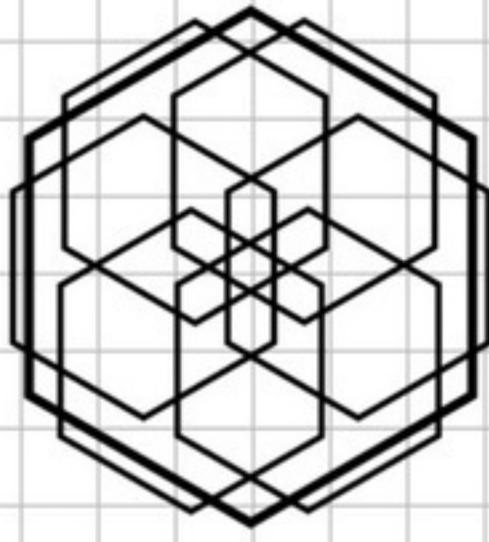




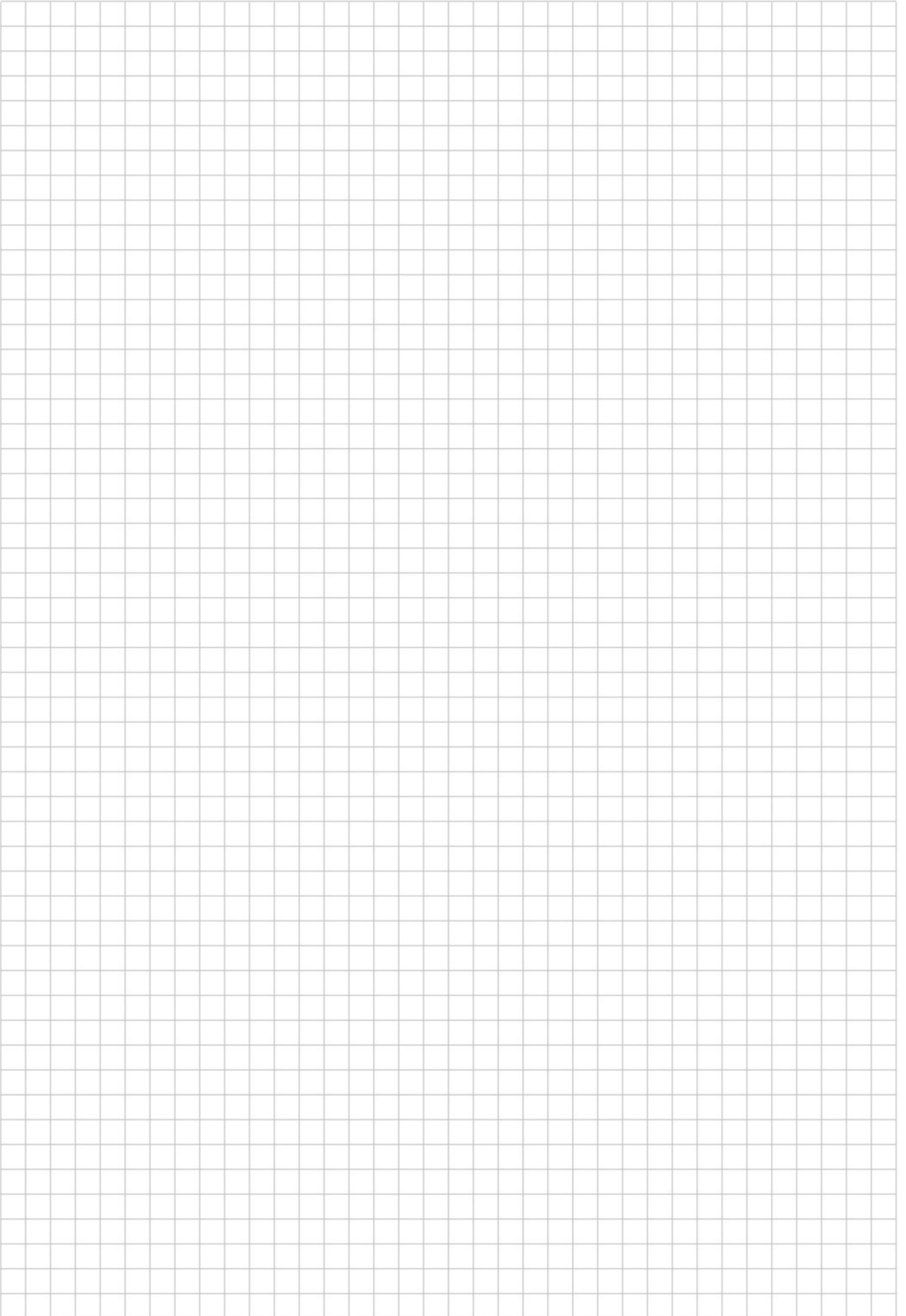
PrivateTeacher
Cours Privés de Science

La Loi Normale

*L'art de chiffrer
le hasard*

Sommaire:

Introduction
Pouvoir prédictif des statistiques
Distribution de probabilité
La distribution gaussienne
Forme mathématique
Terminologie et notation
Calcul de probabilité
Table des quantiles
Propriétés de symétrie
Dans la réalité
Centrer-Réduire
Exercices corrigés





Introduction

des statistiques sont l'art
de compter les événements.

Pourquoi donc compter ?

Pour pouvoir prédire !

Admettons que je m'intéresse de savoir
quelles sont mes chances de rencontrer
un chat noir dans une certaine région.

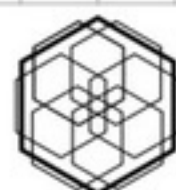
Je me promène donc, et je compte les chats

À la fin de la journée, j'ai croisé 14 chats
et 5 d'entre eux étaient noirs !

À l'aide de ces observations, je peux dire
maintenant que si vous vous promenez
dans cette même région, alors vous
avez vous aussi 5 chances sur 14
de rencontrer un chat noir.

Il s'agit bien là d'une prédiction.

Le pouvoir prédictif des statistiques
cependant, est bien plus grand.





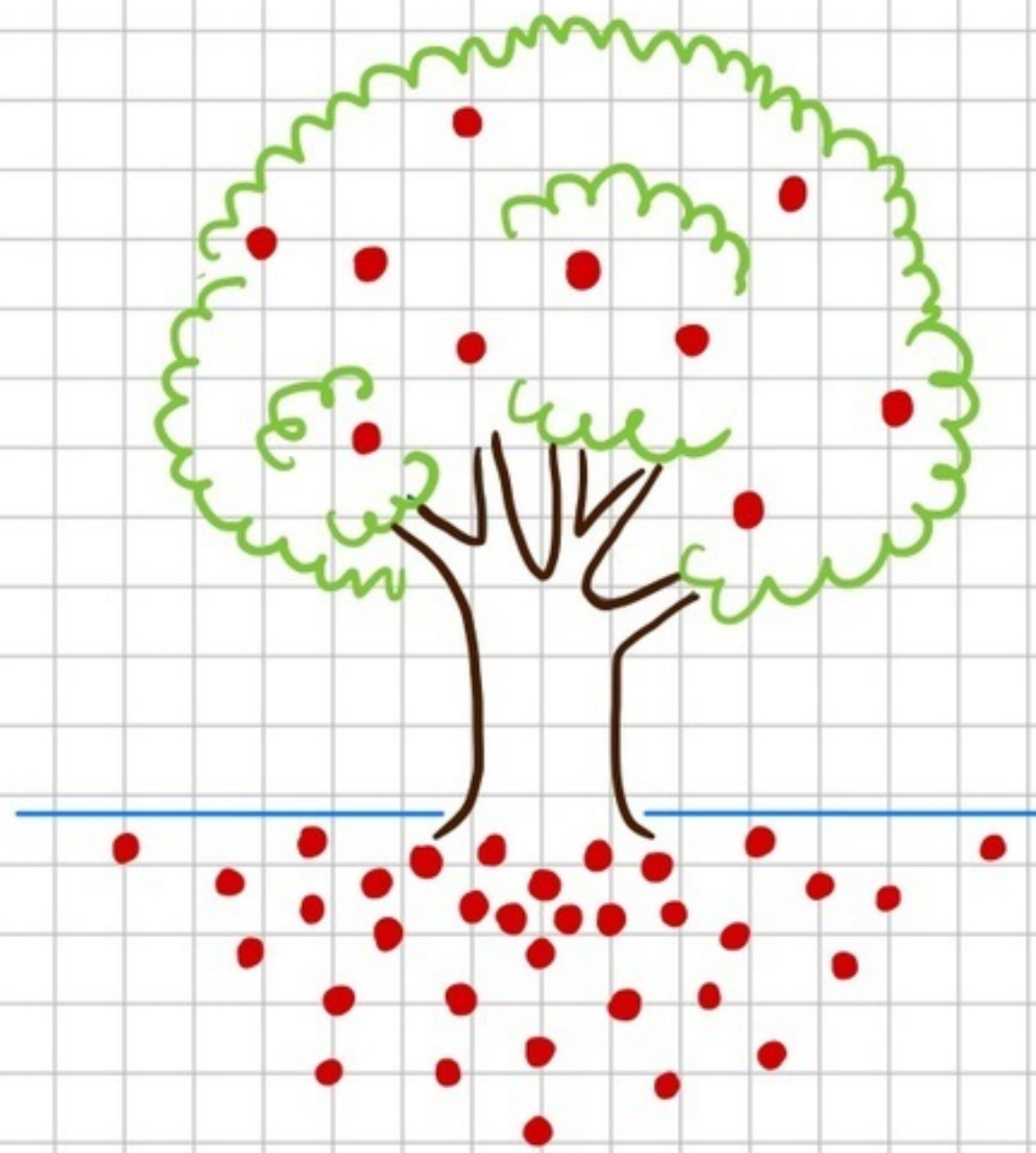
Le pouvoir prédictif des statistiques

Observons des pommes tomber d'un arbre.

Qui peut prévoir à quel endroit précis tombera la prochaine pomme ?

Réponse : Personne !

Une fois que toutes les pommes sont tombées pourtant, on observe qu'elles se répartissent autour de l'arbre de la manière suivante :

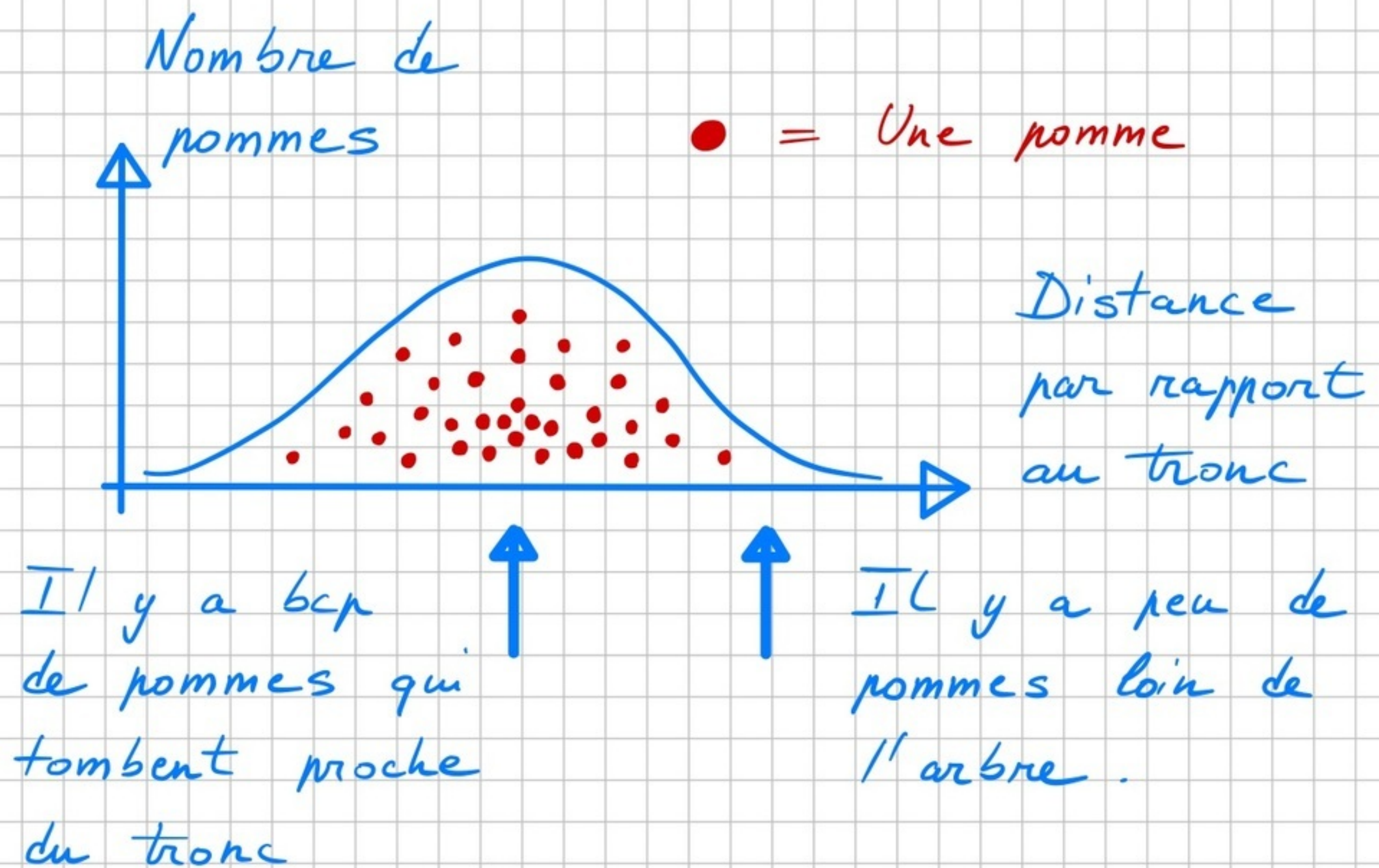


Il y a beaucoup de pommes proches du tronc, et il y en a de moins en moins à mesure que l'on s'éloigne de l'arbre.





