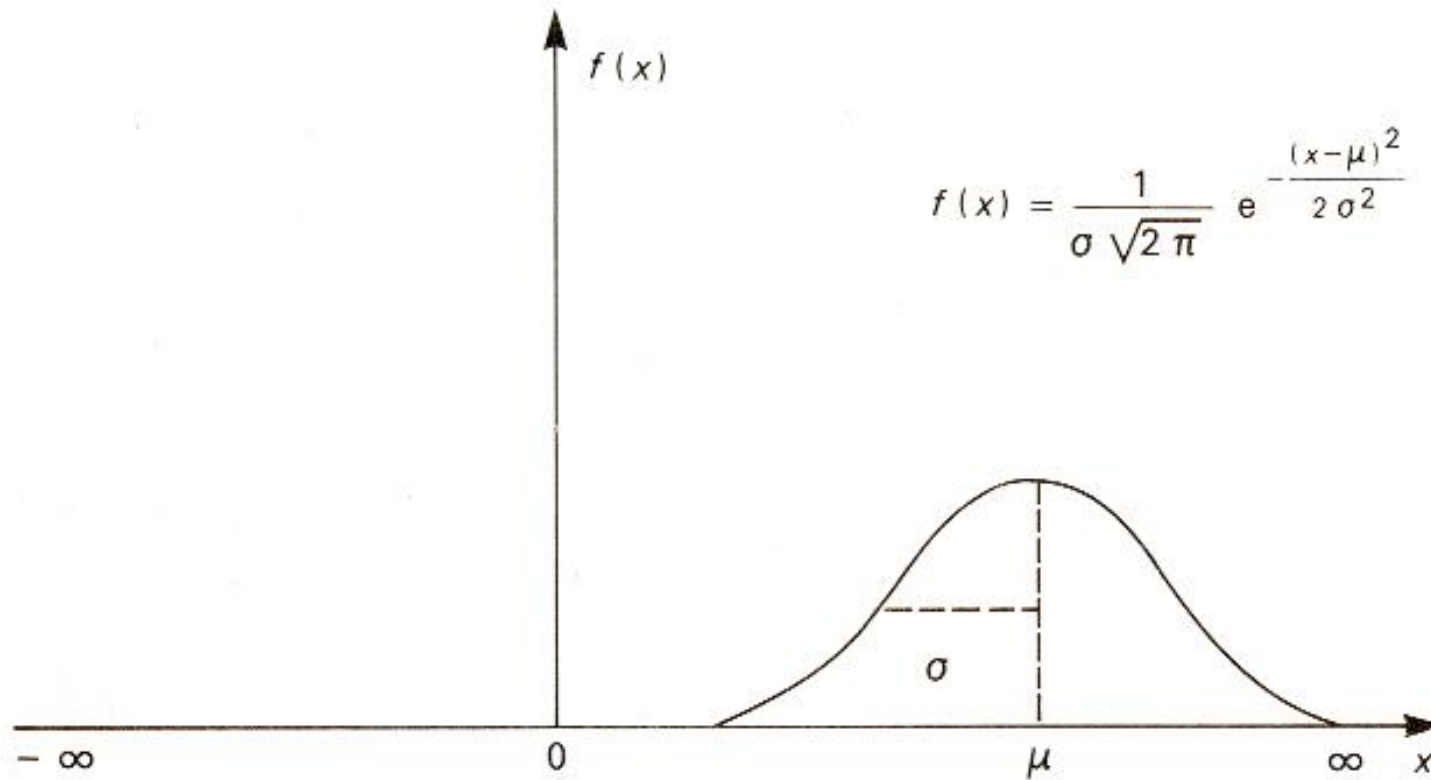
Loi normale centrée réduite: moyenne = 0, écart type = 1.

