Lois Continus

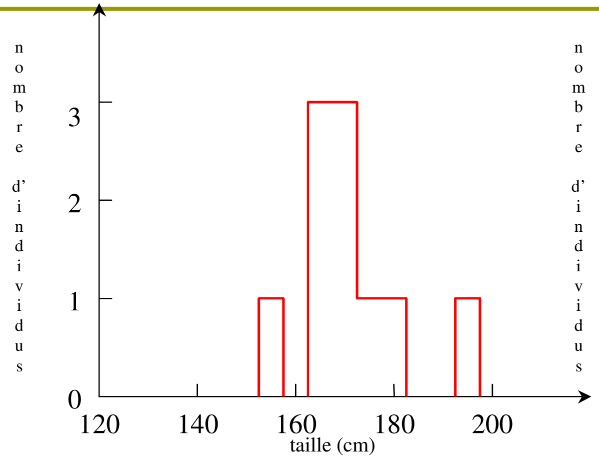
LOI NORMALE

LOI KHI 2

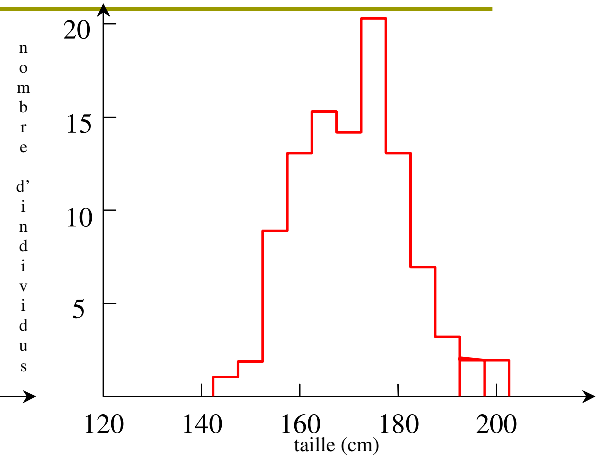
LOI STUDENT

LOI NORMALE

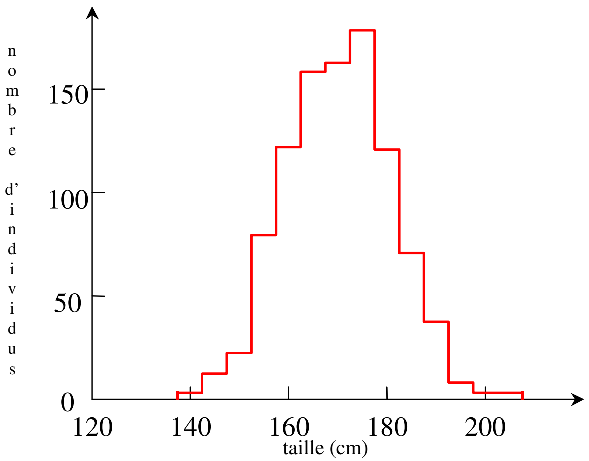
Supposons que nous tirions des échantillons aléatoires d'une population dont la taille moyenne est de 170 cm, avec un écart type de 10 cm.
Traçons l'histogramme de la taille



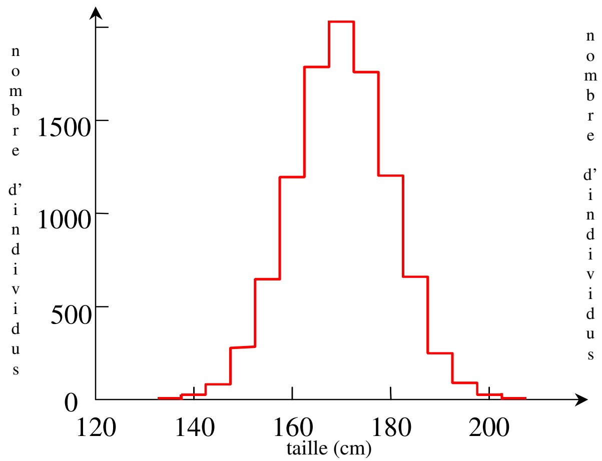
Echantillon de 10 individus



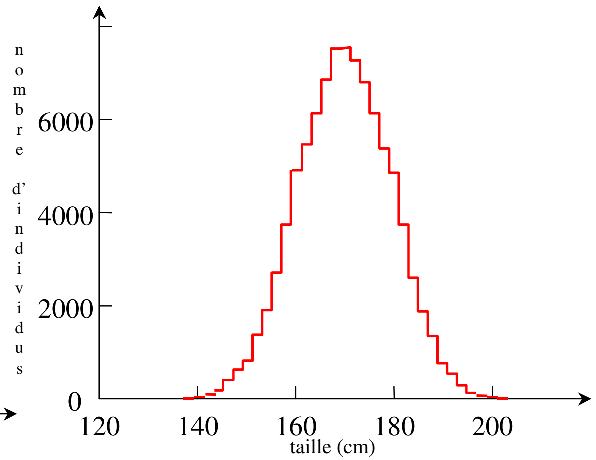
Echantillon de 100 individus



Echantillon de 1000 individus



Echantillon de 10.000 individus



Echantillon de 100.000 individus.

LOI NORMALE

- Au fur et à mesure que la taille de l'échantillon augmente (et que la taille des classes diminue), l'histogramme devient de plus en plus régulier et se rapproche d'une courbe en cloche, appelée loi normale.

LA LOI NORMALE

C'est possiblement la loi la plus importante de la théorie statistique. Pourquoi?

- Elle fournit une bonne description de certains phénomènes réels. Les distributions qui sont souvent proches de la normale:
  - Résultats de tests d'aptitude pris par plusieurs personnes (résultats d'examens, tests de QI).
  - Mesures répétées avec soin de la même quantité avec des équipements de laboratoires.
- Bonne approximation d'expériences impliquant un résultat au hasard, comme le lancer d'une pièce un très grand nombre de fois où l'on regarde le nombre de piles.

Distribution normale

- Une v.a.c X pouvant prendre toutes les valeurs réelles x dans l'intervalle de $-\infty$ à $+\infty$, pour $\mu \in \Re$, pour $\sigma \in \Re^+$ dont la fonction de densité est :

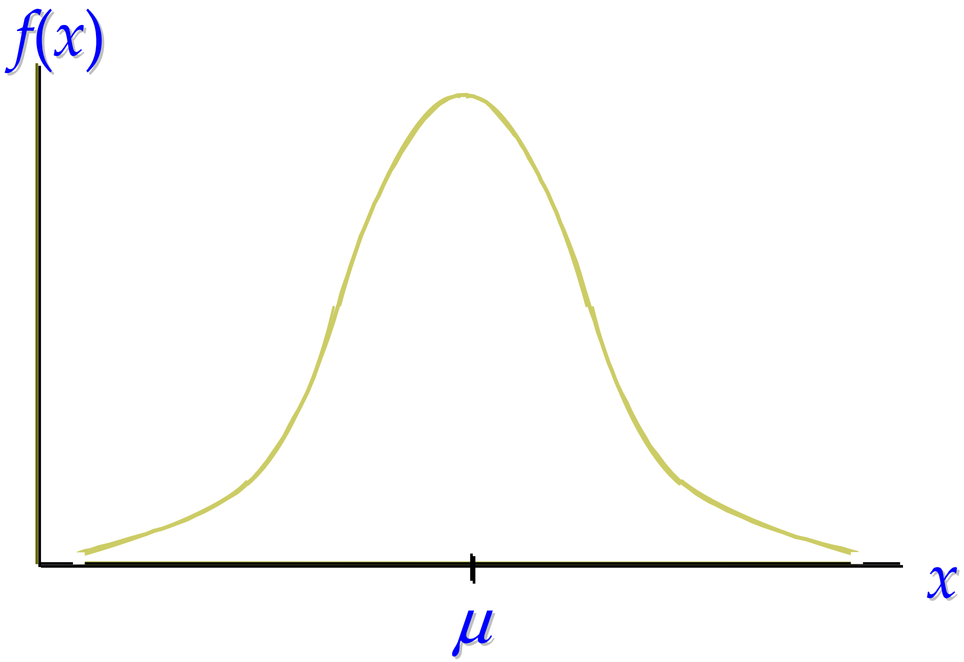
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