On peut aussi représenter la distribution des pommes graphiquement :



Si l'on observe un autre pommier
aucunes pommes ne tombera
au même endroit.

Pourtant, la répartition des pommes
autours du tronc sera la même !

IL y a là ce que l'on appelle
une régularité statistique

C'est de cette régularité
que les statistiques
tirent leur pouvoir prédictif





Distribution de probabilité

Nous avons vu comment le décompte d'événements donne accès à la probabilité de l'observer.

(5 chats noirs sur 14 chats rencontré permet de calculer la probabilité :

$$\frac{5}{14} = 0.357$$

$$= 35.7 \% \text{ de chance}$$

de rencontrer un chat noir)

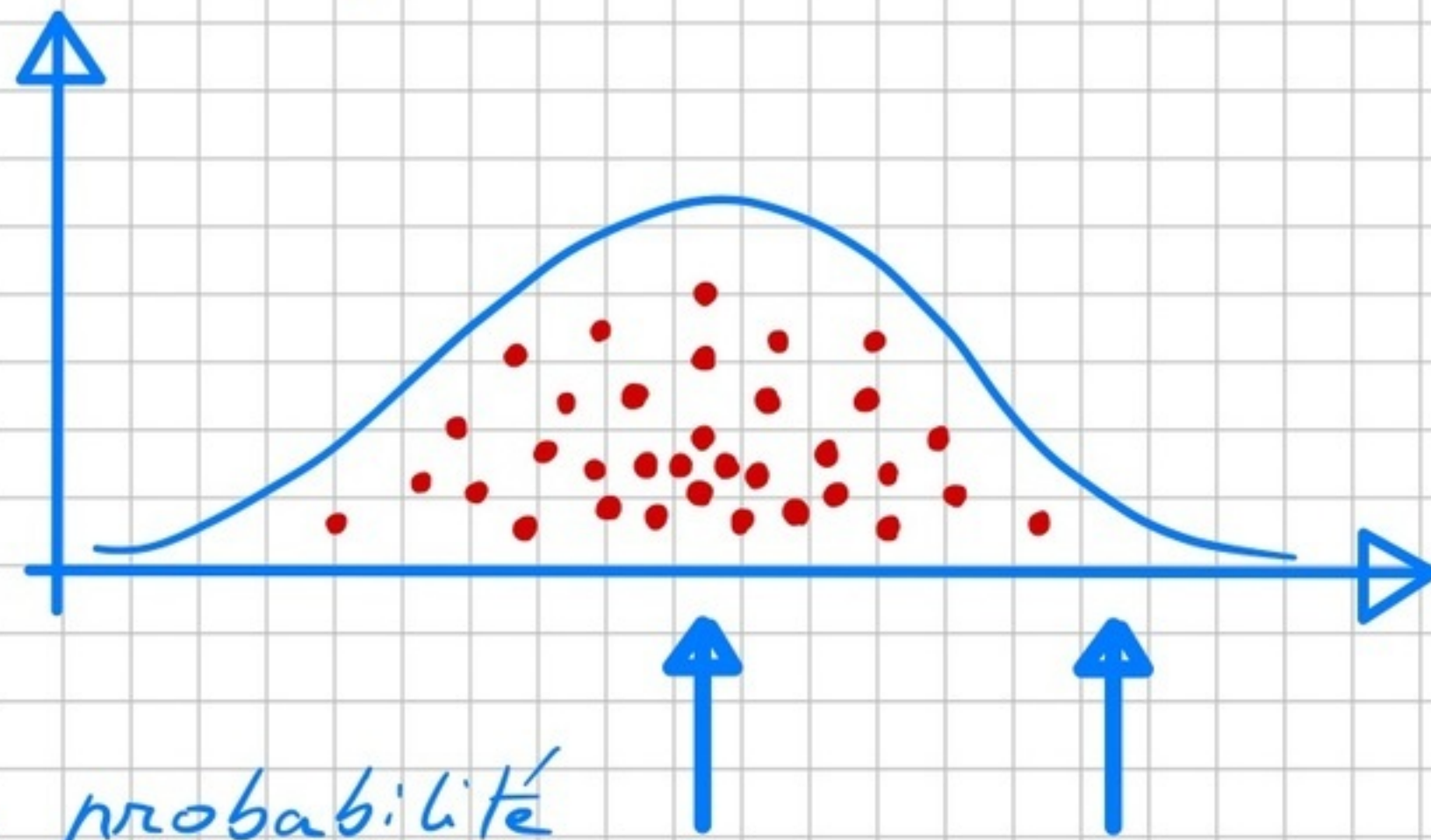
On le voit, il suffit de diviser le nombre d'événements qui nous intéresse par le nombre total d'événements

Si l'on dessine maintenant le même graphique que tout à l'heure avec cette fois-ci la probabilité de voir tomber une pomme à une certaine distance de l'arbre plutôt que simplement le nombre de pommes à cet endroit, on a ce que l'on appelle une distribution de probabilité.





Probabilité de voir
une pomme tomber



Distance
par rapport
au tronc

La probabilité
de voir tomber
une pomme
proche du tronc
est importante

Il est très peu
probable de trouver
une pomme loin
de l'arbre.

Une distribution de probabilité
permet donc de se représenter
un ensemble d'événements

Il existe plusieurs types d'événements
différents. Pour cette raison, il existe
plusieurs familles de distributions.





La distribution gaussienne

La distribution gaussienne est une distribution de probabilité qui apparaît lorsque l'on observe des événements aléatoires et indépendants les uns des autres.

Aléatoire, cela signifie que la réalisation d'un événement est possible, mais non certaine.

(la probabilité est une mesure du caractère aléatoire d'un événement)

Indépendant, cela signifie : sans aucun lien les uns avec les autres

(si une pomme tombe à droite, rien ne permet de prédire ou tombera la pomme suivante)

Attention, tous les phénomènes aléatoires dont les résultats sont indépendants les uns des autres ne se distribuent pas nécessairement de façon gaussienne.
Il existe d'autres distributions.





Forme mathématique

En mathématique, une fonction transforme une valeur x en son image $f(x)$

On note $x \mapsto f(x)$

et on dit : " x a pour image f de x "

Une distribution de probabilité est une fonction qui prend la valeur d'une observation et la transforme en la probabilité de l'observer.

On parle aussi de loi de probabilité

la distribution gaussienne, dite aussi distribution normale ou loi normale à la forme mathématique suivante :

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

x représente la valeur d'une observation et $f(x)$ représente la probabilité de l'observer

Quant à C , il s'agit d'une constante





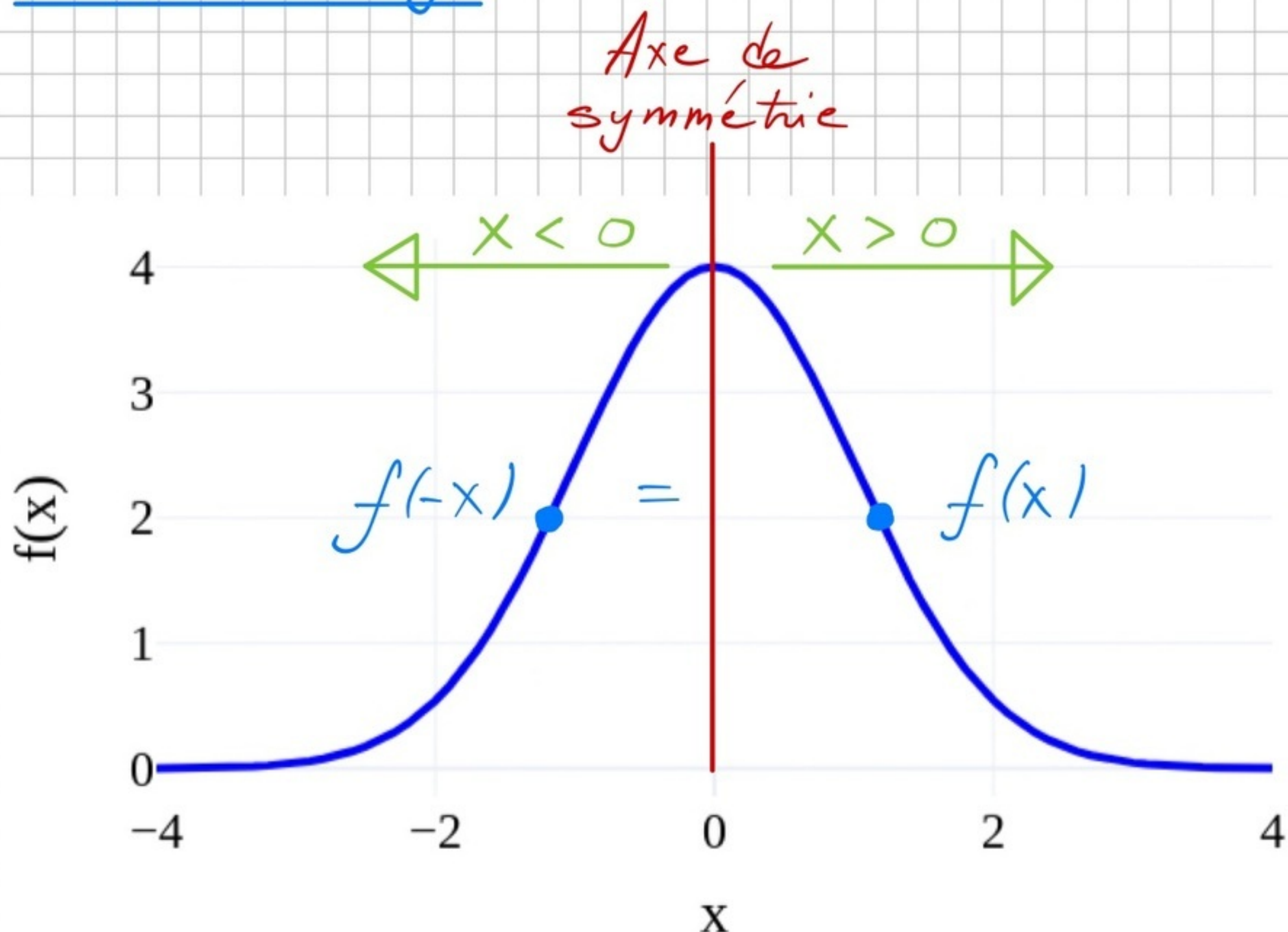
Regardons cette fonction de plus près pour comprendre ce qu'elle signifie :

On remarque tous d'abord que la variable x est élevée au carré.

Tous comme la fct $f(x) = x^2$, la fonction gaussienne retourne donc la même valeur pour les x positifs que pour les x négatifs.

On dit que ces deux fonctions sont paires

Graphiquement, cela se traduit par une courbe symétrique par rapport à l'axe des y

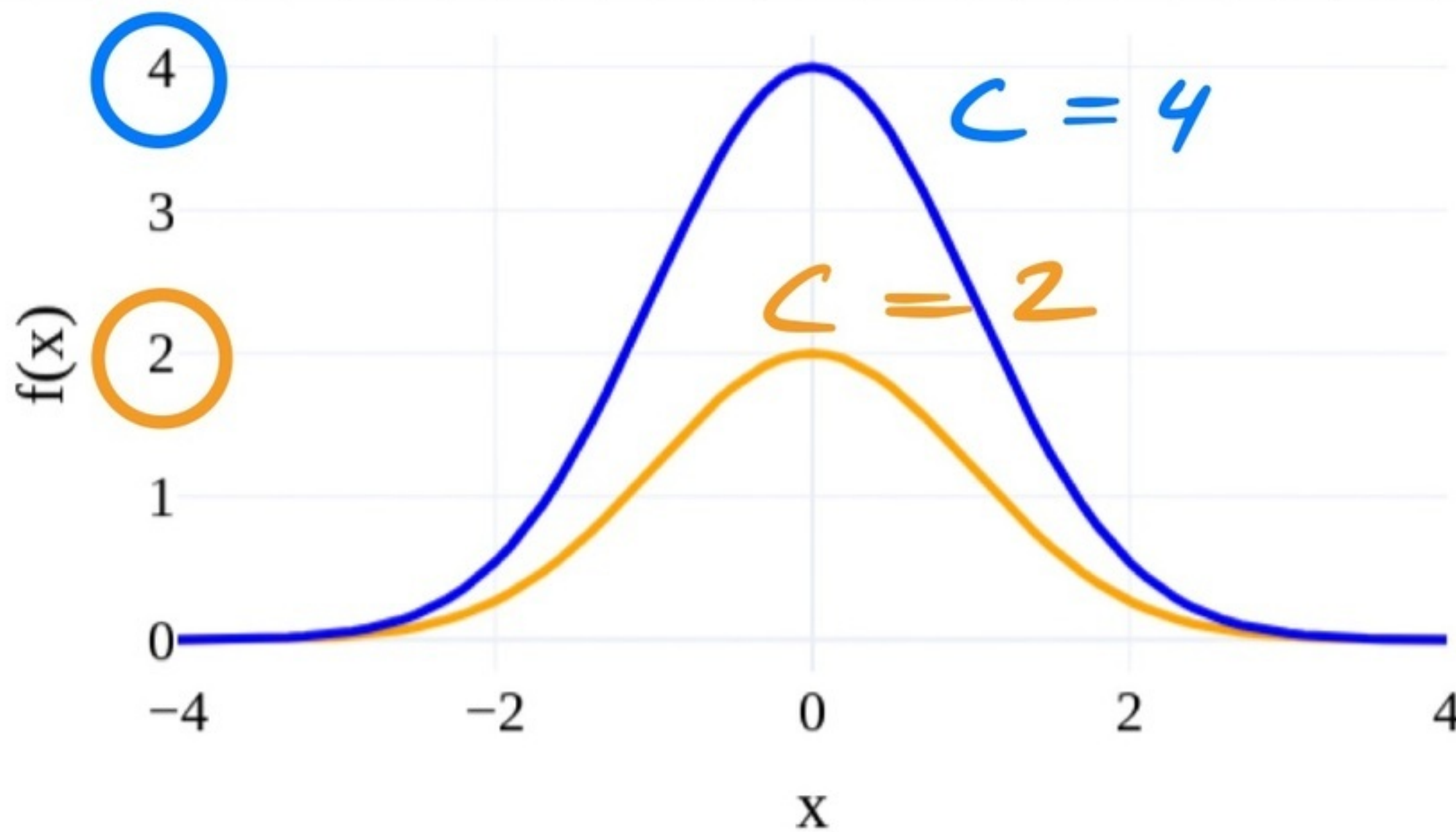




Le facteur C est une constante qui multiplie toute la fonction :

$$f(x) = 4 \cdot e^{-\frac{1}{2}x^2}$$

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



En $x = 0$, $\exp(-\frac{1}{2}x^2) = \exp(0) = 1$

lorsque $A = 2$, la fct retourne donc $2 \cdot \exp(0) = 2 \cdot 1 = 2$

Dans le cadre des statistiques, on souhaite que l'aire totale sous la courbe soit égale à 1.