1^{ere} transformation : $X = x - \mu$

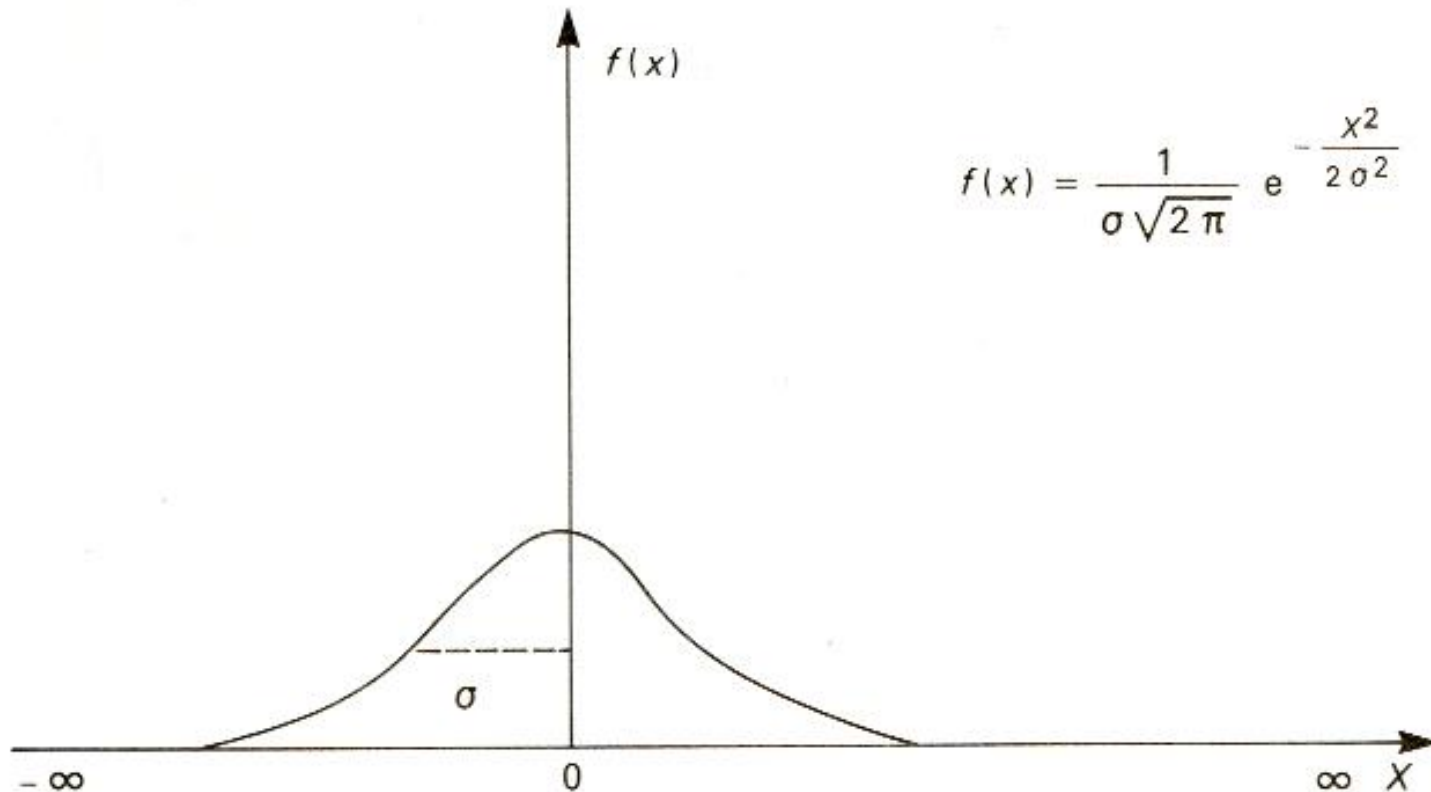
2^{eme} transformation : $Z = X/\sigma$

$$-\infty < Z < +\infty$$

