Ch 4 : Quelques distributions usuelles

I. Quelques Lois continues

I. 1 La loi normale (Laplace Gauss)

La loi normale est le modèle le plus utilisé en statistique. Elle possède l'avantage d'approcher des distributions de sommes de variables indépendantes. Elle fut découverte par Karl Friedrich Gauss (1777-1855).

Définition : On dit qu'une variable aléatoire X suit une loi normale de paramètres μ et σ^2 , si sa densité est de la forme :

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

La loi normale de paramètres μ et σ^2 est notée:

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

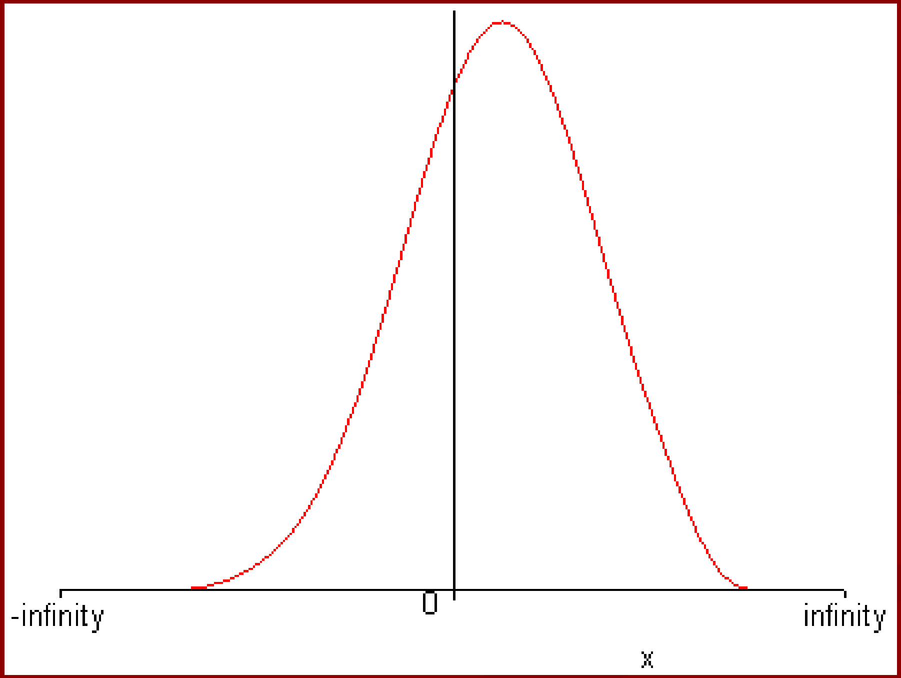
Résultats:

Soit X une variable aléatoire distribuée selon une loi normale de paramètres μ et σ^2 : $X \approx N(\mu,\sigma^2)$.