En effet, la somme des probabilités de tous les événements vaut 100% (= 1)





Afin que l'aire sous une fonction gaussienne vaille 1, on choisit

$$C = \frac{1}{\sqrt{2\lambda}}$$

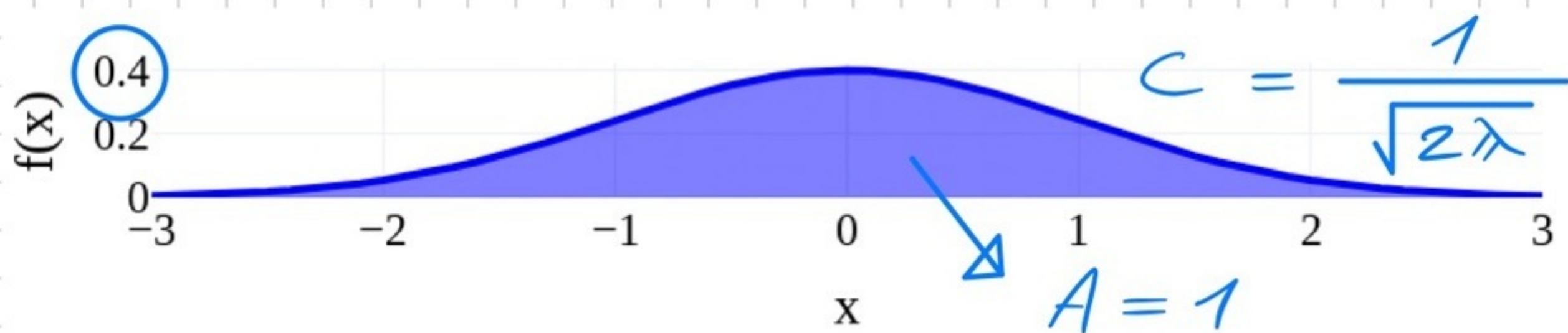
De cette façon, la fonction gaussienne suivante :

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

représente une distribution de probabilité.

x représente la valeur d'une observation, et $f(x)$ représente la probabilité de l'observer.

Sur un graphique, la distribution gaussienne se présente sous la forme d'une courbe très aplatie :



L'aire A sous la courbe s'écrit

$$A = \int_{-\infty}^{\infty} f(x) dx$$





Terminologie et notation

On note X (grand X)

la mesure du caractère observé.

Cela peut être une taille, un poids, une couleur ou tout autre grandeur qui caractérise la population que l'on souhaite étudier.

X se nomme une variable aléatoire

car elle prend ses valeurs dans un processus dont l'issue est incertaine.

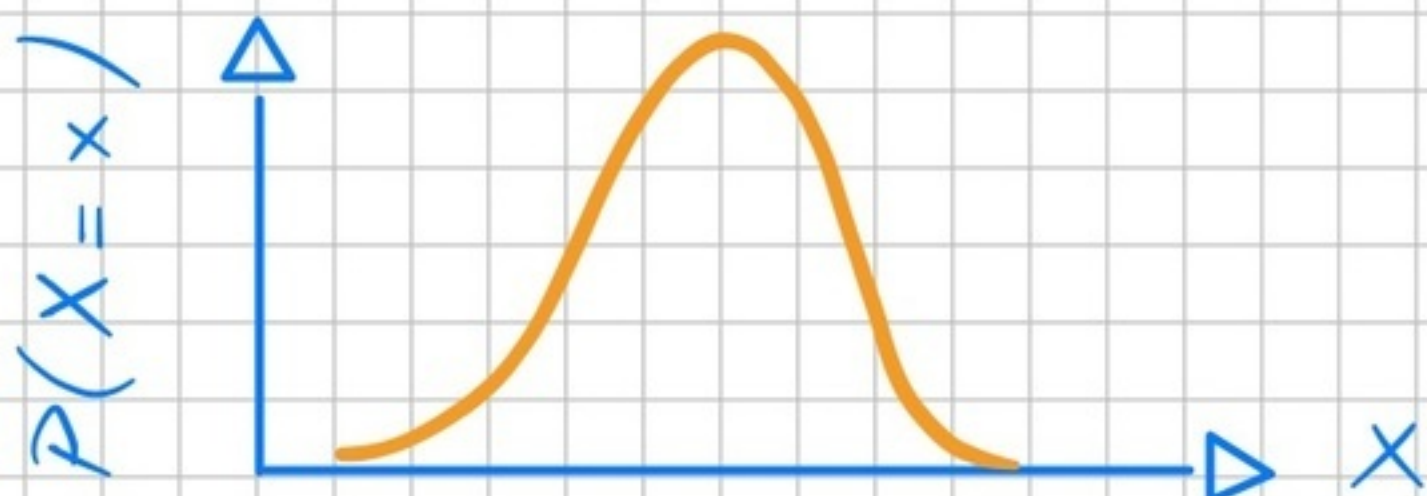
On note x (petit x)

la réalisation de la variable aléatoire X

la probabilité de voir se réaliser une valeur x de la variable aléatoire X se note $P(X=x)$ et se dit :

"probabilité que la variable aléatoire X prenne une valeur x)

Sur un graphique :



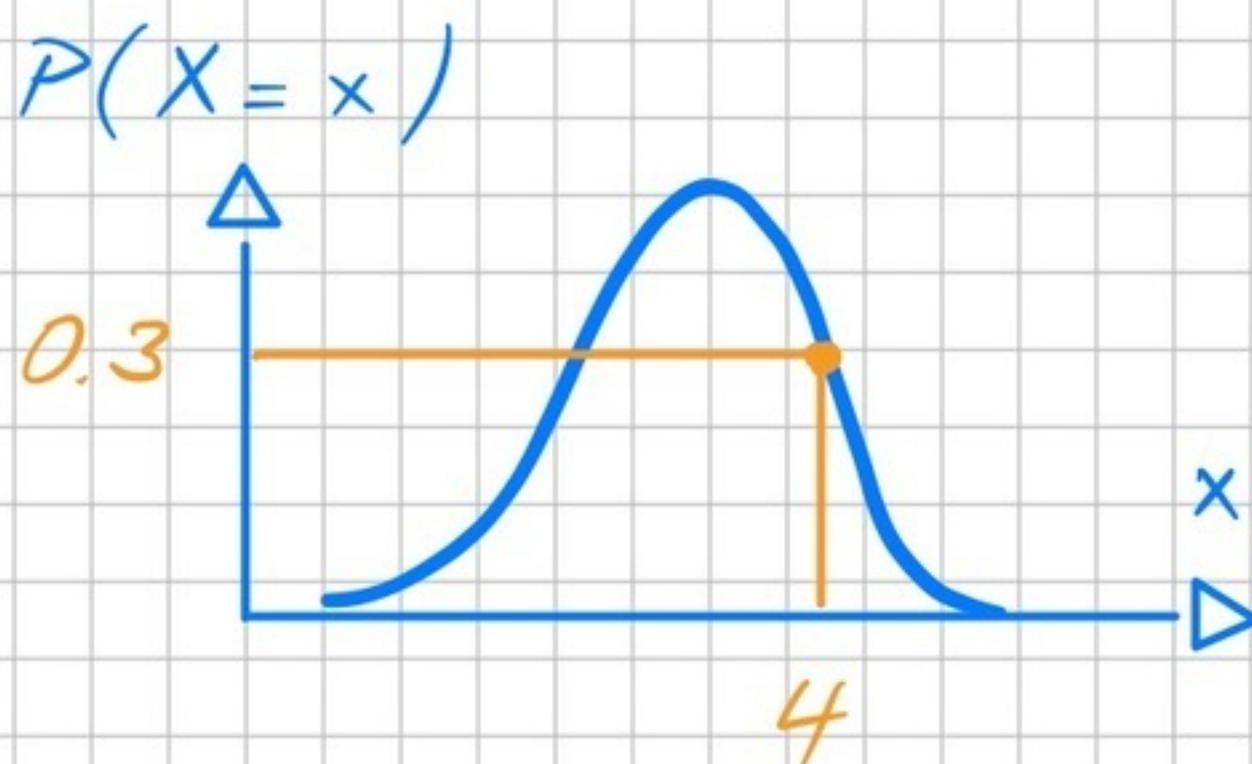


Calcul de probabilité

Toute l'utilité d'une distribution de probabilité est de permettre de calculer ... des probabilités !

Oui !

Si l'on s'intéresse à la probabilité d'un seul événement, il suffit de lire sur le graphique :

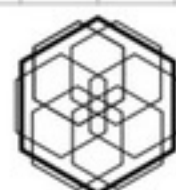


Ici, la probabilité d'observer la valeur 4 est donnée par $f(4) = 0.3$ ($= 30\%$)

On note $P(X=4) = 0.3$

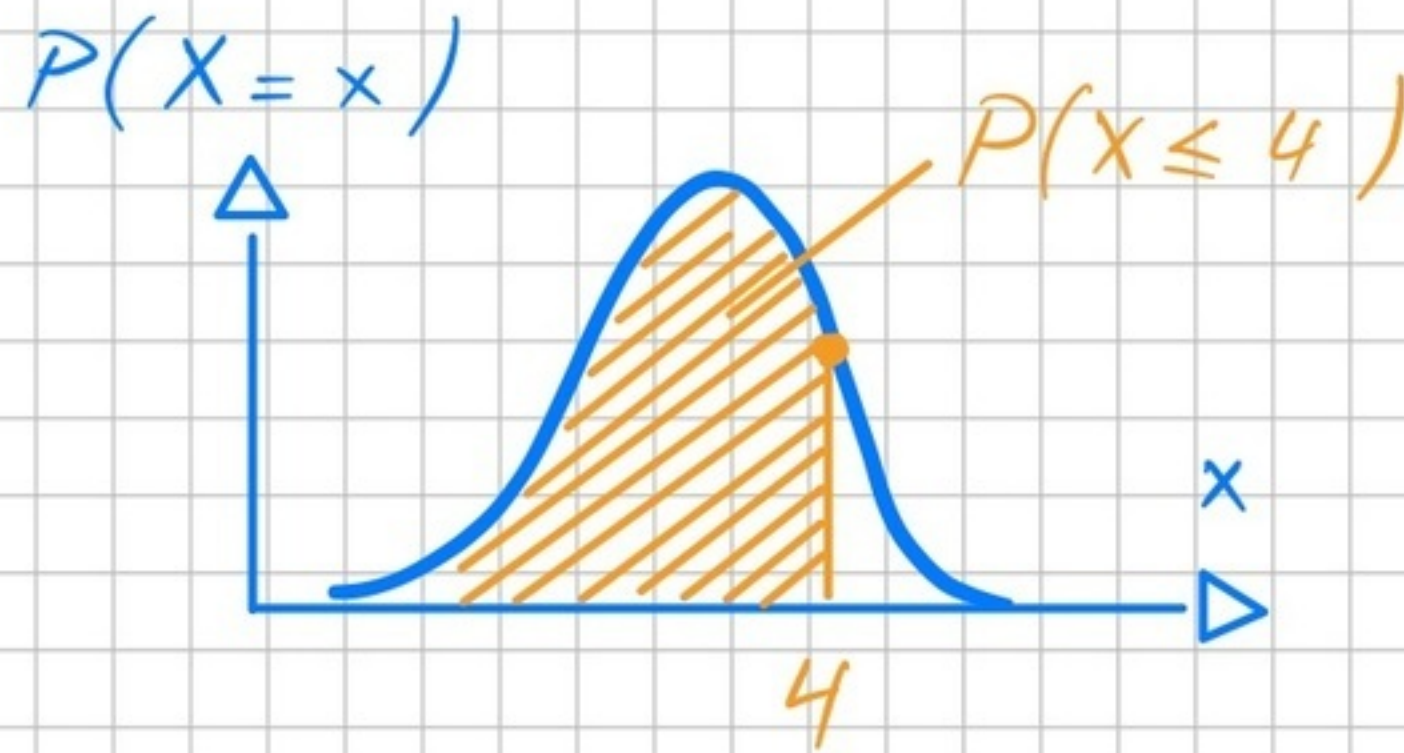
Dans la pratique, il arrive très souvent que l'on s'intéresse à la probabilité d'observer une valeur au moins aussi grande que 4.