s'appelle une v.a. normale de paramètres μ et σ^2: $X \sim N(\mu, \sigma^2)$

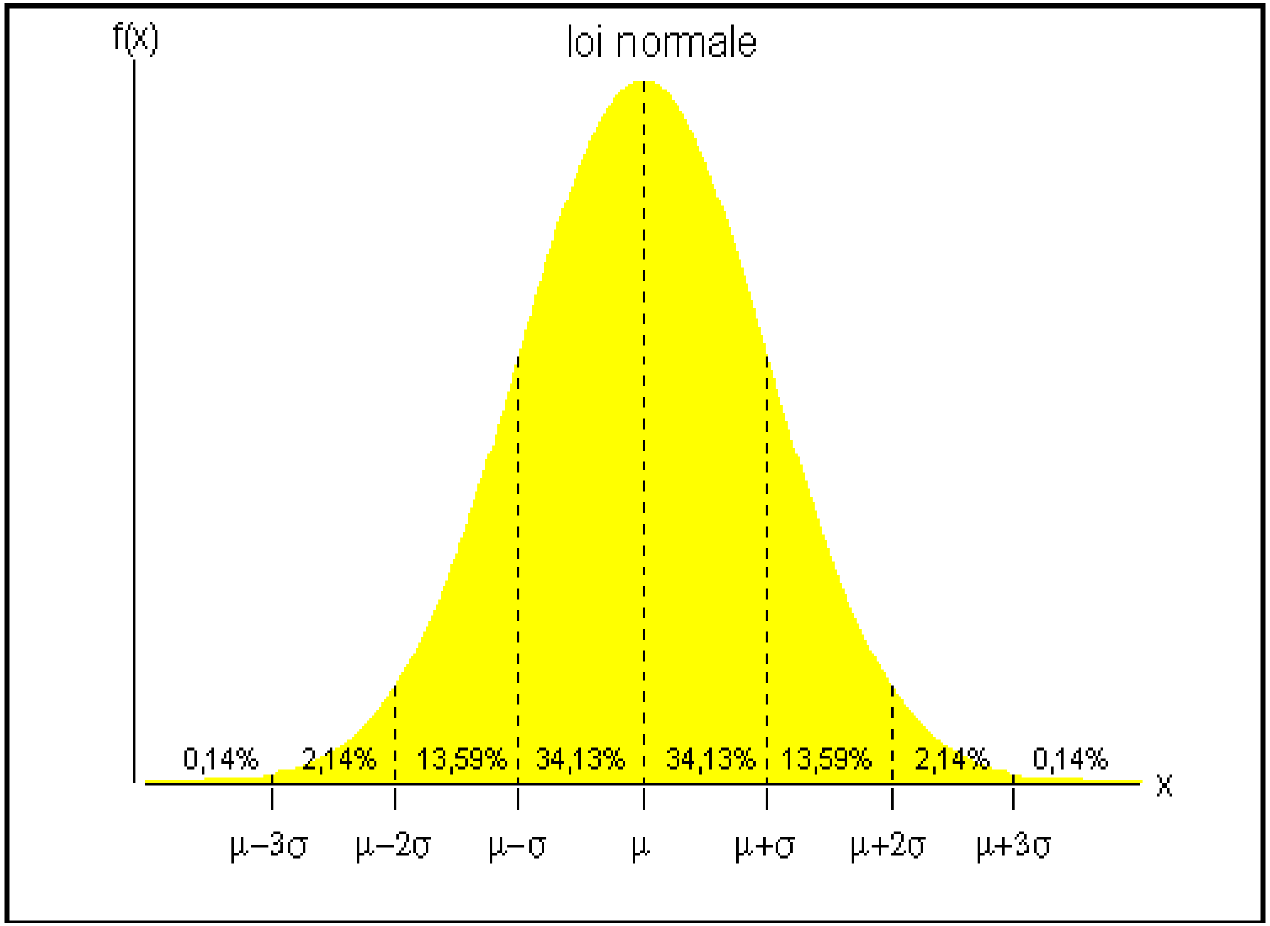
Forme de la distribution normale

- Il existe une famille entière de lois normales. Elles se différencient par leur moyenne et leur variance
- Courbe en cloche
- Courbe symétrique
- La moyenne, le mode et la médiane correspondent au même point (le point le plus élevé)

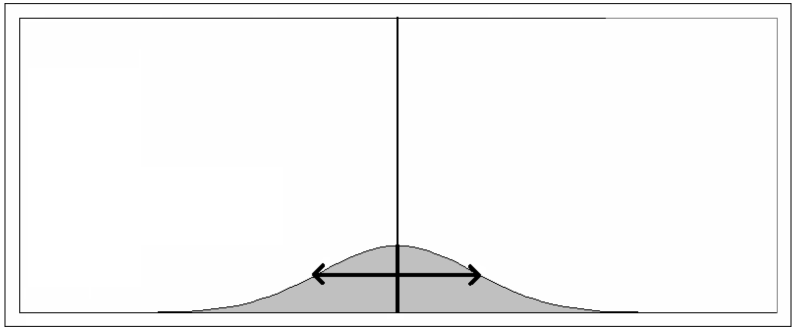
- L'écart type détermine la largeur de la courbe, plus il est grand, plus la courbe sera large et aplatie
- L'aire totale sous la courbe est 1
- Aussi appelée loi Gaussienne ou loi de Gauss