Courbe normale : $\mathcal{N}(\mu, \sigma)$

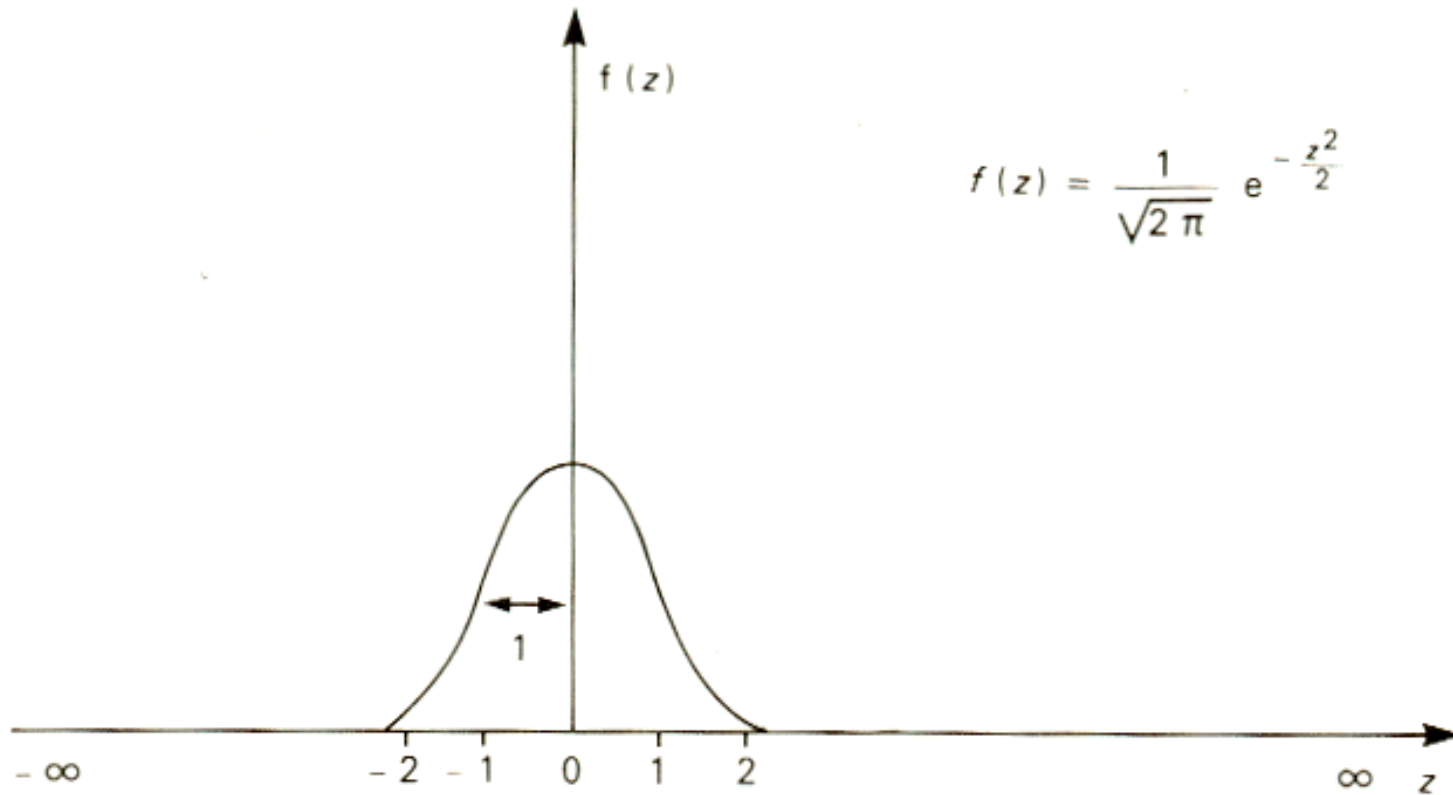


Courbe normale centrée : $\mathcal{N}(0, \sigma)$



On retire la moyenne (on centre sur 0)