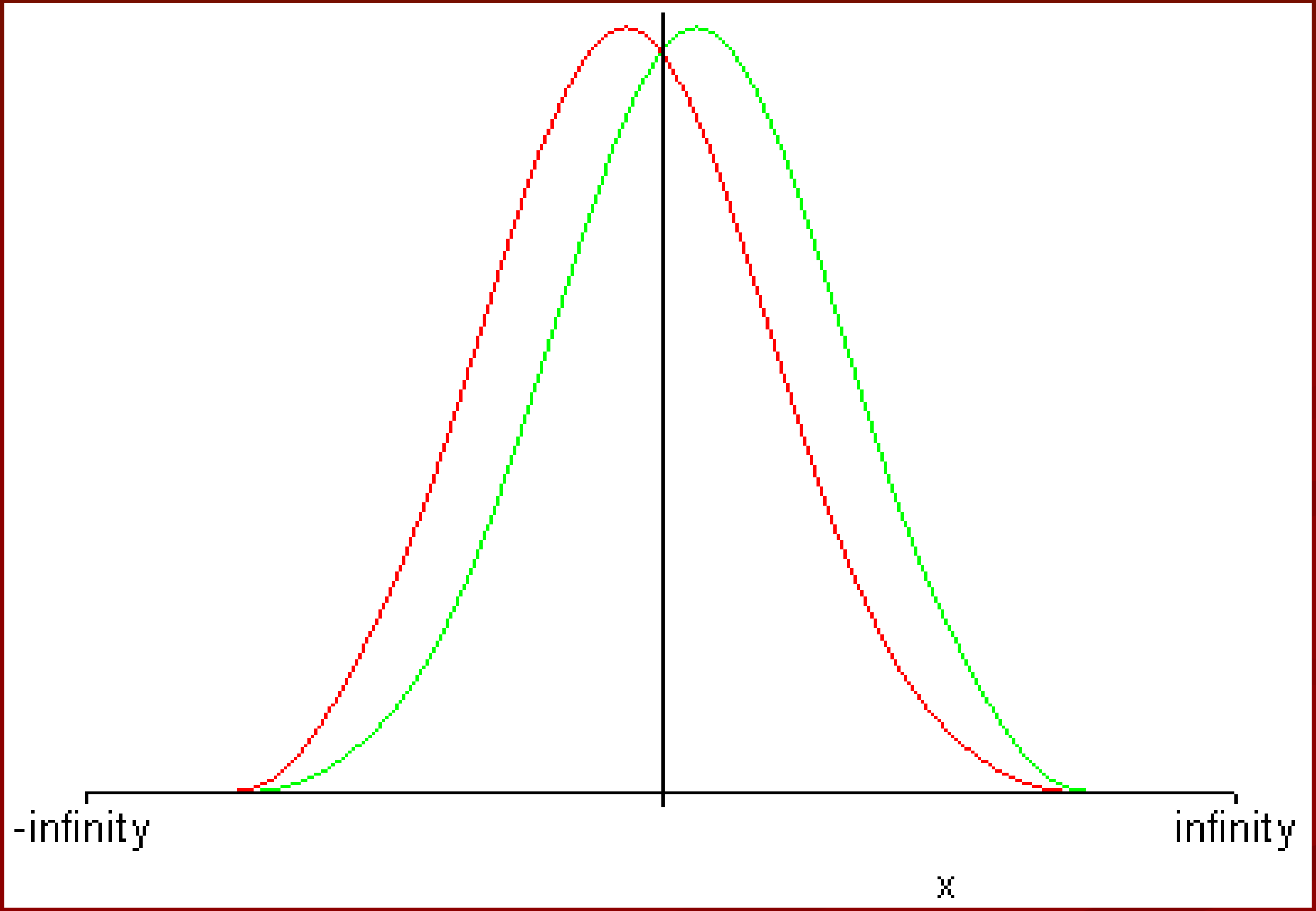
Alors on montre que:

$$E(X) = \mu \qquad \text{et} \qquad Var(X) = \sigma^2$$

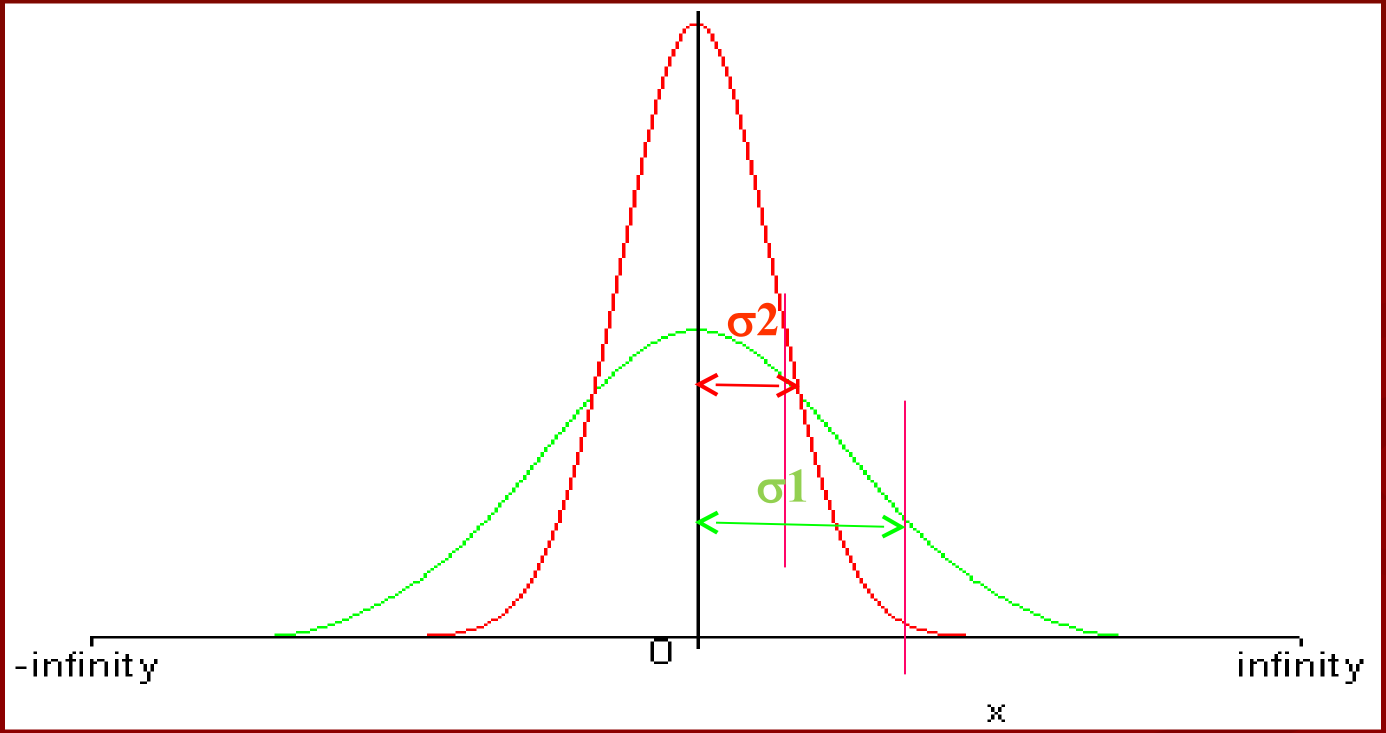
I. 1.1 Graphe d'une loi normale : La courbe a une forme de cloche. Elle est symétrique par rapport à la droite verticale passant par la moyenne. L'écart type représente la distance entre l'axe de symétrie et le point d'inflexion de la courbe (changement de courbure).



Effet de la moyenne μ: lois normales de variances égales mais de moyennes différentes.



Effet de la variance σ^2 : lois normales de même moyenne mais de variances différentes.



I.1.2 Loi normale centrée réduite

La loi normale de moyenne $\mu = 0$ et de variance $\sigma^2 = 1$ est appelée loi normale centrée réduite.
La variable aléatoire qui suit une normale centrée réduite.

Sa densité est alors:

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

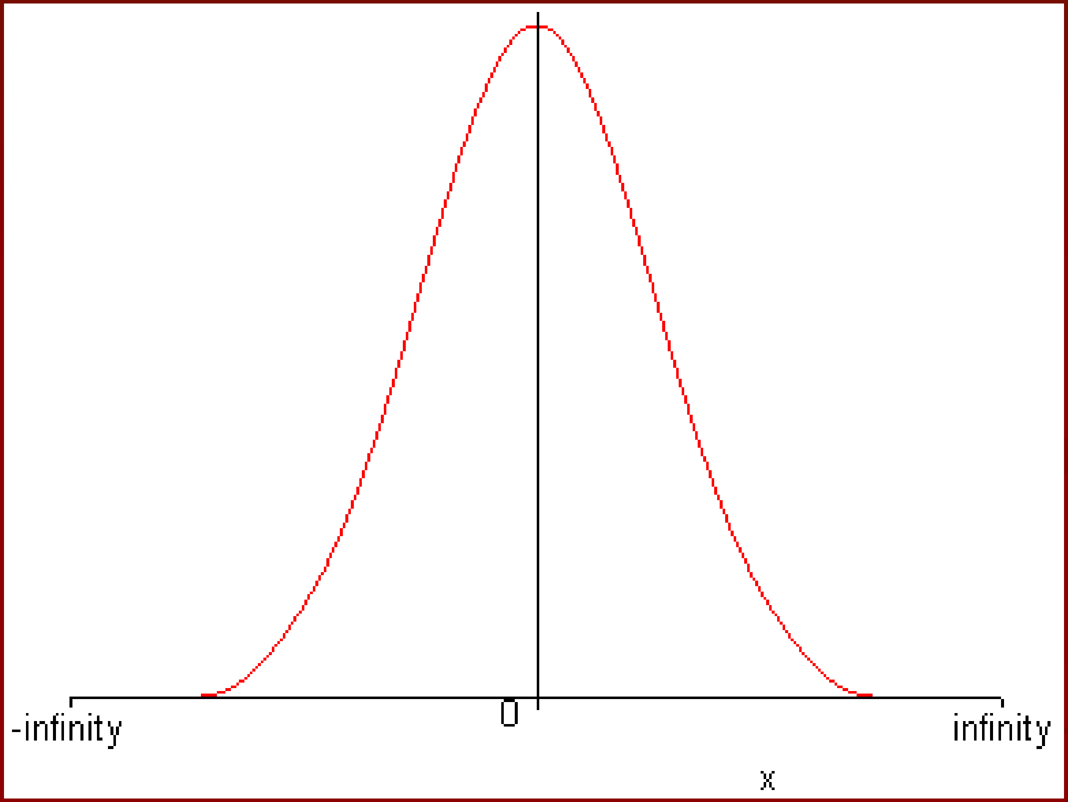
I.1.3 Fonction de répartition de la loi normale centrée réduite