Représentation graphique

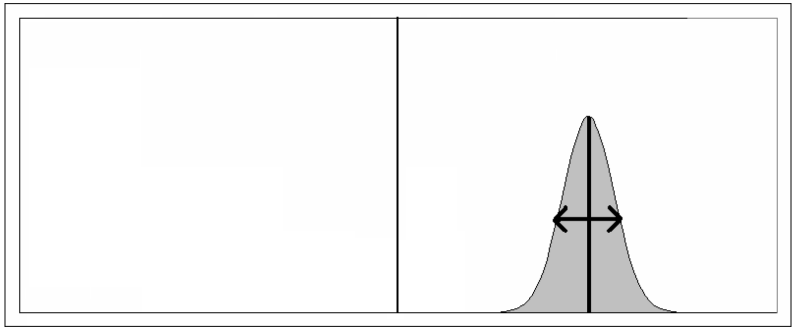


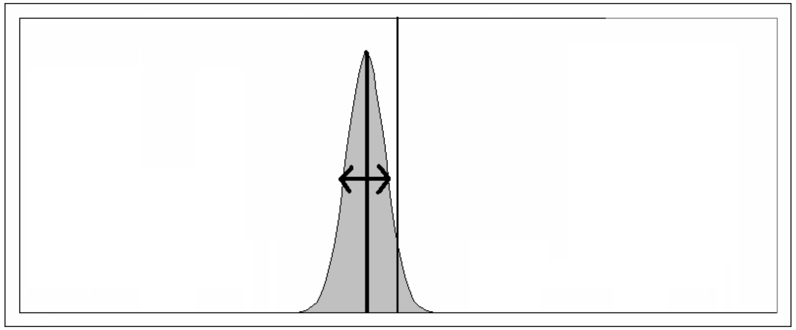
Exemples de lois normales



- Moyenne : 0
- Écart type : 3

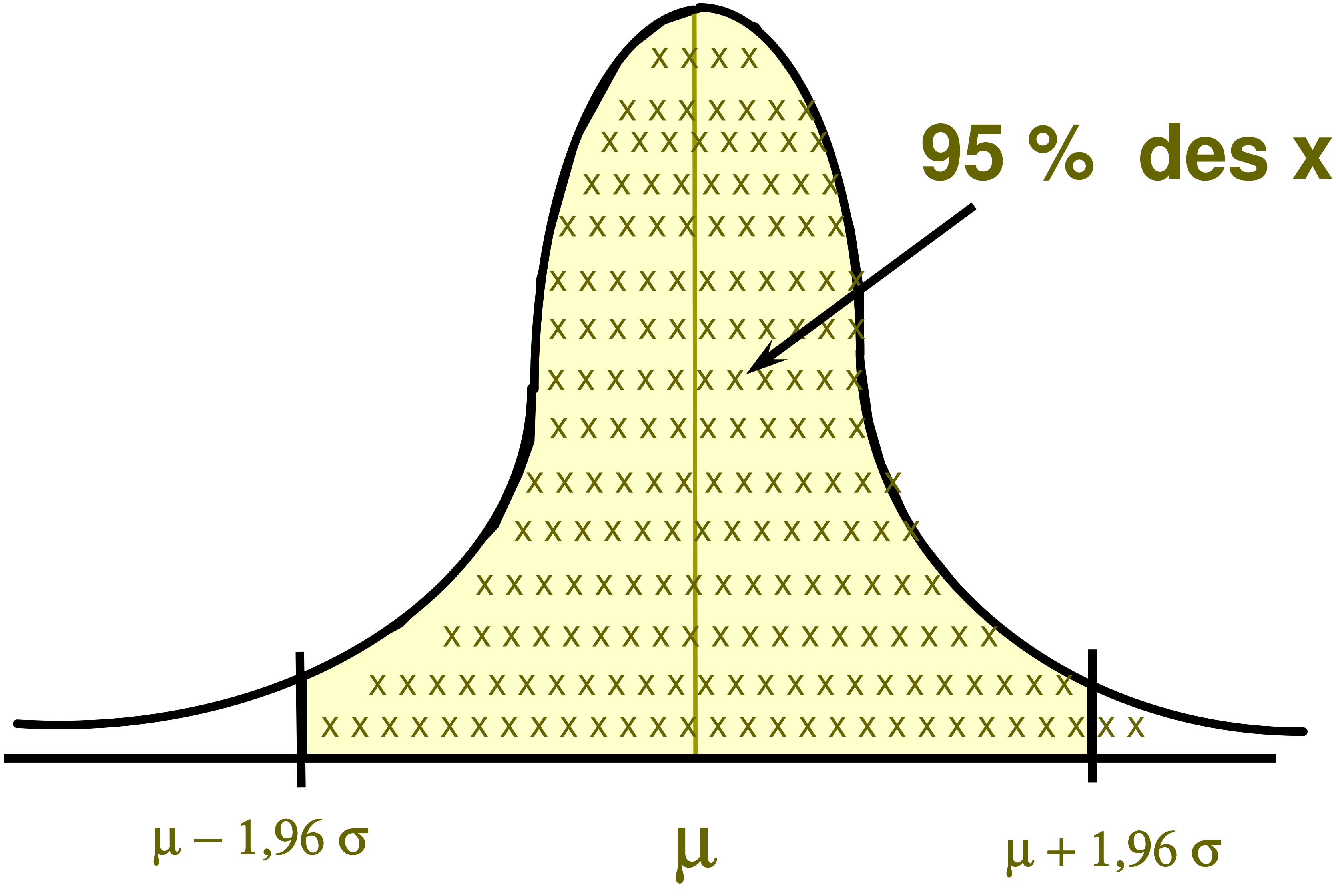
- Moyenne : 4
- Écart type : 1





- Moyenne : -1
- Écart type : 0,5

Cas particulier très utile



Distribution normale

- 68,26% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu-\sigma;\mu+\sigma]$
- 95,44% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu-2\sigma;\mu+2\sigma]$
- 99,72% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu-3\sigma;\mu+3\sigma]$

Distribution normale

$$E(X) = \mu$$

$$VAR(X) = \sigma^2$$

$$\int_a^b f(x)dx = P(a \leq x \leq b) \quad = F(b)\text{- } F(a)$$

- Pas facile de calculer l'aire sous la courbe
- On utilise la table de distribution normale centrée réduite

La loi normale centrée réduite

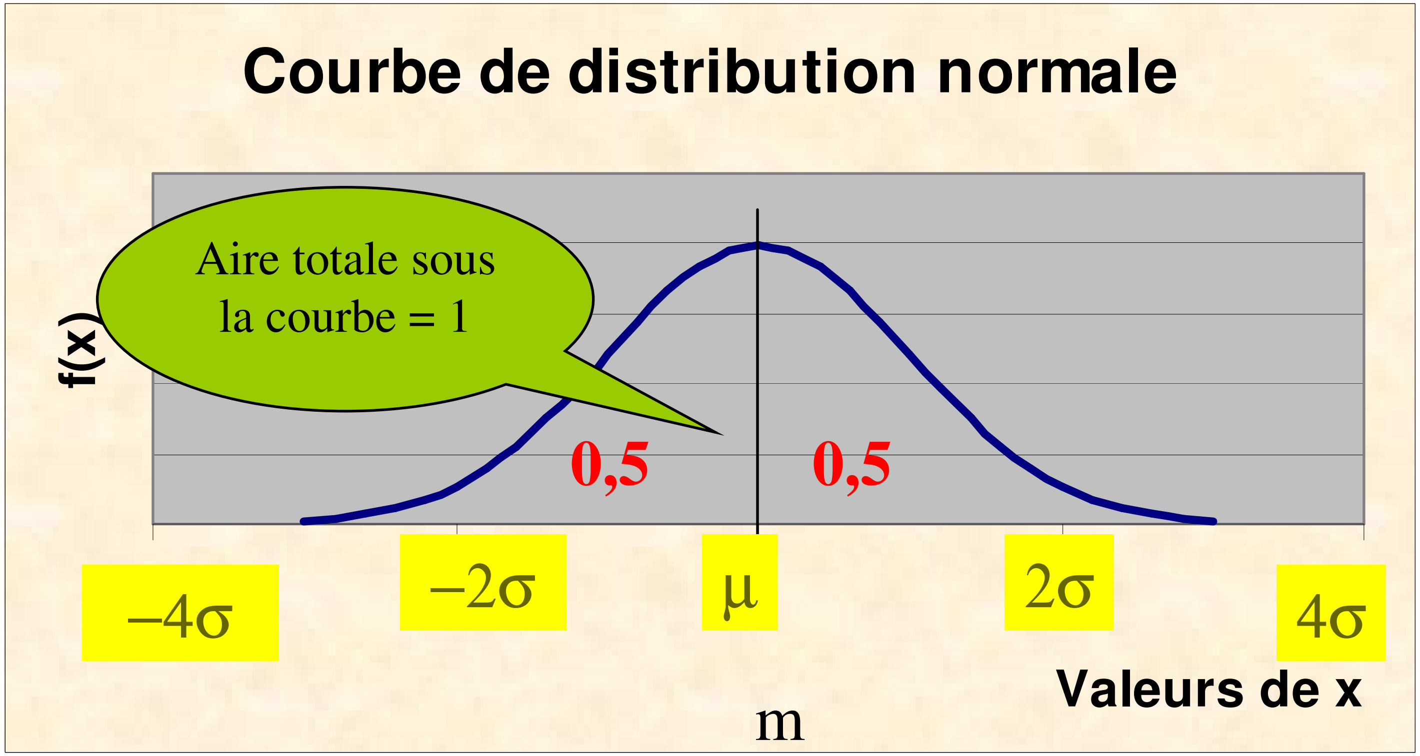
- Une v.a.c. qui a une distribution de probabilité normale de moyenne 0 et écart type 1, suit ce qu'on appelle une loi normale centrée réduite.
- Cette variable est souvent dénotée par la lettre Z
- On peut convertir une v.a.c. X qui suit une loi normale de moyenne μ et écart type σ en une variable normale centrée réduite Z :