Courbe normale centrée réduite : $\mathcal{N}(0, 1)$

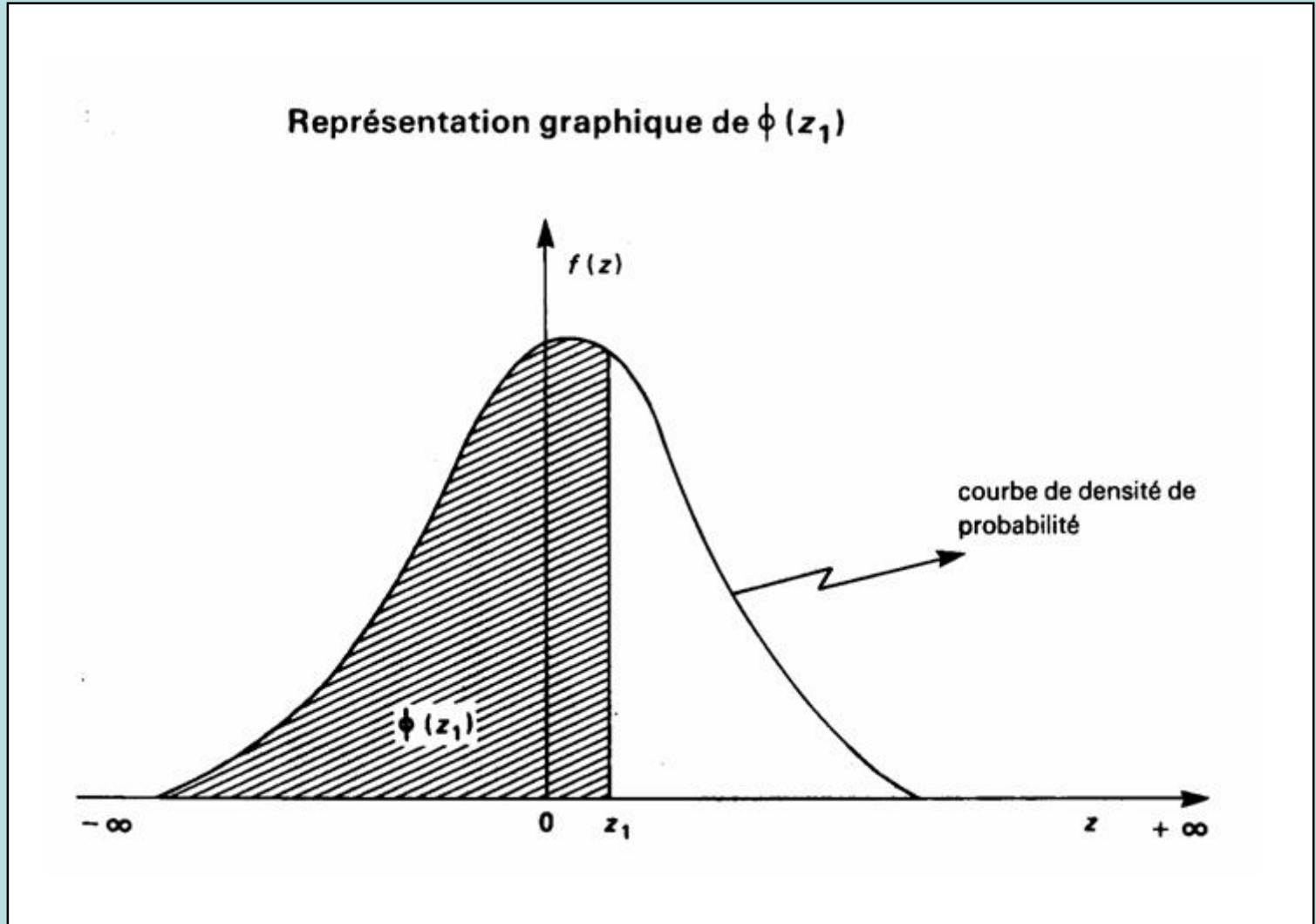


On divise par l'écart type (Nouvel écart type = 1)

L'aire totale comprise entre la courbe et l'axe des abscisses est égale à 1.