On note cela $P(X \leq 4)$





Graphiquement, la probabilité $P(X \leq 4)$ représente l'aire sous la gaussienne entre $-\infty$ et 4



Cette aire est donnée par l'intégrale de la fonction gaussienne $f(x)$ entre $-\infty$ et 4.

$$P(X \leq 4) = \int_{-\infty}^4 f(x) dx$$

Or, il se trouve que la fonction gaussienne ne se laisse pas intégrer facilement. Il existe même un théorème qui démontre qu'il est impossible d'écrire l'intégrale d'une gaussienne (sa primitive) à l'aide des fonctions usuelles simples (théorème de Liouville)

Pour cette raison, on a recours à une intégration numérique.





Table des quantiles

L'intégration numérique est une technique qui permet d'intégrer une fonction sans connaître sa primitive.

Elle repose largement sur la puissance de calcul de nos ordinateurs.

Il fut un temps où ces calculs demandaient du temps et coûtaient donc cher. Pour cette raison, ils avaient été fait une fois pour toute et les résultats ont été publiés dans une table.

Cette table résume à elle seule tout ce qu'il faut savoir pour calculer n'importe quelle probabilité qui suivrait une distribution normale.

On lui donne le nom de table des quantiles

Je vous la donne ici, c'est cadeau !



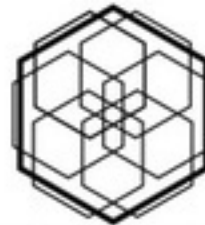


Table des quantiles de la distribution normale

X	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990





Voici maintenant comment utiliser cette table :

On lit la valeur x de la variable aléatoire X dans la première colonne, puis sur la première ligne comme ceci :

① $x = 0.27$

② Puis les centièmes là ↓

On lit d'abord les dixièmes ici →

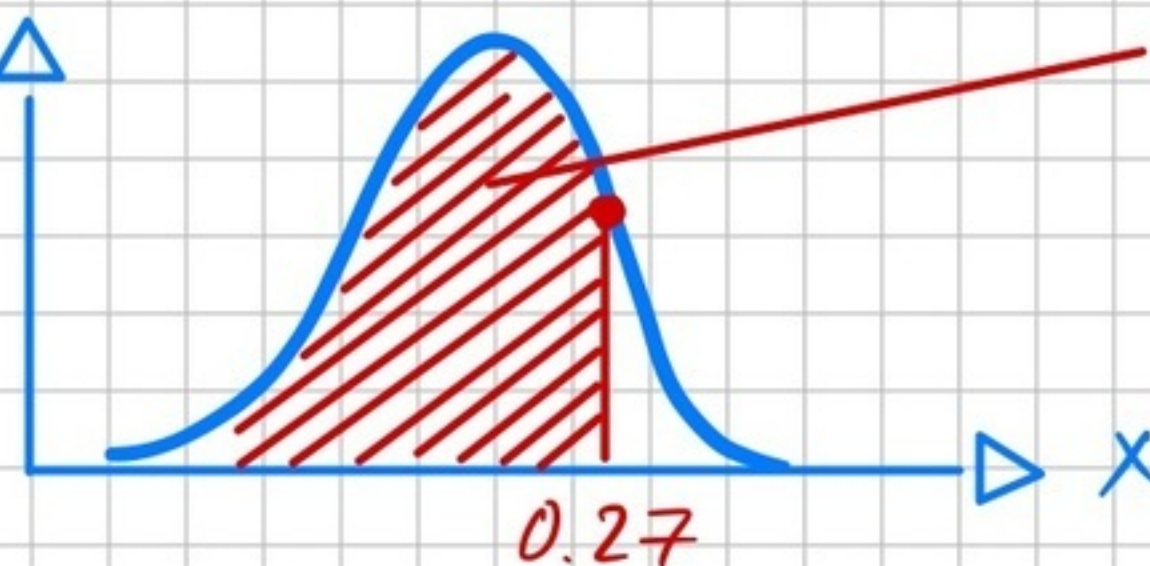
X	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879

À la valeur $x = 0.27$ correspond la valeur 0.6064

On note $\Phi(0.27) = 0.6064$

Il s'agit de la probabilité d'observer une valeur plus petite ou égale à 0.27

$P(X=x)$



$$\begin{aligned}\Phi(0.27) &= P(X \leq 0.27) \\ &= 0.6064\end{aligned}$$





Propriétés de symmétrie

Nous l'avons vu, la fonction gaussienne est symétrique autour de 0.

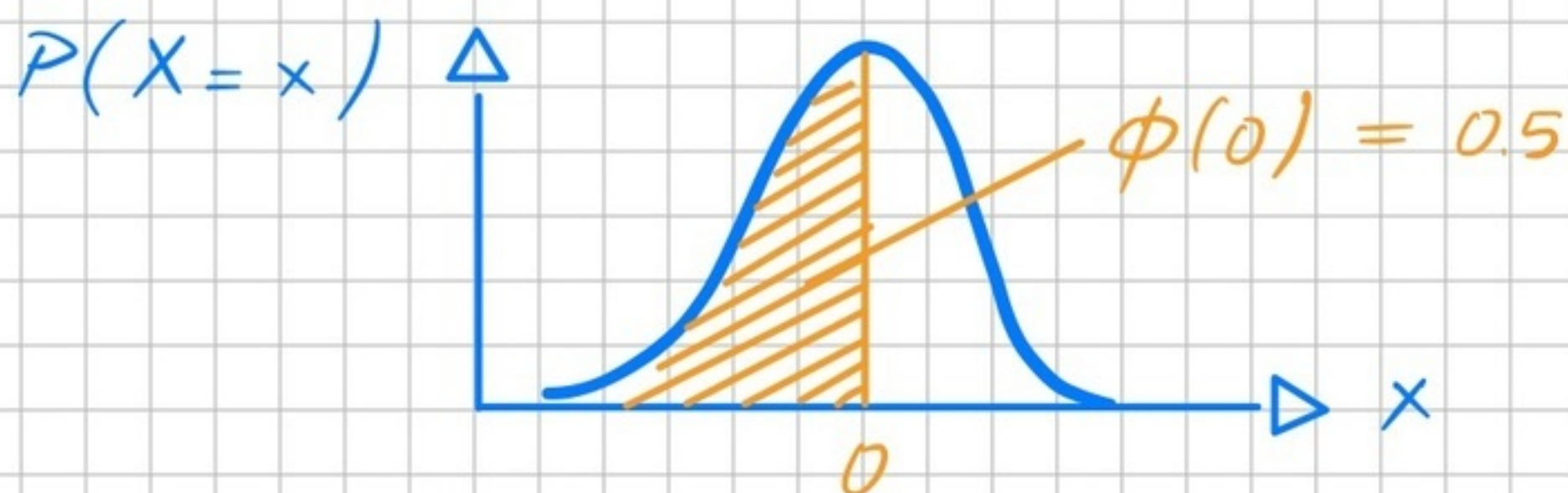
Cette parité est causée par le x élevé à la puissance 2 :

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

Nous savons également que l'aire totale sous la courbe vaut 1 :

$$A_{\text{tot}} = \int_{-\infty}^{\infty} f(x) dx = 1$$

Puisque la fonction est symétrique autour de 0, on s'attend à retrouver la moitié de l'aire totale à gauche de 0 et l'autre moitié à droite.



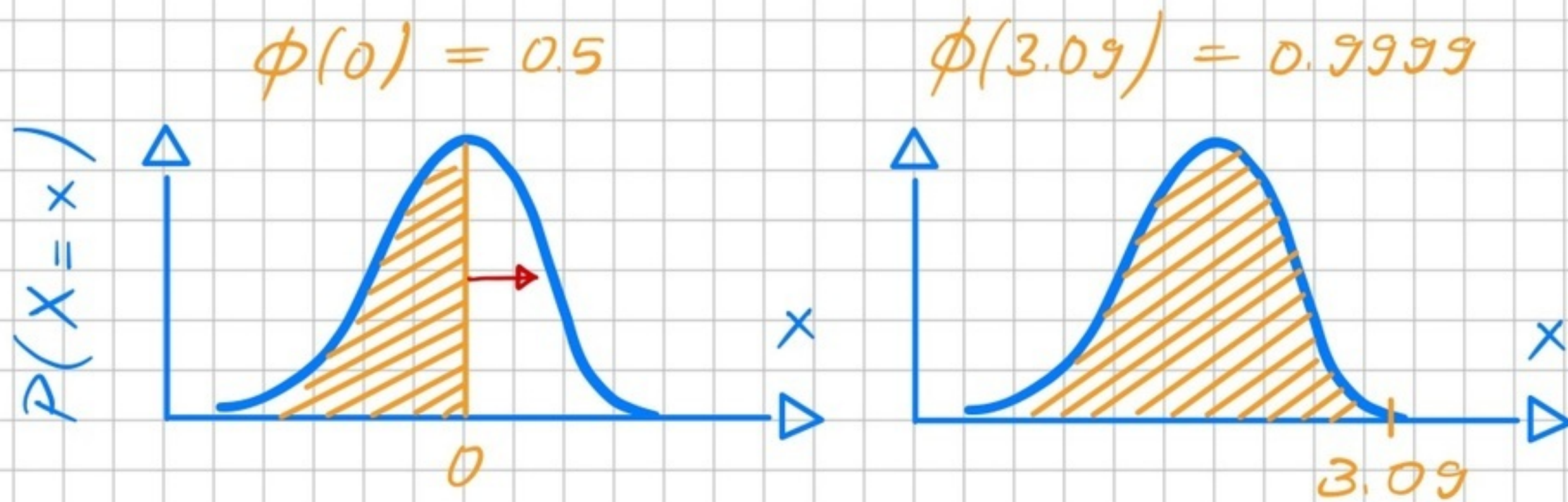
Ainsi, lorsqu'on cherche $\phi(0)$ dans la table, on trouve bien 0.5





Si l'on continue de lire la table au delà de $x=0$, on constate que l'aire sous la courbe ne cesse d'augmenter jusqu'à $x=3.09$.

À cet endroit en effet, c'est toute la distribution qui est derrière, et l'aire sous la courbe vaut $0.9999 \approx 1$ comme on s'y attend pour une distribution de probabilité.



On peut à présent comprendre facilement une propriété de symétrie importante.

Il s'agit pour cela de bien s'imaginer que l'aire sous la courbe augmente vers la droite de 0 de la même manière qu'elle diminue vers la gauche.

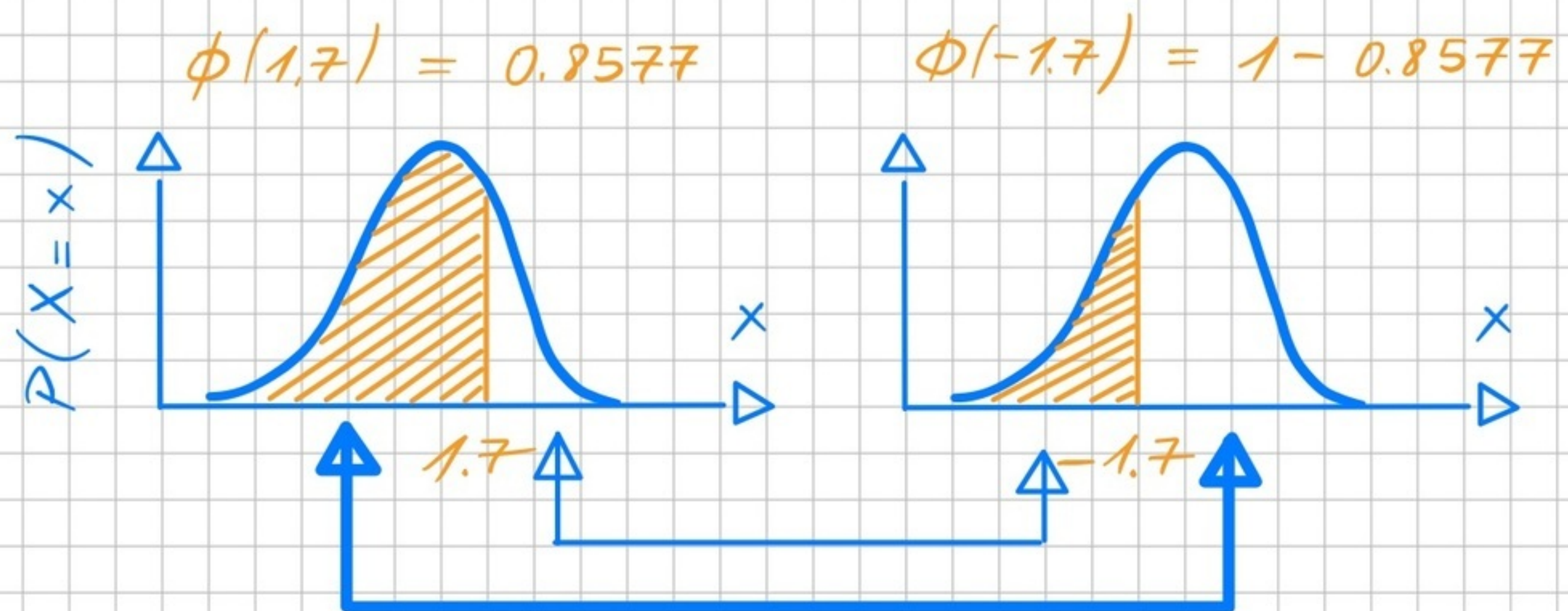




Pour cette raison, les valeurs négatives, (à gauche de 0) ne se trouvent pas dans les tables. On peut les calculer facilement à l'aide de l'égalité suivante :

$$\phi(-x) = 1 - \phi(x)$$

Il s'agit là d'une propriété de symétrie importante que l'on peut représenter de la manière suivante :





Dans la réalité

La gaussienne que nous avons étudiée jusqu'ici est un objet mathématique abstrait.