$$Z = \frac{x - \mu}{\sigma}$$

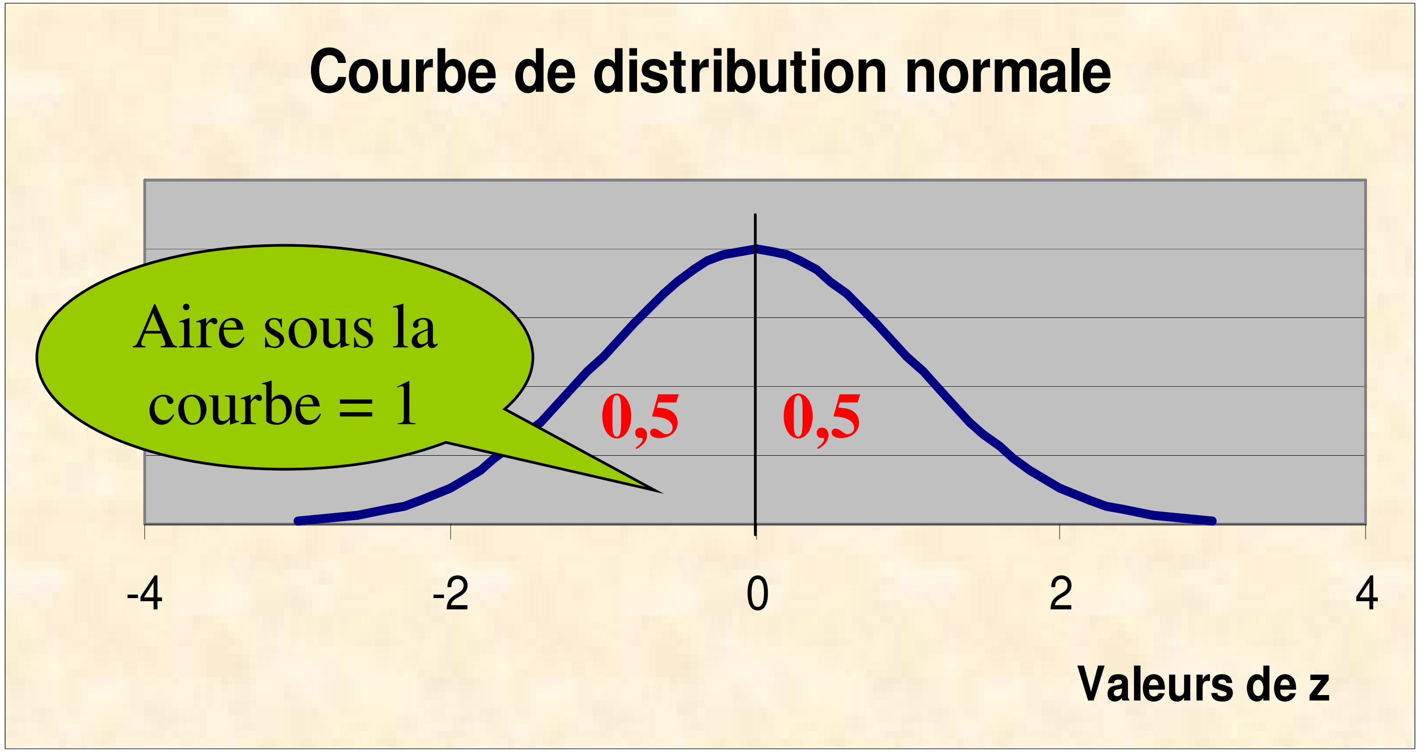
La loi normale centrée réduite

- Étant donné une valeur z, nous utilisons la table normale centrée réduite pour trouver la probabilité (l’aire sous la courbe) qui lui est associée.

La loi normale: N(μ,σ)



La loi normale centrée réduite N(0,1)



$$P(X < x) = P\left(\frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right) = P(Z < z)$$

Exemple

Un marchand vend de l’huile à moteur. Lorsque l’inventaire de l’huile à moteur descend à 20 gallons, une commande est effectuée auprès du fournisseur. Le gérant du magasin a remarqué qu’il perdait des ventes lorsqu’il était en attente du nouveau stock. Il avait déterminé que la demande lorsqu’on est en attente du nouveau stock suit une loi normale de moyenne de 15 gallons et un écart type de 6 gallons. Le gérant aimerait savoir quelle est la probabilité d’une rupture de stock?

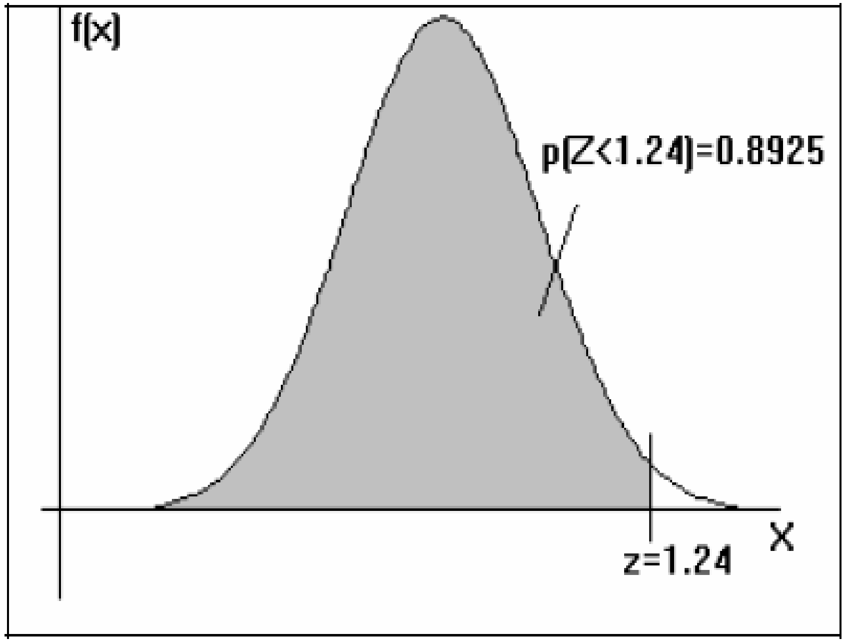
Exemple:

- X est la variable aléatoire qui représente la demande en gallons d'huile à moteur en période d'attente de stock
  - On cherche P(X > 20)

- Définissons: $Z = (X - \mu)/\sigma$
  - P(X > 20)=P(Z > (20 - 15)/6)= P(Z > 0,83)
  - On cherche donc: P(Z> 0,83)

Utilisation de la table numérique

Lecture de la table: Pour z=1.24 (intersection de la ligne 1.2 et de la colonne 0.04), on a la proportion P(Z < 1,24) = 0.8925



P(Z > 1,96) = 0,025
P(Z > 2,58) = 0,005
P(Z > 3,29) = 0,0005

Rappels:

1/ P(Z > z) = 1 - P(Z < z) et 2/ P(Z < -z) = P(Z > z)

Exemple: Sachant P(Z < 1,24) = 0,8925, on en déduit:

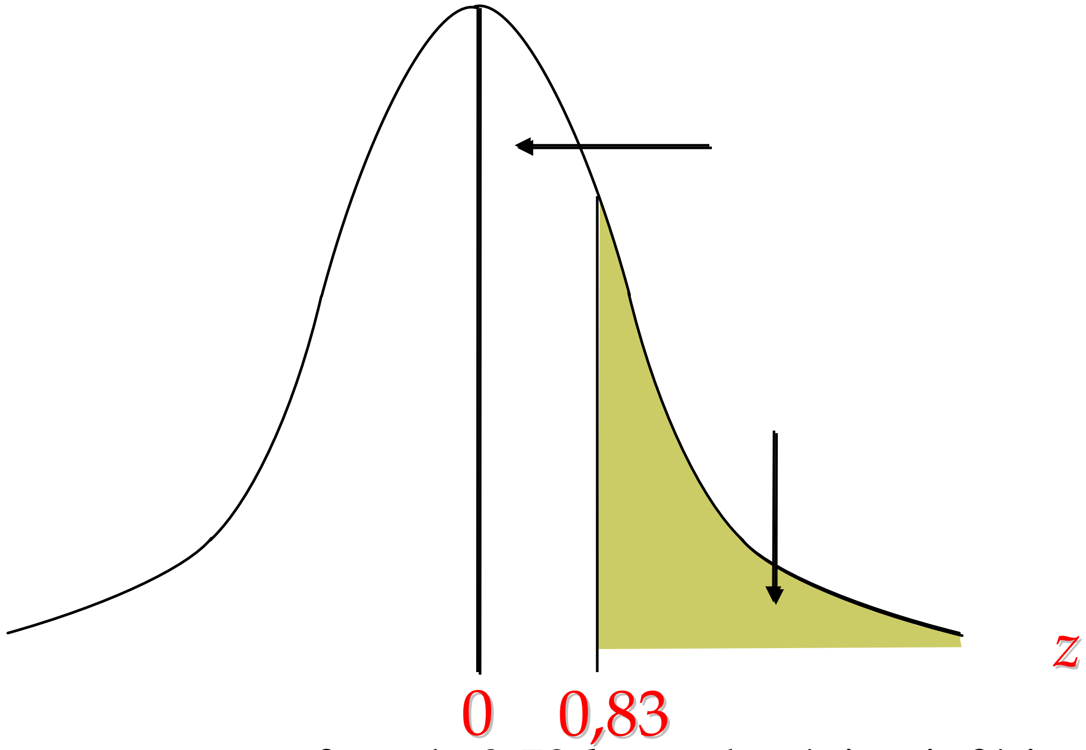
1/ (P(Z > 1,24) = 1 - P(Z < 1,24) = 1- 0,8925 = 0,1075

2/ P(Z < -1,24) = P(Z > 1,24) = 0,1075

Un extrait de la table numérique

	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	50000	50399	50798	51197	51595	51994	52392	52790	53188	53586
0,1	53983	54380	54776	55172	55567	55962	56356	56749	57142	57535
0,2	57926	58317	58706	59095	59483	59871	60257	60642	61026	61409
0,3	61791	62172	62552	62930	63307	63683	64058	64431	64803	65173
0,4	65542	65910	66276	66640	67003	67364	67724	68082	68439	68793
0,5	69146	69497	69847	70194	70540	70884	71226	71566	71904	72240
0,6	72575	72907	73237	73565	73891	74215	74537	74857	75175	75490
0,7	75804	76115	76424	76730	77035	77337	77637	77935	78230	78524
0,8	78814	79103	79389	79673	79955	80234	80511	80785	81057	81327
0,9	81594	81859	82121	82381	82639	82894	83147	83398	83646	83891
1,0	84134	84375	84614	84849	85083	85314	85543	85769	85993	86214
1,1	86433	86650	86864	87076	87286	87493	87698	87900	88100	88298
1,2	88493	88686	88877	89065	89251	89435	89617	89796	89973	90147
1,3	90320	90490	90658	90824	90988	91149	91309	91466	91621	91774
1,4	91924	92073	92220	92364	92507	92647	92785	92922	93056	93189
1,5	93319	93448	93574	93699	93822	93943	94062	94179	94295	94408
1,6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449

Exemple



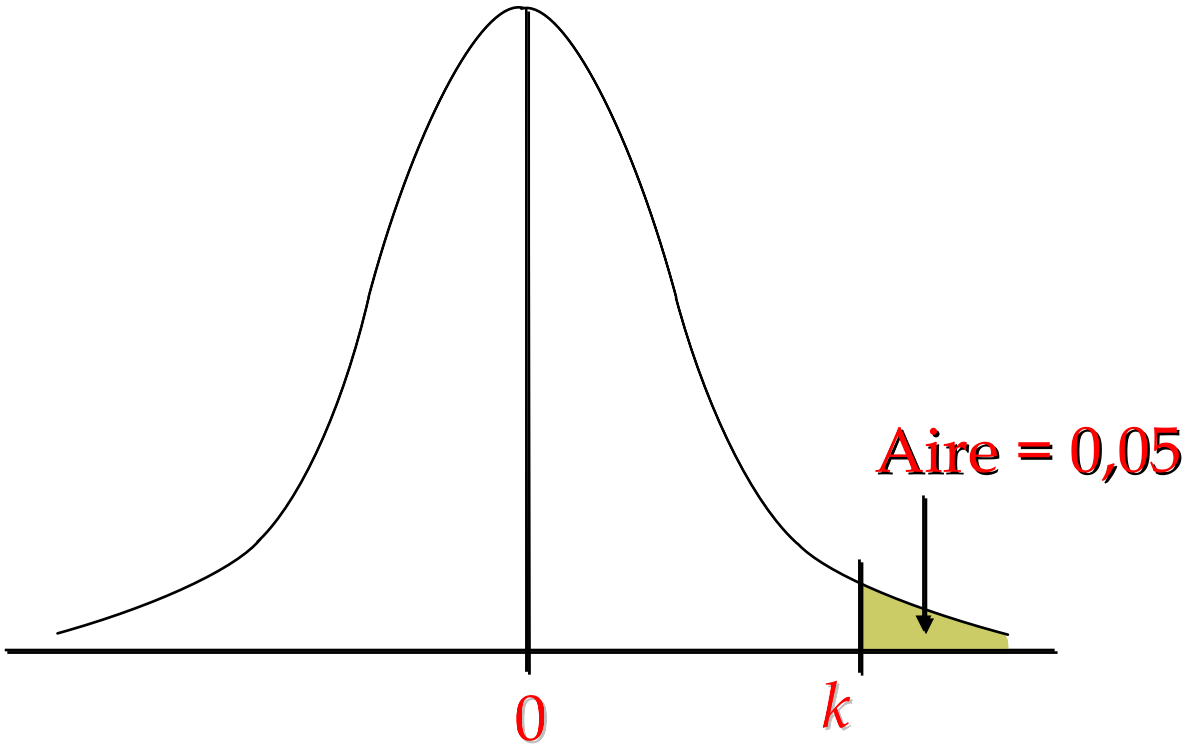
- La table montre une surface de 0,796 pour la région inférieure à z = 0,83.
- La région de l'extrémité en vert correspond donc à 1-0.79673 = 0,203.
- La probabilité d'une rupture de stock est 0,203

Exemple

Si le gérant veut que la probabilité de rupture de stock ne dépasse pas 0.05, à quel niveau d'inventaire devrait-il passer une nouvelle commande?