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}} dz = 1$$

Représentation graphique de $\Phi(z_1)$



Fonction de densité et fonction de répartition

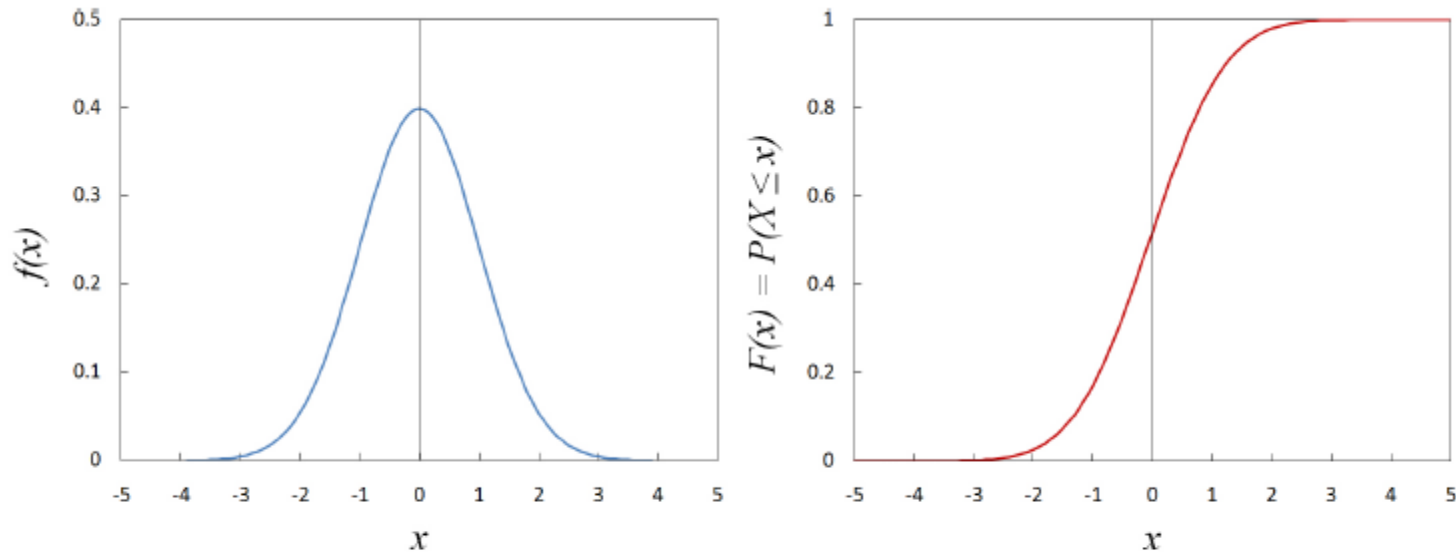


FIG. 5 – Loi de normale centrée réduite ($\mu = 0, \sigma = 1$) : fonction de densité à gauche, et fonction de répartition à droite.