La fonction de répartition de la loi normale **N(0,1)** nous permet le calcul des probabilités par la table de la loi normale. Elle est définie par :

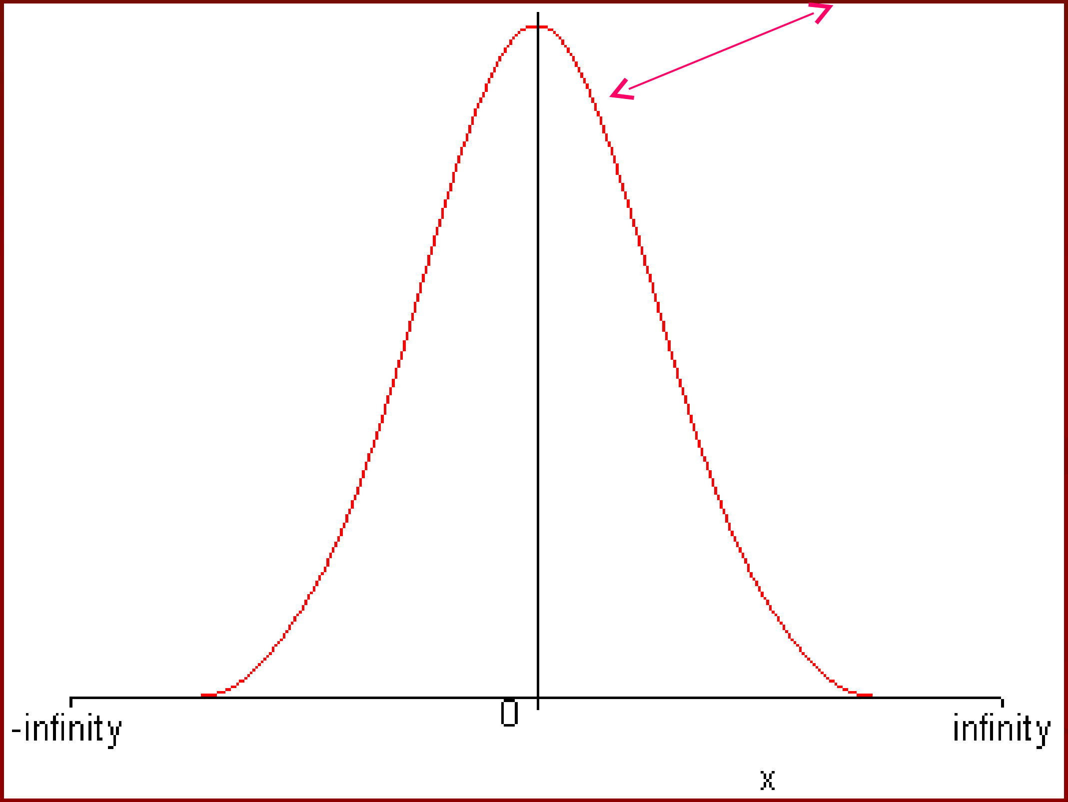
$$\phi(x) = P(Z \leq x)$$

où $Z \approx N(0,1)$,

et représente la surface, située à gauche de x, entre la courbe de densité et l'axe des abscisses :