Exemple:

On cherche k tel que $P(Z>k) \leq 0{,}05$



Exemple:

- La valeur correspondante de x est:

$$x = \mu + k\sigma$$
$$= 15 + 1{,}645(6)$$
$$= 24{,}87$$

- Une commande de 24,87 gallons rendra la probabilité d'une rupture de stock égale à 0,05.
- Donc, le gérant du magasin devrait passer une nouvelle commande lorsque il lui reste 25 gallons pour garder cette probabilité sous 0,05.

Exemple

- Une machine fabrique des rondelles de métal dont le diamètre est distribué normalement avec une moyenne de 2,4 cm et un écart type de 0,05 cm. Des rondelles produites par cette machine, trouvons la proportion de celles dont le diamètre:

  i) excède 2,5 cm;

Rép. 0,0228

  ii) n'excède pas 2,32 cm;

Rép. 0,0548

  iii) est compris entre 2,35 cm et 2,46 cm.

Rép. 0,7262

iv) Trouvons la valeur de x telle que 5 % des rondelles présentent un diamètre qui lui est supérieur.

Rép. 2,48225

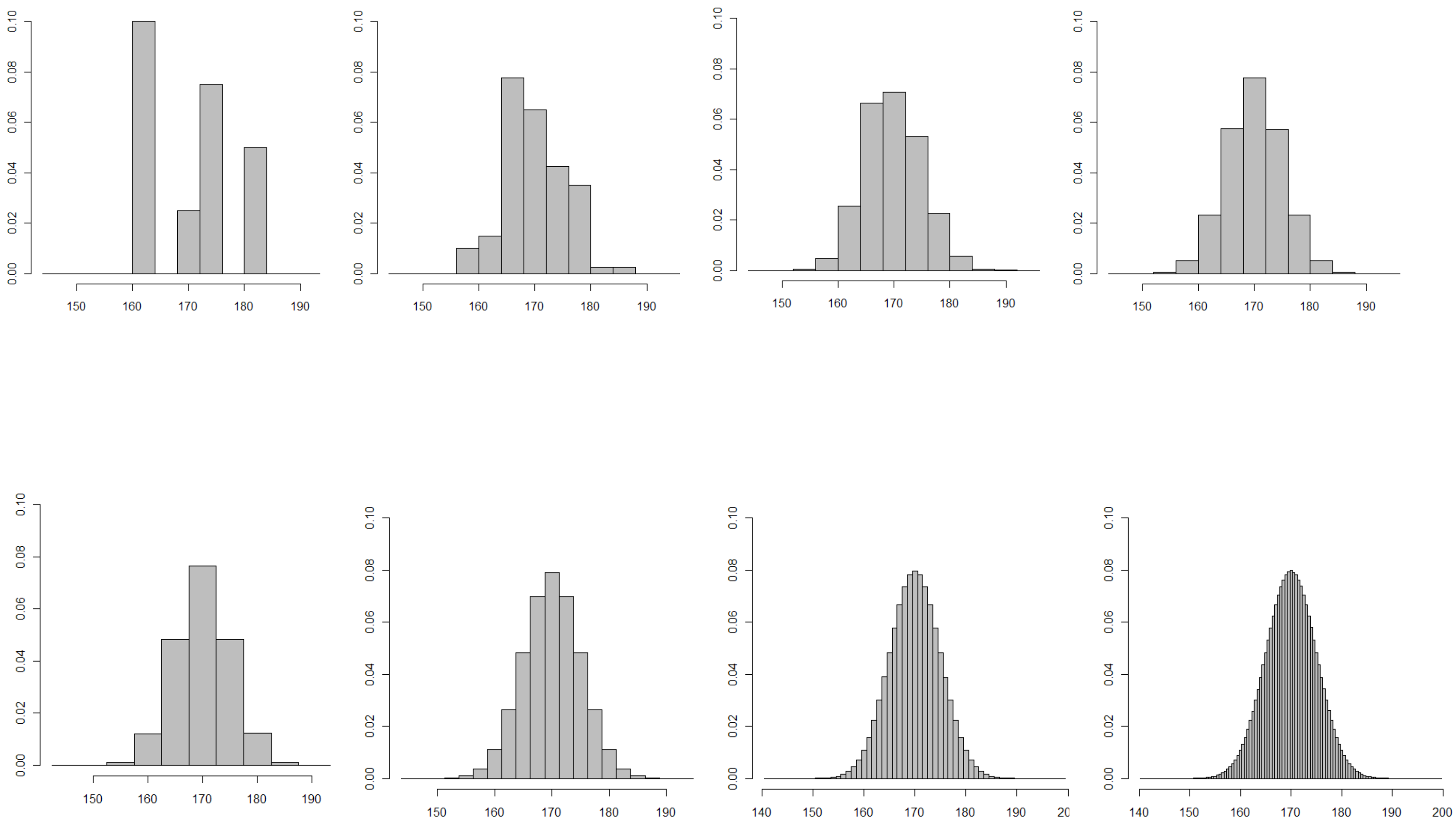
Théorème central-limite

- La loi Normale est une « loi limite » pour toutes les lois.