TABLE III — AIRES LIMITÉES PAR LA COURBE NORMALE CENTRÉE RÉDUITE

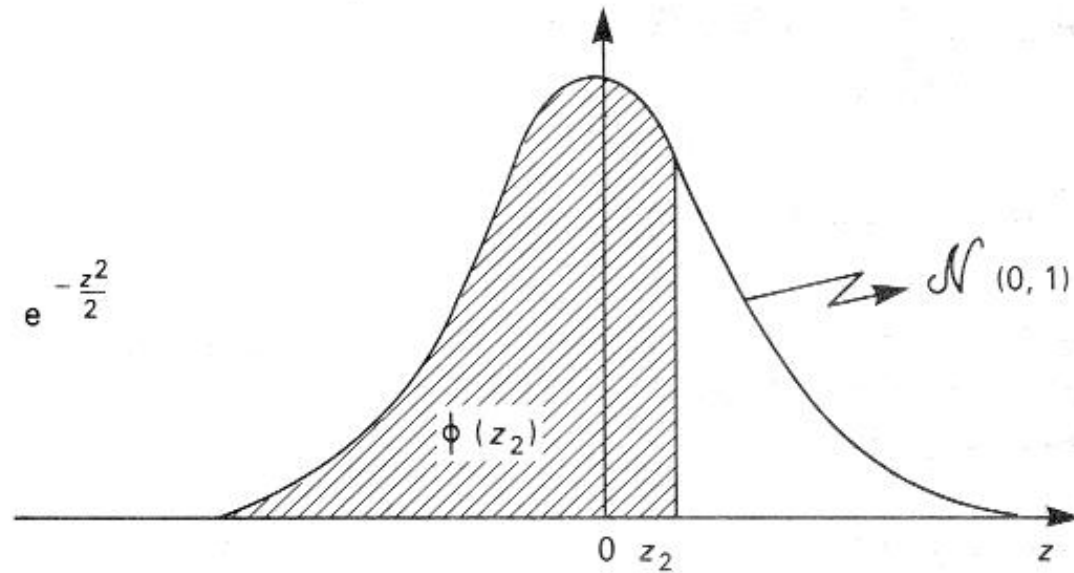
La table fournit les valeurs de $\phi(z)$ pour z positif. Lorsque z est négatif il faut calculer le complément à l'unité de la valeur lue dans la table. La première colonne indique la première décimale de z et la première rangée fournit la deuxième décimale.

Exemples : pour $z = 1,21$, $\phi(z) = 0,8869$ et pour $z = -1,21$, $\phi(z) = 0,1131$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	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	0,5910	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	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	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	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	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	0,8340	0,8365	0,8389
1,0	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	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	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	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
z	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
3	0,9987	0,9990	0,9993	0,9995	0,9997	0,9998	0,9998	0,9999	0,9999	1,0000
4	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

$$P(z < z_2)$$

$$\phi(z_2) = \int_{-\infty}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

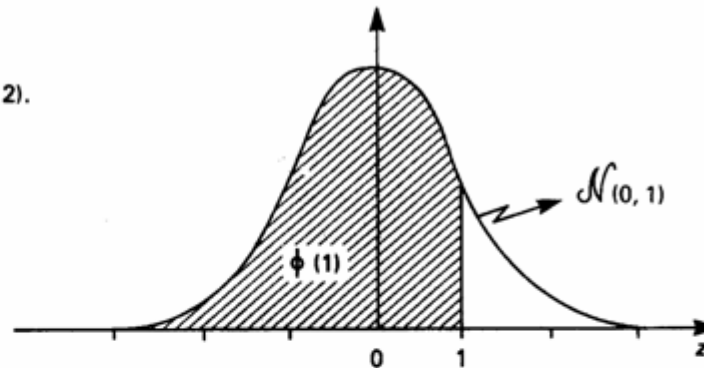


$$P(z_1 < z < z_2) = P(z < z_2) - P(z < z_1)$$

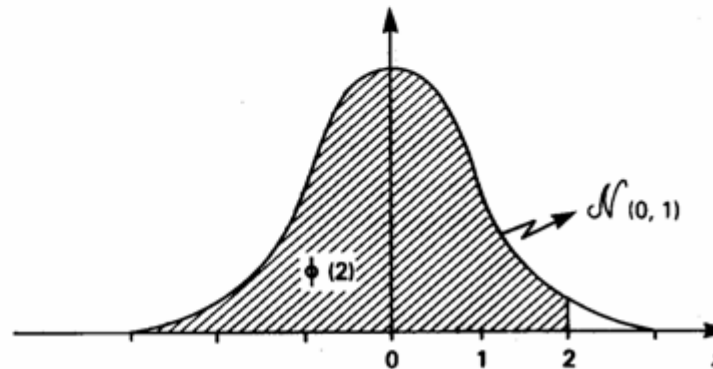
$$P(z_1 < z < z_2) = \phi(z_2) - \phi(z_1)$$

Représentation graphique du calcul de $P(1 < z < 2)$

Il s'agit de trouver $P(1 < z < 2)$.
Pour ce faire, cherchons
 $P(z < 1) = \Phi(1)$
 $\Phi(1)$ se lit dans la table III
pour $z = 1$
 $\Phi(1) = 0,841$