Nous allons voir maintenant comment s'en servir concrètement pour modéliser des observations faites sur le terrain.

Chaque observation cependant sera différente des autres et on ne peut pas imaginer qu'elle donneront toute la même distribution.

Pour reprendre l'exemple de l'arbre, on s'imagine bien que : des pommes ne se répartissent pas autour d'un pommier de la même manière que des pins autour d'un sapin.

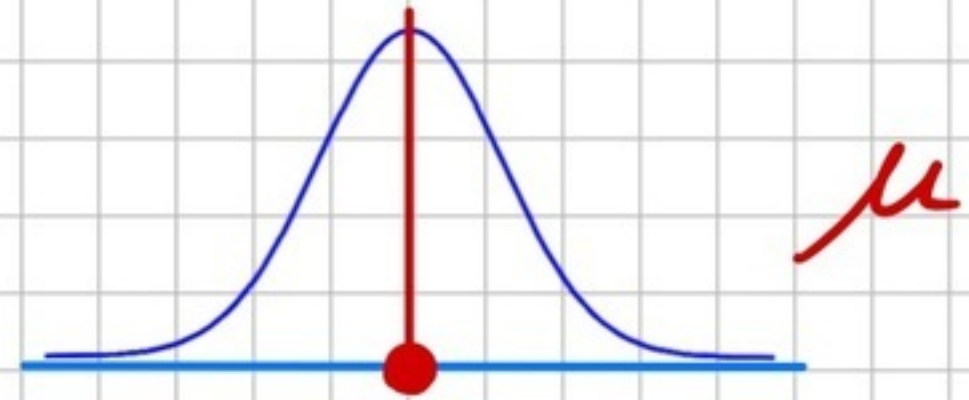
Dans le langage des statistiques, on dit que ces fruits ne se distribuent pas de la même manière.



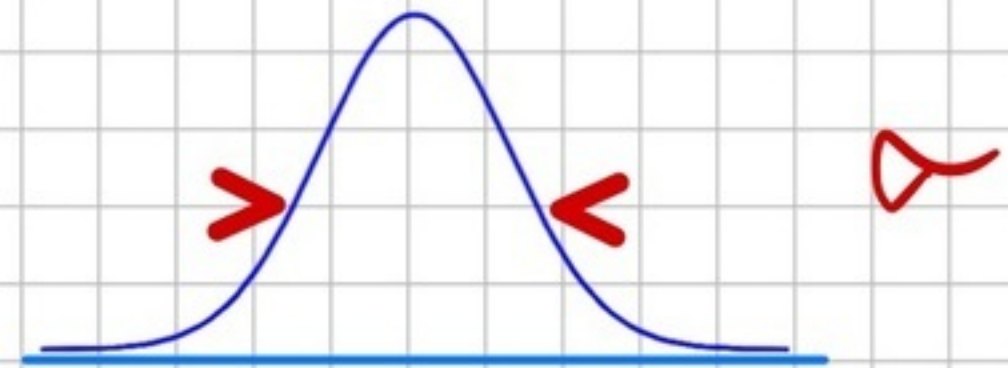


Pour adapter la distribution gaussienne à la réalité, on utilise deux valeurs :

1) la position sur x du maximum



2) la largeur de la distribution



Ces deux grandeurs ont une signification concrète et portent donc un nom.

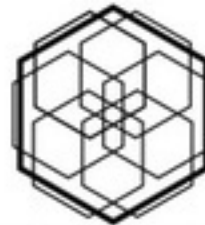
1) On appelle moyenne la position sur x du maximum et on la note μ (mu)

Il s'agit de la valeur autour de laquelle est centrée la gaussienne.

2) On appelle écart-type l'étendue de la distribution et on la note σ (sigma)

Il s'agit d'une mesure de la dispersion des valeurs de x autour de la moyenne





Ces deux valeurs permettent d'ajuster la gaussienne dans le but de représenter des observations réelles.

Elles peuvent donc varier librement d'une situation à l'autre.

Pour cette raison, on les appelle paramètres de la fonction.

On se souvient de la forme mathématique de la fonction gaussienne de base :

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

Voici maintenant comment ses deux paramètres la modifient pour la rendre ajustable :

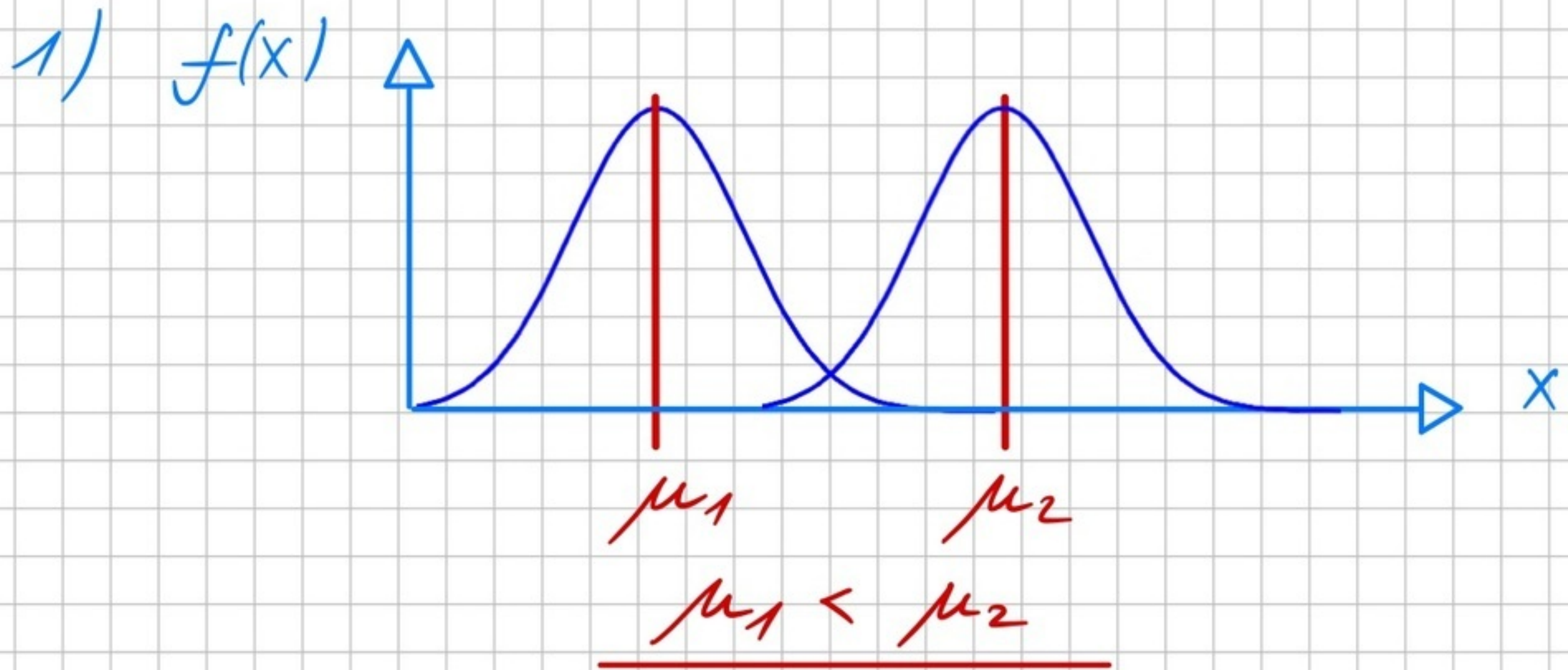
$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2}$$

Nous avons là l'expression mathématique de la loi de probabilité gaussienne utilisée par les statisticiens et statisticiennes.

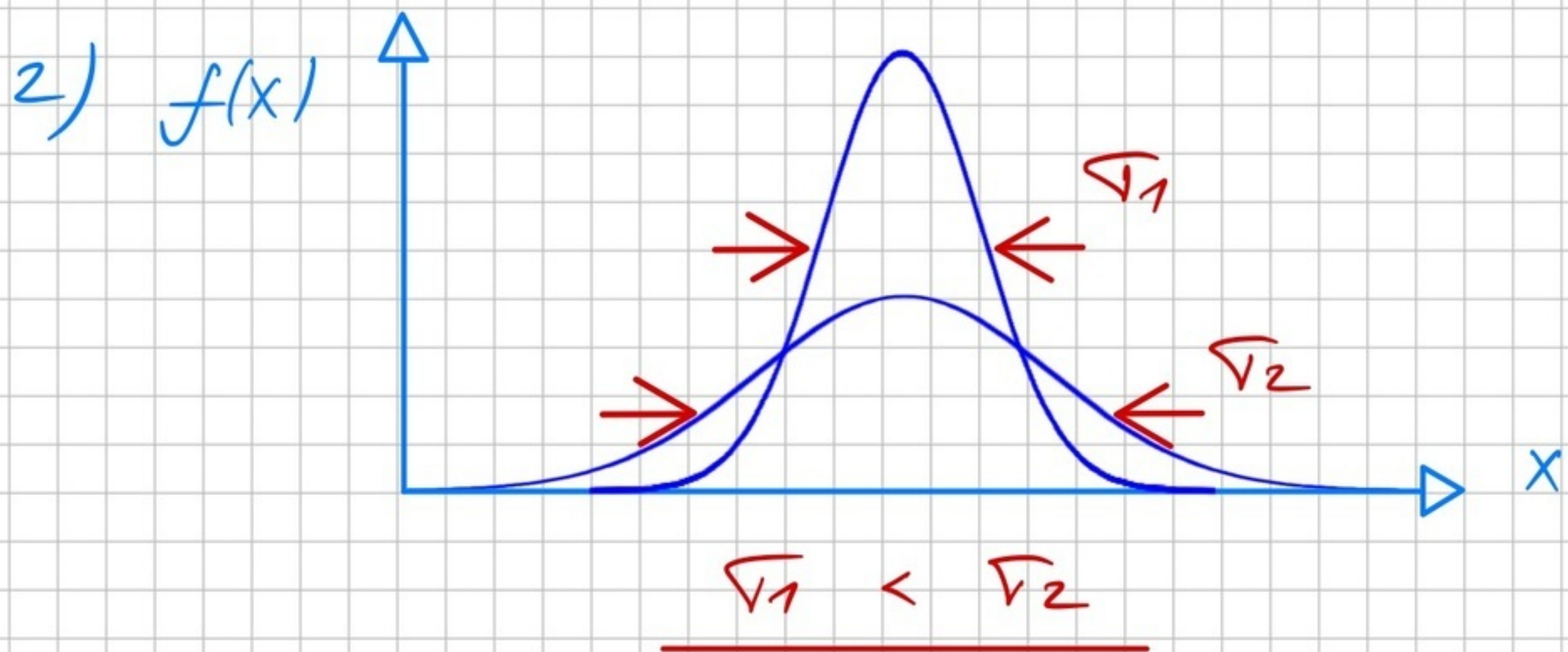




Si l'on cherche à se représenter l'action des paramètres sur la fonction, voici ce que l'on obtient



La moyenne déplace horizontalement la gaussienne (sur l'axe des x)



L'écart-type élargit ou rétrécit la distribution autour de la moyenne. L'aire sous la courbe est conservée et reste donc égale à 1.





Centrer - Réduire

Ces deux paramètres, μ et σ , nous permettent donc d'ajuster la distribution à chaque situation.

Ils sont donc différents d'une situation à l'autre.

On peut dès lors très bien envisager le processus inverse : normaliser chaque distribution particulière vers la même distribution de base.

On appelle ce processus "centrer - réduire"

Il permet de se reporter à la table des quantiles pour calculer des probabilités à la main (sans ordi)

1) Centrer consiste à soustraire la moyenne μ à X

2) Réduire consiste à diviser X par son écart-type σ





Ainsi, si une variable aléatoire X suit une loi normale de moyenne μ et d'écart-type σ , on note :

$$X \sim N(\mu; \sigma^2)$$

Alors la variable X^* centrée - réduite s'obtient en effectuant les opérations 1) et 2)

$$X^* = \frac{X - \mu}{\sigma}$$

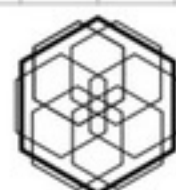
Cette nouvelle variable a la particularité de suivre une loi normale de moyenne $\mu = 0$ et d'écart-type $\sigma = 1$.

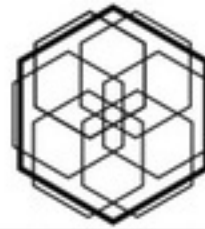
On note : $X^* \sim N(0; 1)$

C'est la distribution qui a servi à calculer la table des quantiles.