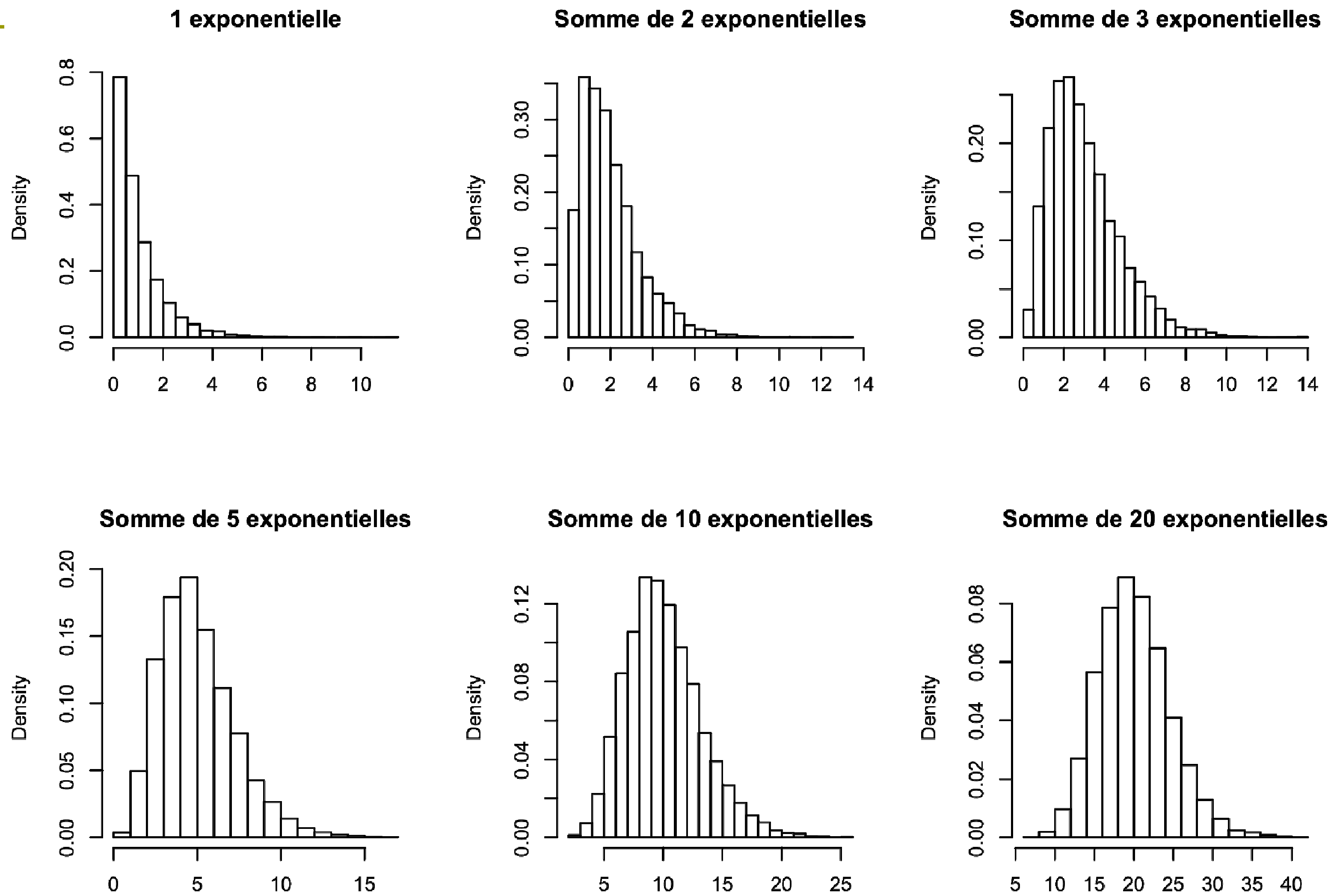
- Soient $X_1, X_2, \ldots, X_n$, n variables aléatoires indépendantes et de même type de loi.

  Si n est assez grand, la somme $Y = X_1 + \ldots + X_n$ suit approximativement une **loi normale** de paramètres
  - moyenne = somme des moyennes des X_i
  - variance = somme des variances des X_i

CONVERGENCE VERS UNE LOI



Exemple de loi d'une somme de v.a.i



Approximations

- Approximation normale de la loi binomiale

  La loi binomiale $B(n\ ;\ p)$ suivra à peu prés la loi normale de moyenne np et de variance $np(1-p)$ si $n>30$, $np> 10$ et $np(1-p)>10$.

- Approximation normale de la loi de Poisson

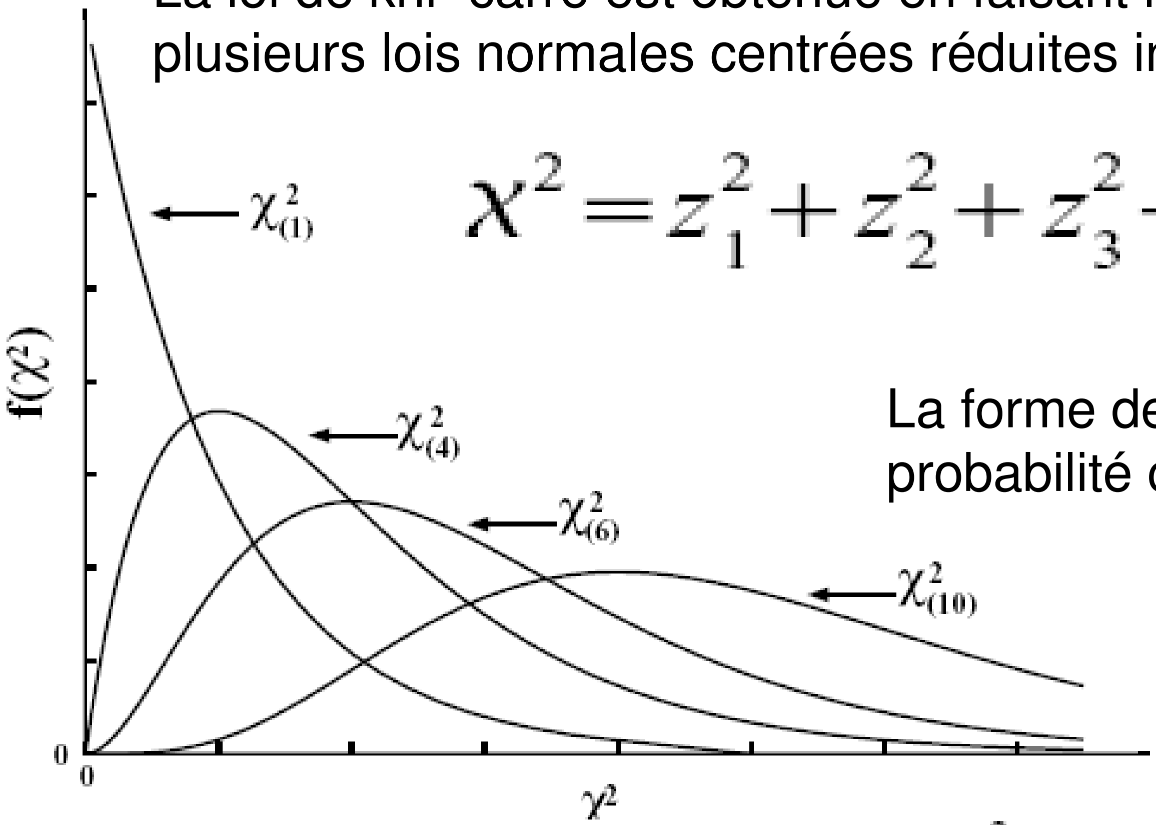
  On a $P(m) = N(m\ ;\ m)$ si $m>20$

La loi du χ^2

Définition :

La loi de khi–carré est obtenue en faisant la somme des carrés de plusieurs lois normales centrées réduites indépendantes :

$$\chi^2 = z_1^2 + z_2^2 + z_3^2 + \ldots + z_\nu^2$$



La forme de la courbe de densité de probabilité change en fonction de υ

Distribution de densité des lois de χ^2 à $\nu=1$, $\nu=4$, $\nu=6$ et $\nu=10$ degrés de liberté.