Puis
 $P(z < 2) = \Phi(2)$
qui se lit dans la
table III pour $z = 2$
 $\Phi(2) = 0,977$



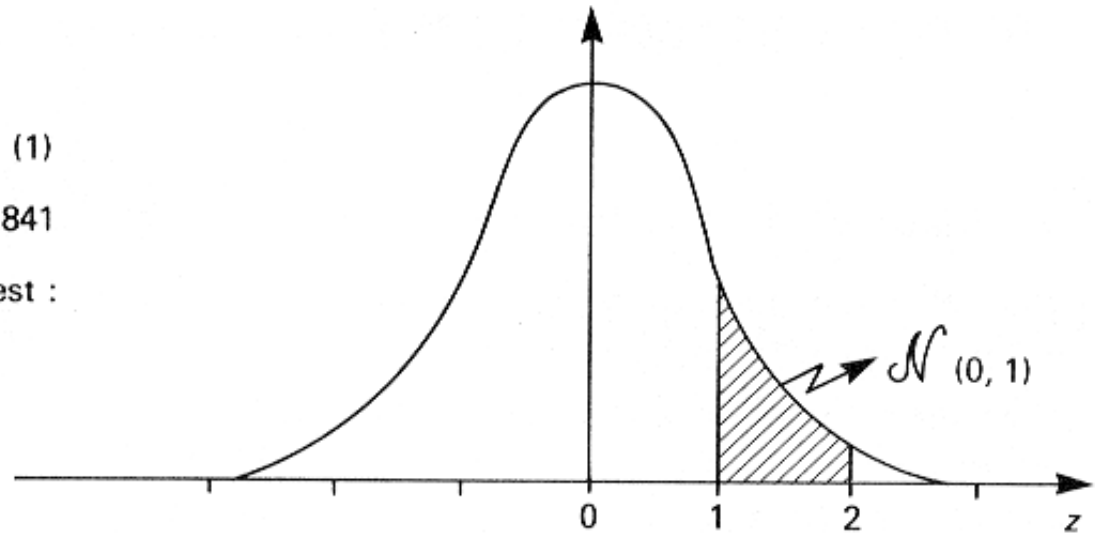
alors

$$P(1 < z < 2) = \Phi(2) - \Phi(1)$$

$$\Phi(2) - \Phi(1) = 0,977 - 0,841$$

La probabilité recherchée est :

$$P(1 < z < 2) = 0,136$$



Exemple: $N(5, 0.5)$, $P(4 < x < 6.5)$?

$$Z_1 = \frac{4-5}{0.5} = -2 \qquad Z_2 = \frac{6.5-5}{0.5} = 3$$

$P(-2 < Z < 3)$?

Réponse : calcul de $P(0 < Z < 3)$ et de $P(-2 < Z < 0)$ et addition des deux termes.