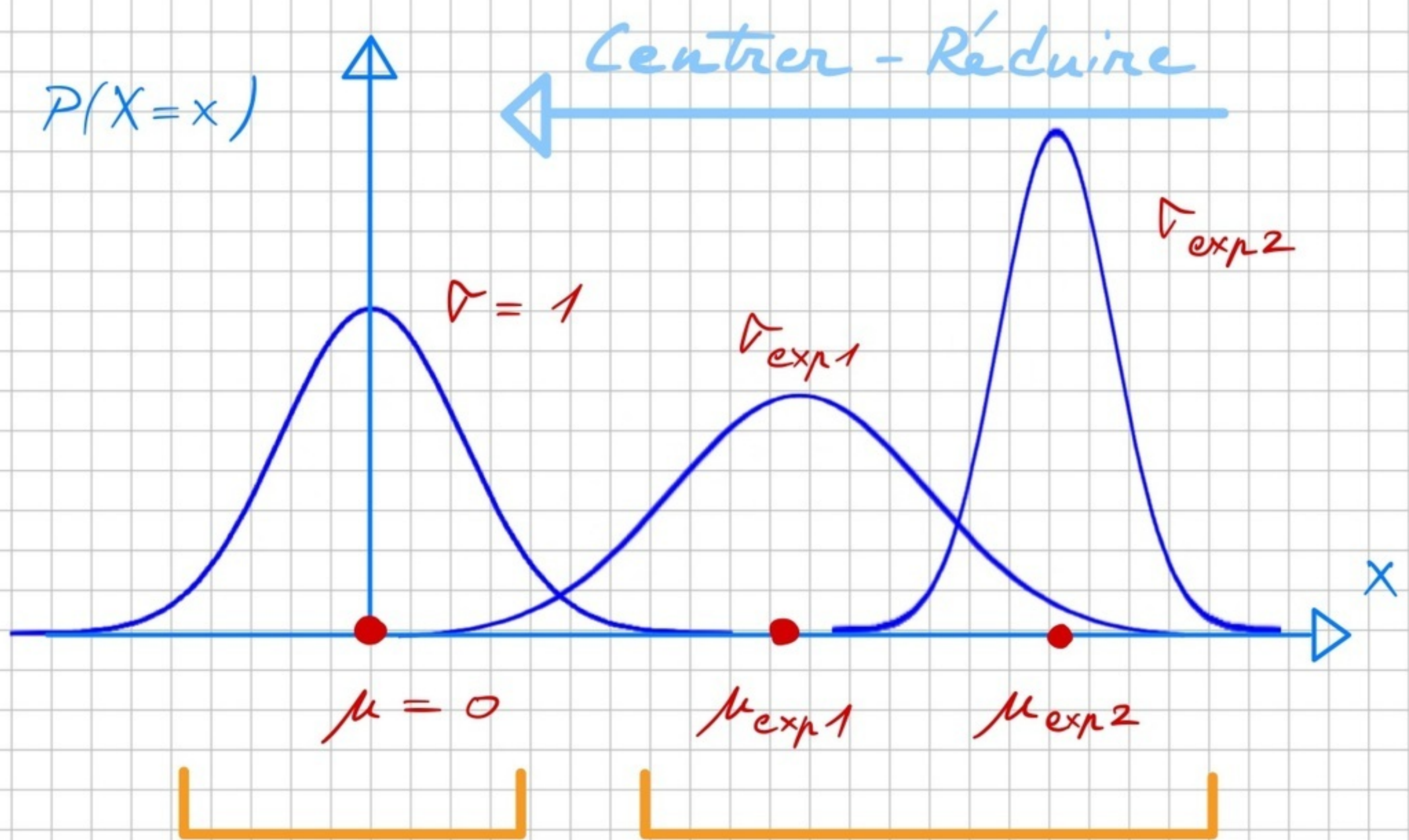
La grandeur σ^2 se nomme variance.

Tout comme l'écart-type σ , la variance est une mesure de la dispersion.





On peut représenter graphiquement le processus de centrer - réduire de la manière suivante :



Une seule et même distribution centrée - réduite

$$\mu = 0$$

$$\sigma = 1$$

Deux distributions expérimentales différentes

$$\mu = \mu_{exp}$$

$$\sigma = \sigma_{exp}$$

x	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6025	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

Centrer - Réduire





Exercices corrigés

Exercice 001

Une firme fait passer des tests aux postulants pour certains postes de l'entreprise. Elle a pu établir que les résultats à ces tests sont distribués normalement avec une espérance = 73.2 et un écart-type = 7. Quelle est la probabilité que le prochain postulant obtienne un résultat inférieur à 70.0? Quelle est la probabilité que le prochain postulant obtienne un résultat supérieur à 82.8? Au-dessous de quelle note trouve-t-on 25% des postulants jugés les moins aptes? Réponses : 32.4%, 8.5%, 68.5%

les résultats sont normalement distribués

$$\Leftrightarrow X \sim N(\mu, \sigma^2)$$

$$\mu = 73.2$$

$$\sigma = 7$$

a) Quelle est la probabilité que le prochain postulant obtienne une note inférieure à 70?

$$\Leftrightarrow \text{Que vaut } P(X \leq 70) ?$$

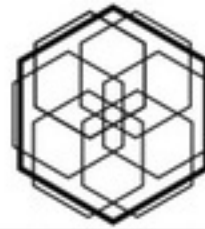
$$P(X \leq 70) = P\left(\frac{X - \mu}{\sigma} \leq \frac{70 - \mu}{\sigma}\right)$$

On a centré et réduit

$$= P\left(X^* \leq \frac{70 - \mu}{\sigma}\right)$$

$$= P\left(X^* \leq \frac{70 - 73.2}{7}\right)$$





$$= P\left(X^* \leq \frac{70 - 73.2}{7}\right)$$

$$= P\left(X^* \leq -0.4571\right)$$

par symétrie on a
 $P(X \leq -A) = 1 - P(X \leq A)$

$$= 1 - P\left(X^* \leq 0.4571\right)$$

l'aire sous la courbe
entre $-\infty$ et 0.4571
se note $\phi(0.4571)$

$$= 1 - \phi(0.4571)$$

on lit $\phi(0.4571)$ dans
la table des quantiles

$$= 1 - 0.6772$$

$$= 0.3228$$

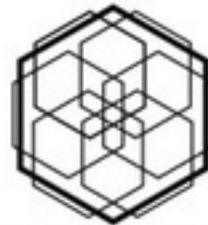
$$= \underline{\underline{32.3\%}}$$

$$0.4571 \approx 0.46$$



X	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879





b) Quelle est la probabilité que le prochain postulant obtienne une note supérieure à 82.8

$\Leftrightarrow$ Que vaut $P(X > 82.8)$

$$P(X > 82.8) = 1 - P(X \leq 82.8)$$

$$= 1 - P\left(\frac{X - \mu}{\sigma} \leq \frac{82.8 - \mu}{\sigma}\right)$$

$$= 1 - P(X^* \leq \frac{82.8 - \mu}{\sigma})$$

$$= 1 - P(X^* \leq \frac{82.8 - 73.2}{7})$$

$$= 1 - P(X^* \leq 1.37)$$

$$= 1 - \Phi(1.37)$$

$$= 1 - 0.9147$$

$$= 0.0853$$

$$= \underline{\underline{8.5\%}}$$

1.37



X	0	1	2	3	4	5	6	7	8	9
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319





c) En dessous de quelle note
trouve-t-on 25% des postulants ?

Cette question est posée à l'envers.

D'habitude on demande :

" quelle est la probabilité
d'observer telle valeur ? "

Et cette fois on demande :

" quelle valeur me donnerait
telle probabilité ? "

On fait comme si on connaissait
cette valeur, et on la note $x_{\min}$

On peut maintenant formuler
la question ainsi :

Trouver $x_{\min}$ tel que $P(X \leq x_{\min}) = 25\%$

$$P(X \leq x_{\min}) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x_{\min} - \mu}{\sigma}\right)$$

$$= P(X^* \leq x_{\min}^*)$$

$$= \Phi(x_{\min}^*)$$

$$= 25\%$$

$$= 0.25$$





$$P(X^* \leq x_{\min}^*) = 0.25$$

$$= 1 - 0.75$$

on utilise cette astuce
car 0.25 ne se trouve
pas dans la table.
il s'agit d'une
propriété de symmétrie

$$= \phi(x_{\min}^*)$$

$$1 - 0.75 = \phi(x_{\min}^*)$$

$$0.75 = \phi(-x_{\min}^*)$$

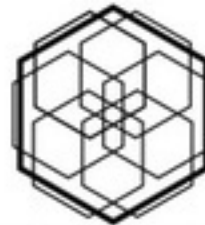
propriété de symmétrie

En lisant maintenant la table
à l'envers, on trouve :

$$\phi(0.67) = 0.7486 \approx 0.75$$

X	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389





$$\begin{aligned} 0.75 &= \phi(-x_{\min}^*) \\ &= \phi(0.67) \end{aligned}$$

$\Leftrightarrow$

$$-x_{\min}^* = 0.67$$

Or, on a

$$x_{\min}^* = \frac{x_{\min} - \mu}{\sigma}$$

$$\begin{aligned} -x_{\min}^* &= \frac{-x_{\min} + \mu}{\sigma} \\ &= 0.67 \end{aligned}$$

$\Leftrightarrow$

$$\frac{-x_{\min} + \mu}{\sigma} = 0.67$$

$$-x_{\min} + \mu = 0.67 \cdot \sigma$$

$$-x_{\min} = 0.67 \cdot \sigma - \mu$$

$$x_{\min} = \mu - 0.67 \cdot \sigma$$

$$= 73.2 - 0.67 \cdot 7$$

$$= \underline{\underline{68.51}}$$





Exercice 002

En admettant que les températures du mois de juillet ont en un endroit une distribution normale de moyenne $18,2^{\circ}\text{C}$ et d'écart-type $3,6^{\circ}\text{C}$, calculer la probabilité que la température soit comprise un jour de juillet entre 20°C et 25°C .

La température suit une loi normale
d'espérance $\mu = 18.2$
et d'écart-type $\sigma = 3.6$

$$\Leftrightarrow X \sim N(18.2, 3.6^2)$$

Quelle est la probabilité que
la température soit comprise
entre 20 et 25 degrés ?

$\Leftrightarrow$

Que vaut $P(20 < X < 25)$?

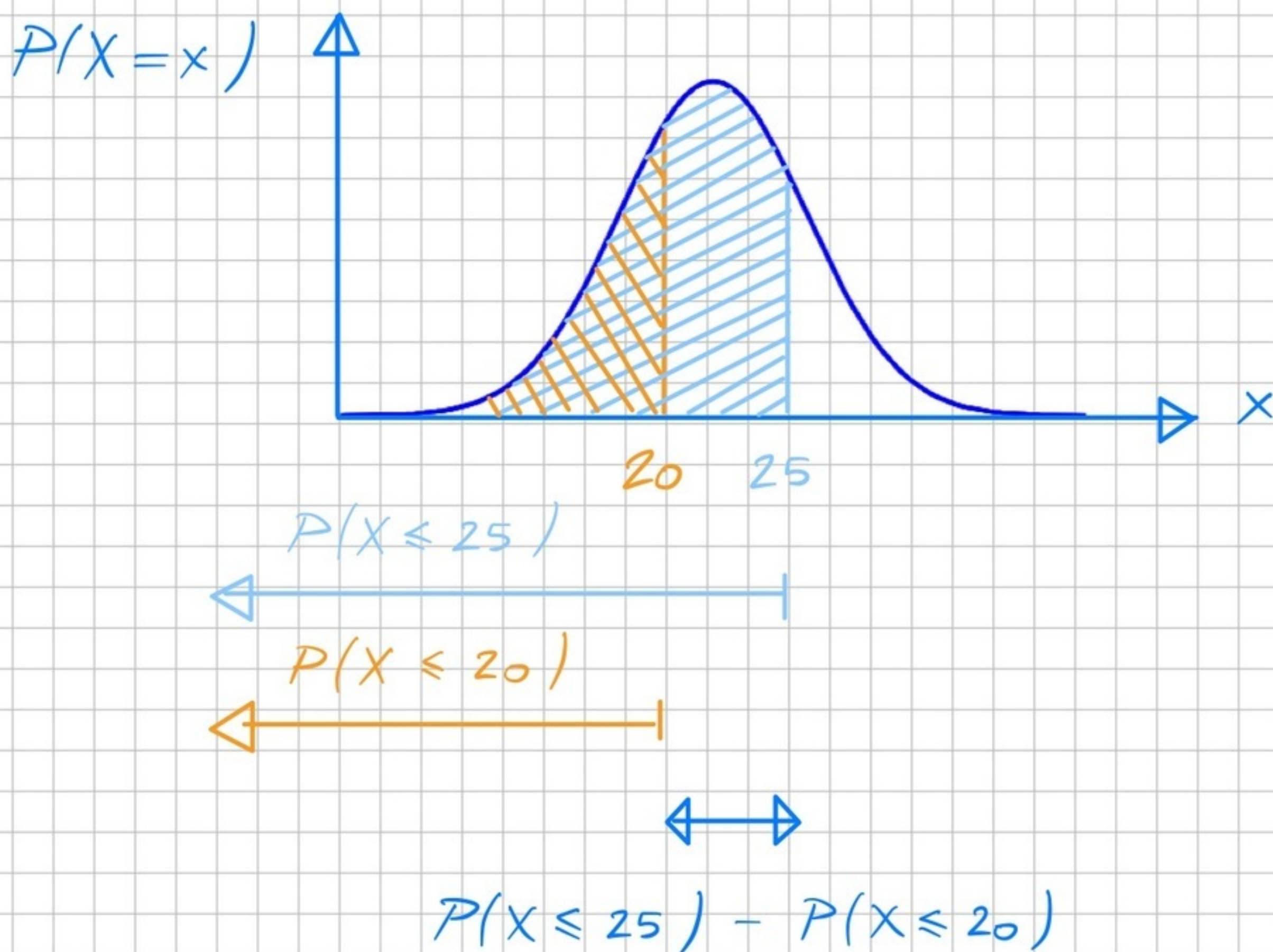




On,

$$P(20 < X < 25) = P(X \leq 25) - P(X \leq 20)$$

On peut se représenter graphiquement cette égalité de la manière suivante :