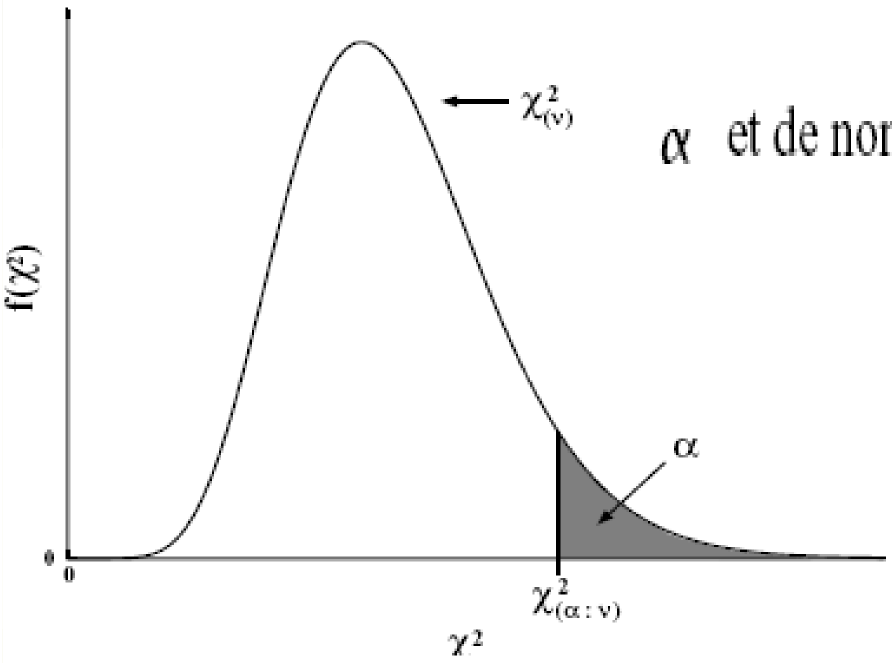
Propriétés

- Si X suit la loi de Khi-2 à n degrés de liberté alors son espérance et sa variance sont définis par:

$$E(X) = n$$

$$Var(X) = 2n$$

Table de la loi de khi–carré $\chi^2_{(\alpha:\nu)}$



α et de nombre de degrés de liberté ν :

$$P\left(\chi^2_\nu > \chi^2_{(\alpha:\nu)}\right) = \alpha$$

La probabilité donnée dans la table est donc ***unilatérale à droite***.

Exemples d'utilisation de la table

- Soit X une variable aléatoire qui suit la loi Khi-2 à 5 degrés de liberté, déterminer a telle que:

$P(X<a)=0,5\%$; $P(0<X<a)=95\%$;

$P(-a<X<a)=97,5\%$; $P(|X|>a)=2\%$.

Solution:

$a=0.412$; $a=11.07$; $a=12.83$; $a=13.388$

La loi de Student ou loi de *t*

- La distribution de Student s'apparente à la distribution de la loi Normale. Elle varie cependant avec la valeur du degré de liberté ν (nu). Si ν est grand ($\nu \geq 30$) alors la distribution de Student est très proche de la distribution normale centrée réduite.
- De façon générale on dira qu'une variable aléatoire continue T suit une loi de Student de degré ν de liberté si :

$$T = \frac{Z}{\sqrt{X / \nu}}$$

où Z est une variable aléatoire distribuée selon une loi normale centrée réduite et X est une variable aléatoire suivant une loi du Khi carré.

L'espérance mathématique de T est : **E(*T*) = 0 si** ν **>1**

La variance de T est : $Var(T) = \dfrac{\nu}{\nu - 2}$

L'objectif ici sera d'apprendre à utiliser la table de t.

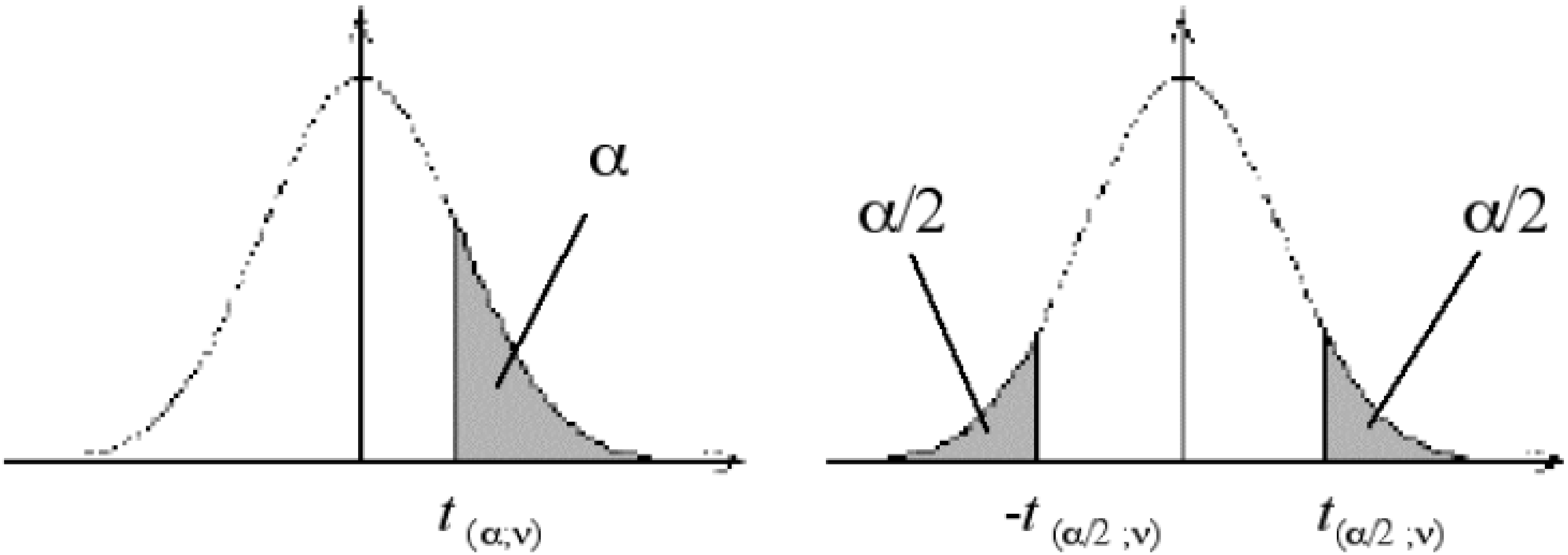
Description de la table

- <u>Description de la table</u>

  Les valeurs dans la table sont des valeurs de *t* <u>*et non* </u>des surfaces sous la courbe. On note ces valeurs $t(\alpha: \nu)$. La table ne donne que les valeurs positives car la distribution de *t* est symétrique.
- Les *valeurs* de la table sont des limites définies sur *l'abscisse* de la courbe.
- Les *probabilités* α ou 1- α/2 (= les deux lignes d'en-tête du tableau)sont des *surfaces* sous la courbe.
- La distribution change en fonction du nombre de *degrés de liberté* ν. Lorsque ν tend vers l'infini, la courbe de *t* converge vers une courbe normale centrée réduite.

Le **seuil α** correspond à **P(|*t*| > *t*** $_{(\alpha/2;\upsilon)}$), c'est-à-dire la probabilité que *t* égale ou dépasse une certaine *valeur critique*, définie en fonction du seuil de probabilité et du nombre de degrés de liberté.
Attention, le seuil peut être **unilatéral** ou **bilatéral**!!!

Si le seuil est bilatéral, la notation est la suivante: $P(|t| > t_{(\alpha/2;\upsilon)})$



Exemples d'utilisation de la table de student

- En utilisant la table de la loi de Student, trouver la valeur de *c* telle que :
  - $P(T>c)=0,01$, si n = 20, n degré de liberté
  - $P(T>c)=0,25$, si n = 1
  - $P(T>c)=0,005$, si n = 10
  - $P(T>c)=0,05$, si n = 200
- Toujours à l'aide de la table de Student, donner la valeur de :
  - $P(T>2,0796)=$? si n = 21
  - $P(T<1,7959)=$? si n = 11
  - $P(0,6844<T<1,7081)$ = ? si n = 25
- **Solutions**

1) 2.5280; 1; 3.1693

2) 2.5%; 95%; 20%