$$P(0 < Z < 3) = \Phi(3) - \Phi(0) = 0.999 - 0.5 = 0.499$$

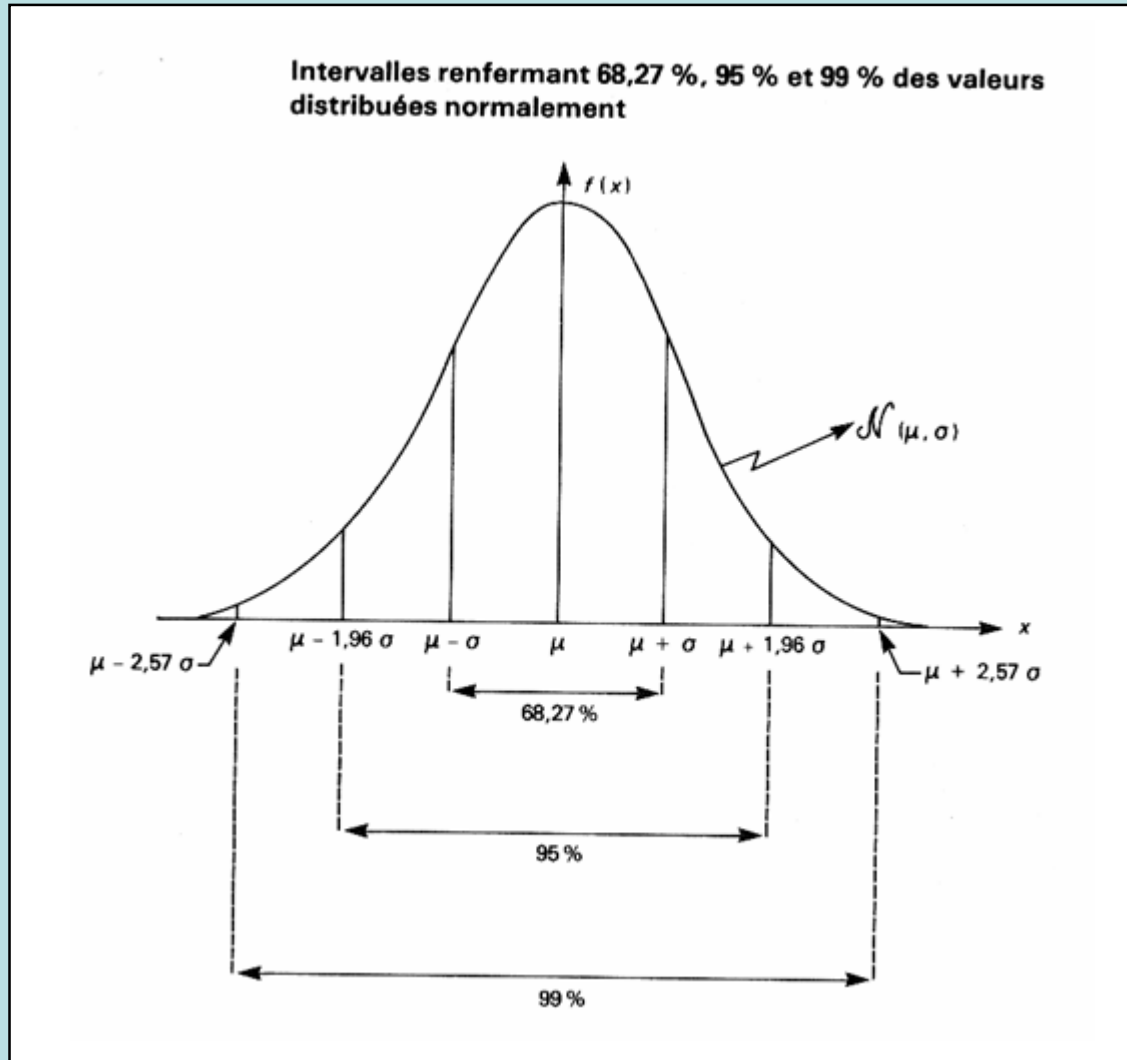
$$P(-2 < Z < 0) = \Phi(2) - \Phi(0) = 0.977 - 0.5 = 0.477$$

$$P(-2 < Z < 3) = 0.499 + 0.477 = \mathbf{0.976}$$

Caractéristiques de la courbe normale

- $f(Z)$ devient de + en + faible quand Z croît en valeur absolue.
- $f(Z)$ toujours > 0 .
- Courbe symétrique : $f(Z) = f(-Z)$
- Asymptote sur l'axe des x : $\exp(-z^2/2) \rightarrow 0$ quand $Z \rightarrow \text{infini}$
- La dérivée s'annule pour $Z=0$. Maximum à $Z=0$

La loi normale: $N(\mu, \sigma)$



Espérance : $E(x) = \mu$

Variance : $\text{Var}(x) = \sigma^2$

Symétrique par rapport à μ

Pour une **variable continue**, la probabilité de tomber exactement sur une valeur est nulle.

Qui mesure exactement 1,758942563 m?

On considère la toujours la probabilité pour que x soit contenue dans un intervalle ($P(x_1 \leq x \leq x_2)$)