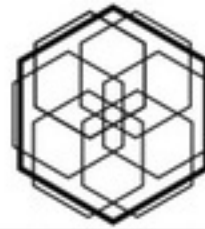


On calcule donc chacune de ces deux probabilités séparément

1) $P(X \leq 25)$

2) $P(X \leq 20)$





$$\begin{aligned}
 P(X \leq 25) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{25 - \mu}{\sigma}\right) \\
 &= P\left(X^* \leq \frac{25 - \mu}{\sigma}\right) \\
 &= P\left(X^* \leq \frac{25 - 18.2}{3.6}\right) \\
 &= P(X^* \leq 1.89) \\
 &= \Phi(1.89) \\
 &= \underline{0.9706}
 \end{aligned}$$

1.89



X	0	1	2	3	4	5	6	7	8	9
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

$$\begin{aligned}
 P(X \leq 20) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{20 - \mu}{\sigma}\right) \\
 &= P\left(X^* \leq \frac{20 - \mu}{\sigma}\right) \\
 &= P\left(X^* \leq \frac{20 - 18.2}{3.6}\right) \\
 &= P(X^* \leq 0.50)
 \end{aligned}$$





$$\begin{aligned}
 P(X \leq 20) &= P(X^* \leq 0.50) \\
 &= \Phi(0.50) \\
 &= \underline{0.6915}
 \end{aligned}$$

0.50



X	0	1	2	3	4	5	6	7	8	9
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389

On peut maintenant calculer :

$$\begin{aligned}
 P(20 < X < 25) &= P(X \leq 25) - P(X \leq 20) \\
 &= 0.9706 - 0.6915 \\
 &= \underline{0.2791}
 \end{aligned}$$

Ainsi, la probabilité de mesurer une température comprise entre 20 et 25 degrés est de 27.91%





Exercice 003

Les articles produits par une machine ont un poids qui suit une loi normale de moyenne 5.15g et d'écart-type 0.37g. Trouver la probabilité pour qu'un ensemble de 100 articles ait une masse (totale) comprise entre 514 et 516g? supérieure à 516g? Attention à l'écart-type à considérer pour la distribution des 100 articles. Réponses : 21.3%, 39.3%

Nous allons voir ici comment étudier le comportement de 100 articles à partir des caractéristiques d'un seul d'entre eux.

Le poids X d'un seul article suit une loi normale :

$$X \sim N(\mu, \sigma^2)$$

$$\mu = 5.15$$

$$\sigma = 0.37$$

On s'imagine bien que si le poids moyen d'un article vaut 5.15, alors le poids moyen de 100 articles vaut 100 fois plus, soit 515.

On note $\mu_{100} = 100 \cdot \mu$ le poids moyen de 100 articles





Nous allons donc dès à présent travailler avec une nouvelle variable : X_{100}
 X_{100} désigne le poids de 100 articles.

Tout comme la variable aléatoire X ,
la variable aléatoire X_{100} suit
une loi normale.

Pour trouver ses paramètres, on utilise
une propriété d'additivité de la gaussienne

Nous venons de voir :

$$\mu_{100} = 100 \mu$$

Concernant la mesure de dispersion,
c'est la variance σ^2 qui est additive
et non pas l'écart-type σ .

On a $\sigma^2 = 0.37^2$

$$= 0.1369$$

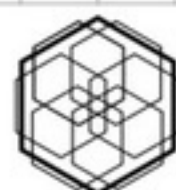
et donc

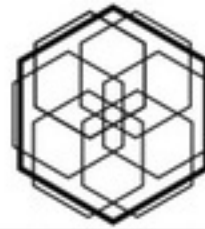
$$\sigma_{100}^2 = 100 \cdot \sigma^2$$

$$(100\sigma)^2 \neq 100\sigma^2$$

$$= 100 \cdot 0.1369$$

$$= 13.69$$





On dit que :

la loi normale est additive par rapport à ses deux paramètres μ et σ^2 .

$$\mu_{100} = 100 \mu$$

$$\sigma_{100}^2 = 100 \sigma^2$$

Dans notre cas

$$\mu_{100} = 5/5$$

$$\sigma_{100}^2 = 13.69$$

On a donc :

$$X_{100} \sim N(\mu_{100}, \sigma_{100}^2)$$

$$\sim N(5/5, 13.69)$$

et nous pouvons dès maintenant répondre aux questions comme nous l'avons fait jusqu'à maintenant :

a) Quelle est la probabilité qu'un ensemble de 100 objets aie un poids compris entre 5/4 et 5/6

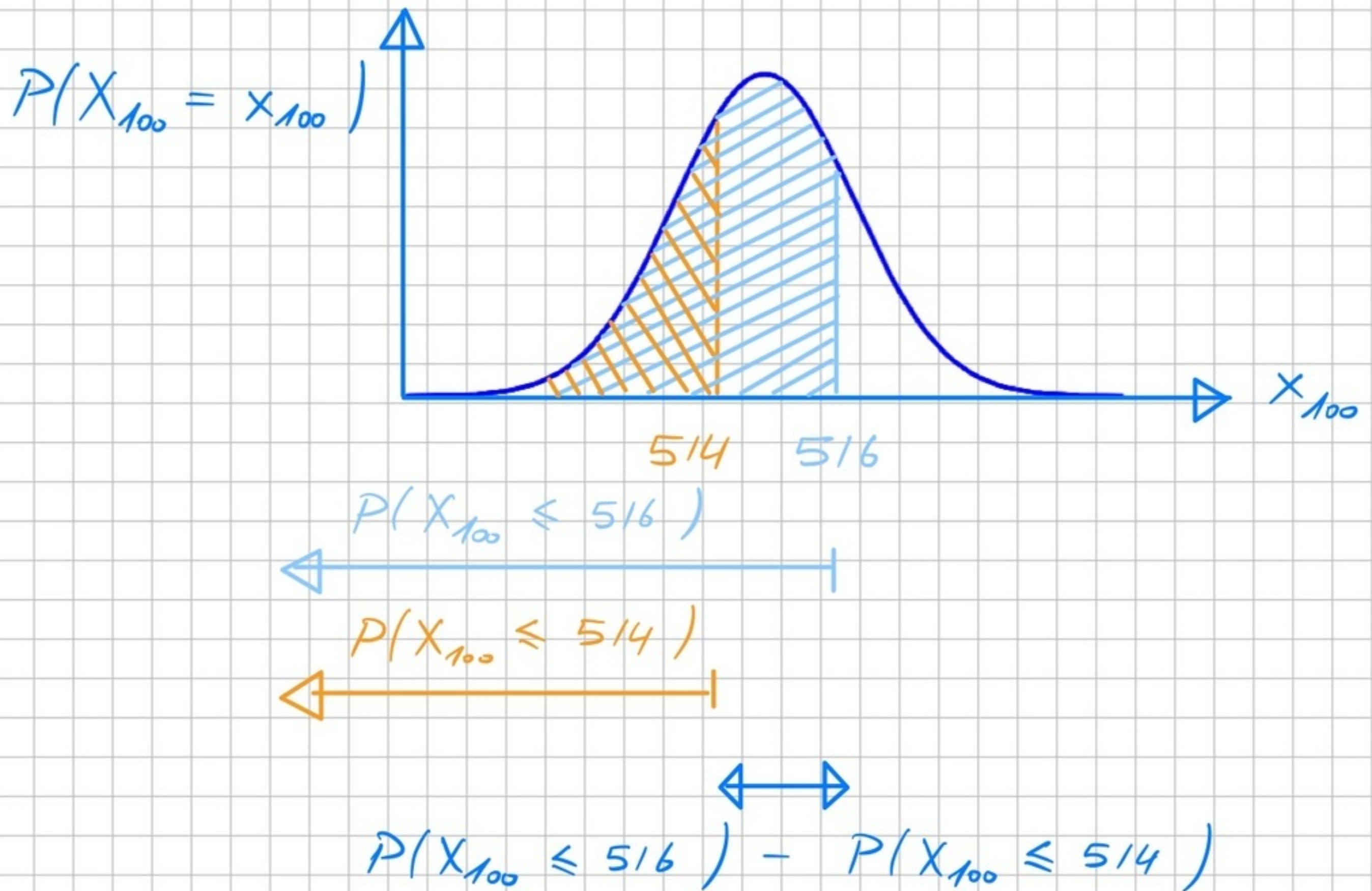
$\Leftrightarrow$ Que vaut $P(5/4 \leq X_{100} \leq 5/6)$?





$$P(514 \leq X_{100} \leq 516)$$

$$= P(X_{100} \leq 516) - P(X_{100} \leq 514)$$



On calcule chacune de ces deux
quantité séparément :

1) $P(X_{100} \leq 516)$

2) $P(X_{100} \leq 514)$





$$\begin{aligned}
 P(X_{100} \leq 516) &= P\left(\frac{X_{100} - \mu_{100}}{\sqrt{\sigma_{100}^2}} \leq \frac{516 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right) \\
 &= P\left(X_{100}^* \leq \frac{516 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right) \\
 &= P\left(X_{100}^* \leq \frac{516 - 515}{3.7}\right) \\
 &= P(X_{100}^* \leq 0.27) \\
 &= \Phi(0.27) \\
 &= \underline{0.6064}
 \end{aligned}$$

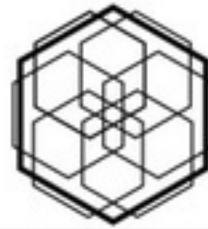


→

X	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879

$$\begin{aligned}
 P(X_{100} \leq 514) &= P\left(\frac{X_{100} - \mu_{100}}{\sqrt{\sigma_{100}^2}} \leq \frac{514 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right) \\
 &= P\left(X_{100}^* \leq \frac{514 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right) \\
 &= P\left(X_{100}^* \leq \frac{514 - 515}{3.7}\right)
 \end{aligned}$$





$$\begin{aligned} P(X_{100} \leq 514) &= P\left(X_{100}^* \leq \frac{514 - 515}{3.7}\right) \\ &= P\left(X_{100}^* \leq -0.27\right) \\ &= 1 - P\left(X_{100}^* \leq 0.27\right) \\ &= 1 - \Phi(0.27) \\ &= 1 - 0.6064 \\ &= \underline{0.3936} \end{aligned}$$

On a donc :

$$\begin{aligned} P(514 \leq X_{100} \leq 516) &= P(X_{100} \leq 516) - P(X_{100} \leq 514) \\ &= 0.6064 - 0.3936 \\ &= \underline{0.2128} \end{aligned}$$

Ainsi, la probabilité que le poids d'un ensemble de 100 articles soit compris entre 514 et 516 est de 21.28 %





b) Quelle est la probabilité
qu'un ensemble d'objet
aie un poids de plus de 516 g ?

$\Leftrightarrow$ Que vaut $P(X_{100} > 516)$?

$$\begin{aligned}P(X_{100} > 516) &= 1 - P(X_{100} \leq 516) \\&= 1 - P\left(\frac{X_{100} - \mu_{100}}{\sqrt{\sigma_{100}^2}} \leq \frac{516 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right) \\&= 1 - P\left(X_{100}^* \leq \frac{516 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right) \\&= 1 - P\left(X_{100}^* \leq \frac{516 - 515}{3.7}\right) \\&= 1 - P(X_{100}^* \leq 0.27) \\&= 1 - \Phi(0.27) \\&= 1 - 0.6064 \\&= \underline{\underline{0.3936}}\end{aligned}$$

La probabilité qu'un ensemble d'objet
aie un poids de plus de 516 g
est de 39.36 %





Exercice 004

Le contenu en nicotine d'une cigarette d'une marque précise est une variable aléatoire d'espérance (moyenne) $\mu = 0.8$ mg et d'écart-type $\sigma = 0.1$ mg. Un certain fumeur de cette marque consomme 5 paquets par semaine (1 paquet = 20 cigarettes).

- Quelle est la probabilité qu'il absorbe en une semaine plus de 82 mg de nicotine ?
- Quelle est la probabilité qu'il absorbe en une semaine moins de 79 mg de nicotine ?

Désignons par X le contenu en nicotine d'une seule cigarette

X suit une loi normale de moyenne $\mu = 0.8$ et d'écart-type $\sigma = 0.1$

On note $X \sim N(0.8, 0.1^2)$

Or, on s'intéresse au contenu en nicotine de 100 cigarettes.

On utilise la propriété d'additivité de la fonction gaussienne :

la loi normale est additive par rapport à ses deux paramètres μ et σ^2 .

On peut dès lors calculer :

$$\mu_{100} = 100 \mu$$

$$\sigma_{100}^2 = 100 \sigma^2$$





$$\begin{aligned}\text{On a : } \mu_{100} &= 100 \mu \\ &= 100 \cdot 0.8 \\ &= 80\end{aligned}$$

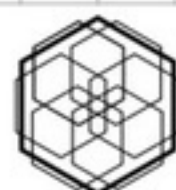
$$\begin{aligned}\sigma_{100}^2 &= 100 \sigma^2 \\ &= 100 \cdot 0.1^2 \\ &= 100 \cdot 0.01 \\ &= 1\end{aligned}$$

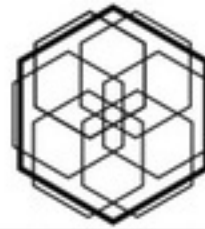
Le contenu en nicotine de 100 cigarettes suit donc une loi normale de moyenne $\mu_{100} = 80$ et d'écart-type $\sqrt{\sigma_{100}^2} = 1$

$$\text{On note } X_{100} \sim N(80, 1^2)$$

a) Quelle est la probabilité d'absorber plus de 82g de nicotine en fumant 100 cigarettes ?

$$\Leftrightarrow \text{Que vaut } P(X_{100} > 82) ?$$





$$\begin{aligned}
P(X_{100} > 82) &= 1 - P(X_{100} \leq 82) \\
&= 1 - P\left(\frac{X_{100} - \mu_{100}}{\sqrt{\sigma_{100}^2}} \leq \frac{82 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right) \\
&= 1 - P\left(X_{100}^* \leq \frac{82 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right) \\
&= 1 - P\left(X_{100}^* \leq \frac{82 - 80}{1}\right) \\
&= 1 - P(X_{100}^* \leq 2) \\
&= 1 - \Phi(2) \\
&= 1 - 0.9772 \\
&= \underline{\underline{0.0228}}
\end{aligned}$$

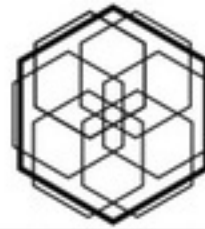
2.00



X	0	1	2	3	4	5	6	7	8	9
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936

Ainsi, la probabilité d'absorber plus de 82 g de nicotine en fumant 100 cigarettes est de 2.28 %





b) Quelle est la probabilité d'absorber moins de 79 g de nicotine avec 100 cigarettes ?

$\Leftrightarrow$ Que vaut $P(X_{100} \leq 79)$?

$$P(X_{100} \leq 79) = P\left(\frac{X_{100} - \mu_{100}}{\sqrt{\sigma_{100}^2}} \leq \frac{79 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right)$$

$$= P\left(X_{100}^* \leq \frac{79 - \mu_{100}}{\sqrt{\sigma_{100}^2}}\right)$$

$$= P\left(X_{100}^* \leq \frac{79 - 80}{1}\right)$$

$$= P(X_{100}^* \leq -1)$$

$$= 1 - P(X_{100}^* \leq 1)$$

$$= 1 - \Phi(1)$$

$$= 1 - 0.8413$$

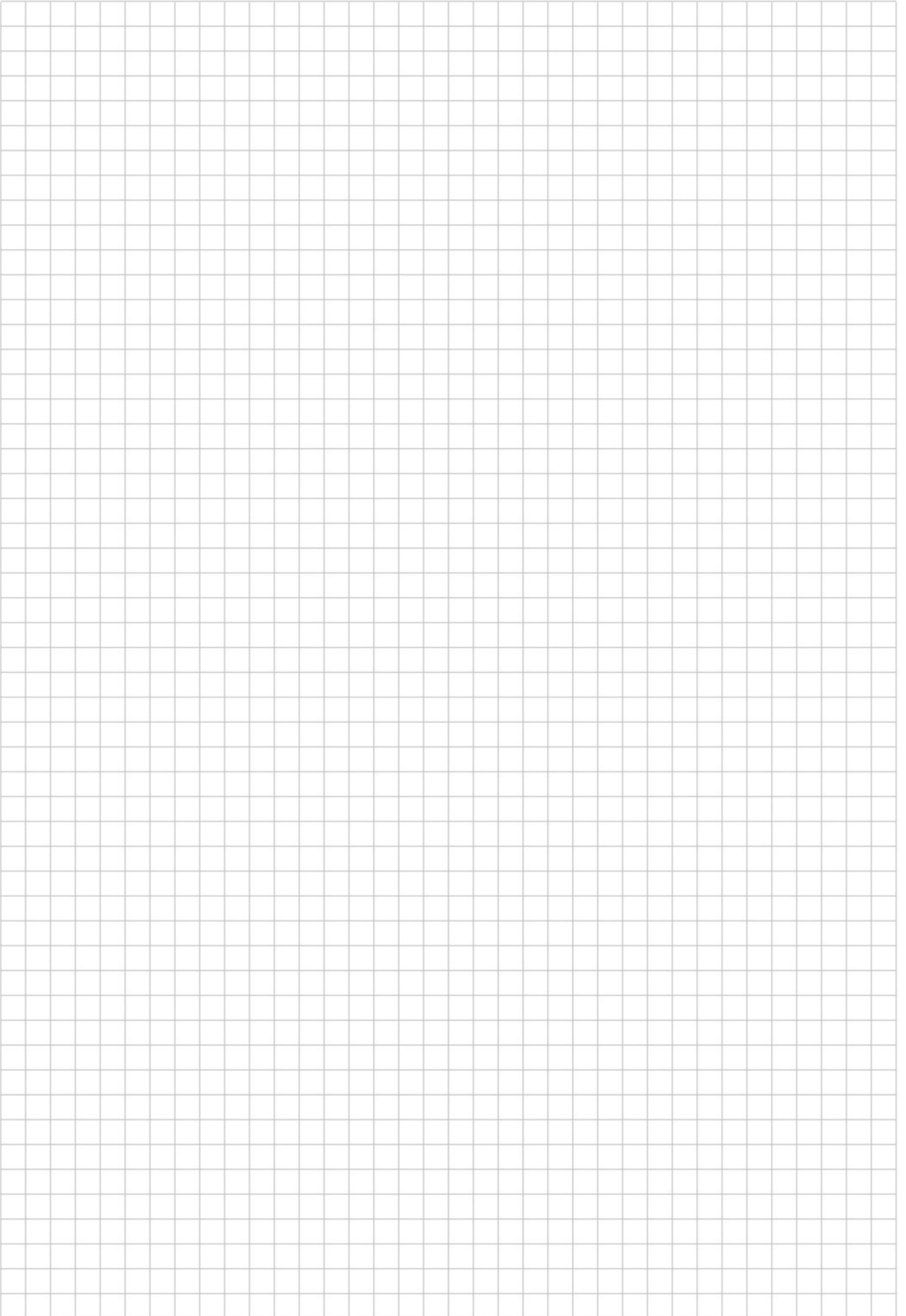
$$= \underline{\underline{0.1587}} = \underline{\underline{15.87\%}}$$

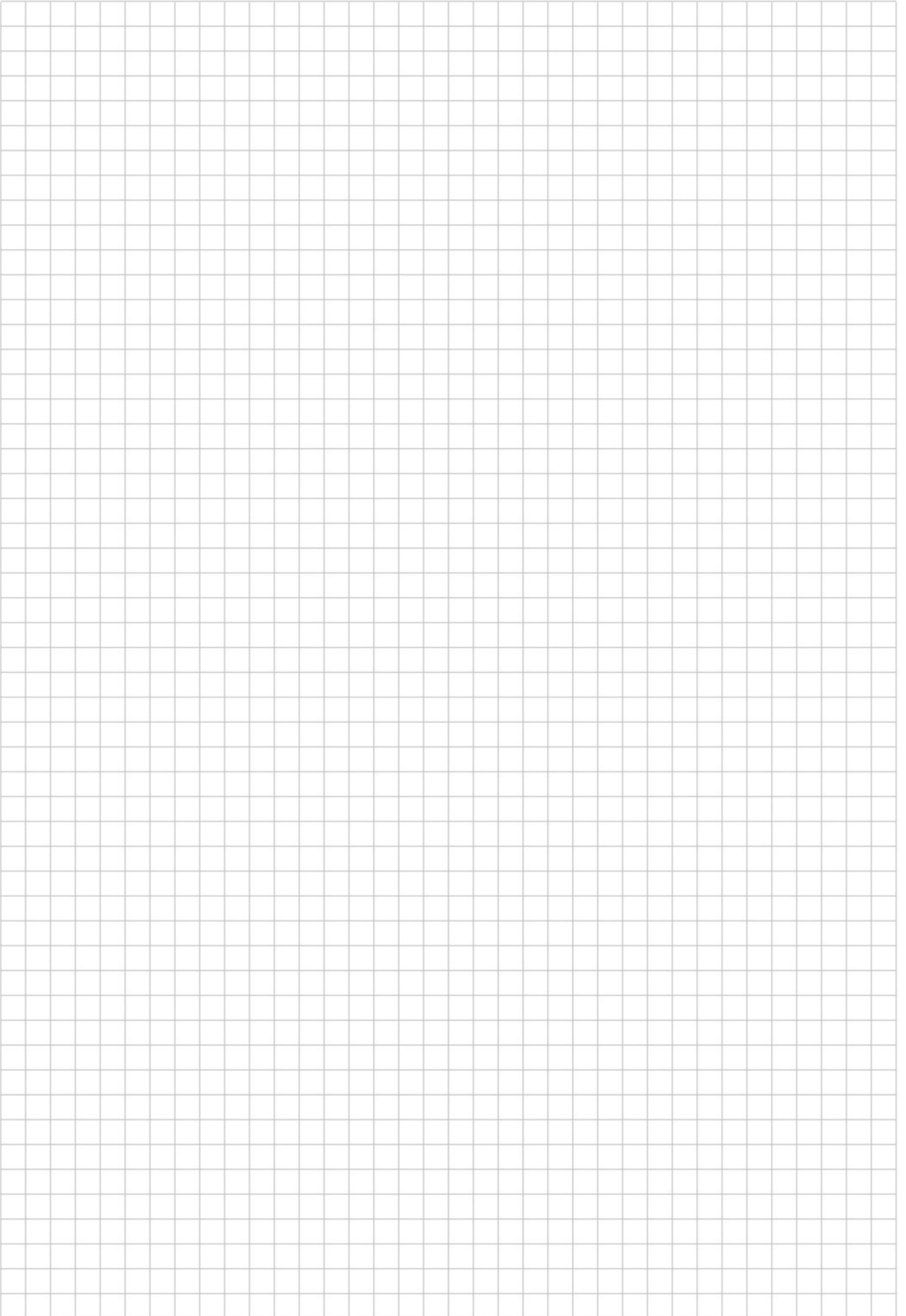
1



X	0	1	2	3	4	5	6	7	8	9
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319





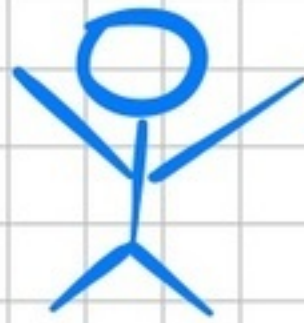




Les Statistiques

Expliquées Simplement

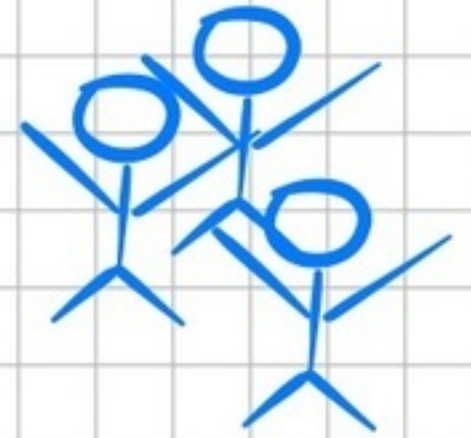
Cours Privés Individuels



- Progrès rapide
- Réponses précises
- Cours sur mesure

60 CHF / heure

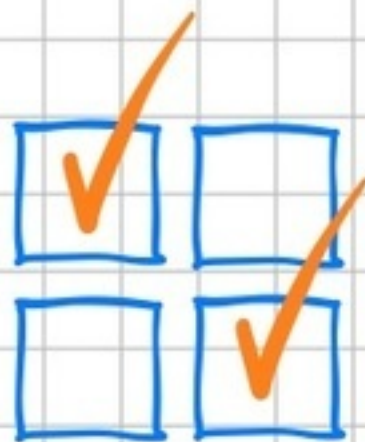
Cours Privés Collectifs



- Travail d'équipe
- Réponse aux questions
- Cours préparé d'avance
- Dès deux personnes

40 CHF / h / Pers.

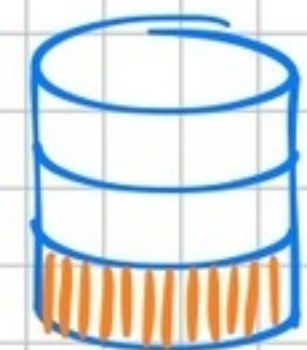
Modules à Choix



- Maîtriser 1 sujet précis
- Formation intensive
- Exercices pratiques
- Groupe de 5 pers. max
- Durée: 2 heures

40 CHF / heure

Pack 12 Heures



- Se mettre à jour
- Suivi personnalisé
- Méthode de travail
- Exces de renforcement
- **10h + 2h Offertes**

50 CHF / heure





Voici mes disponibilités

Genève UNIMAIL :

Lundi et Vendredi
14h00 - 18h00

Bern ANTHROPOLE :

Mardi, Mercredi et Jeudi
14h00 - 18h00

Visio TELEGRAM :

Lundi - Vendredi
9h00 - 11h00
18h00 - 20h00

Réservation Calendrier
(Min 2 jours à l'avance)



Horaires WEEKEND

J'ai la possibilité de mettre à disposition
des périodes durant le weekend si la
situation le demande.

Mes tarifs sont alors majoré de 50%