-infinity
0
infinity
x

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



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

alors $$\frac{X-\mu}{\sigma} \sim N(0,1)$$

X v.a suivant N(m, σ^2) et Z v.a suivant N(0;1)

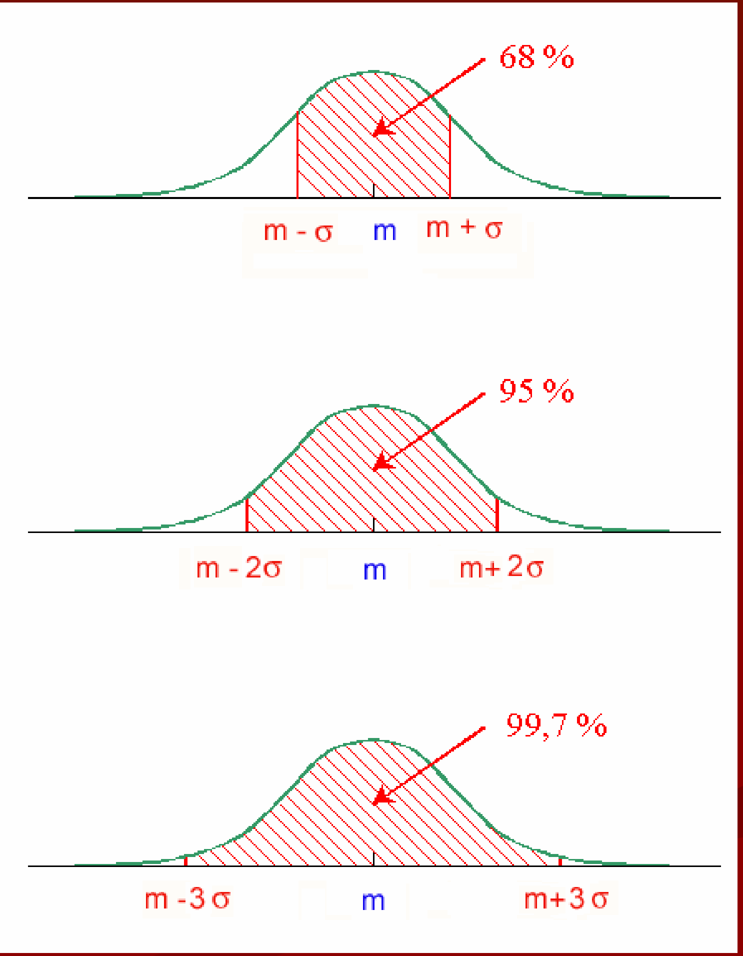
Les intervalles suivants ou plages de normalité se calculent grâce aux égalités ci-dessous, obtenues grâce à la table de la loi N(0;1) :

$P(m-\sigma < X < m+ \sigma) = P(-1 < Z < +1) = 0{,}68.$

$P(m-1{,}6\,\sigma < X < m+1{,}6\,\sigma) = P(-1{,}6 < Z < +1{,}6) = 0{,}90.$

$P(m-1{,}96\,\sigma < X < m+1{,}96\,\sigma) = P(-1{,}96 < Z < +1{,}96) = 0{,}95.$

$P(m-3{,}09\,\sigma < X < m+3{,}09\,\sigma) = P(-3{,}09 < Z < +3{,}09) = 0{,}99.$



I.2. Loi exponentielle

Une loi exponentielle modélise la durée de vie d'un phénomène sans mémoire, ou sans vieillissement, ou sans usure : la probabilité que le phénomène dure au moins s + t heures sachant qu'il a déjà duré t heures sera la même que la probabilité de durer s heures à partir de sa mise en fonction initiale. En d'autres termes, le fait que le phénomène ait duré pendant t heures ne change rien à son espérance de vie à partir du temps t.

Plus formellement, soit X est une variable aléatoire définissant la durée de vie d'un phénomène, d'espérance mathématique . On suppose que :

$$P(X > t + s \mid X > t) = P(X > s)$$

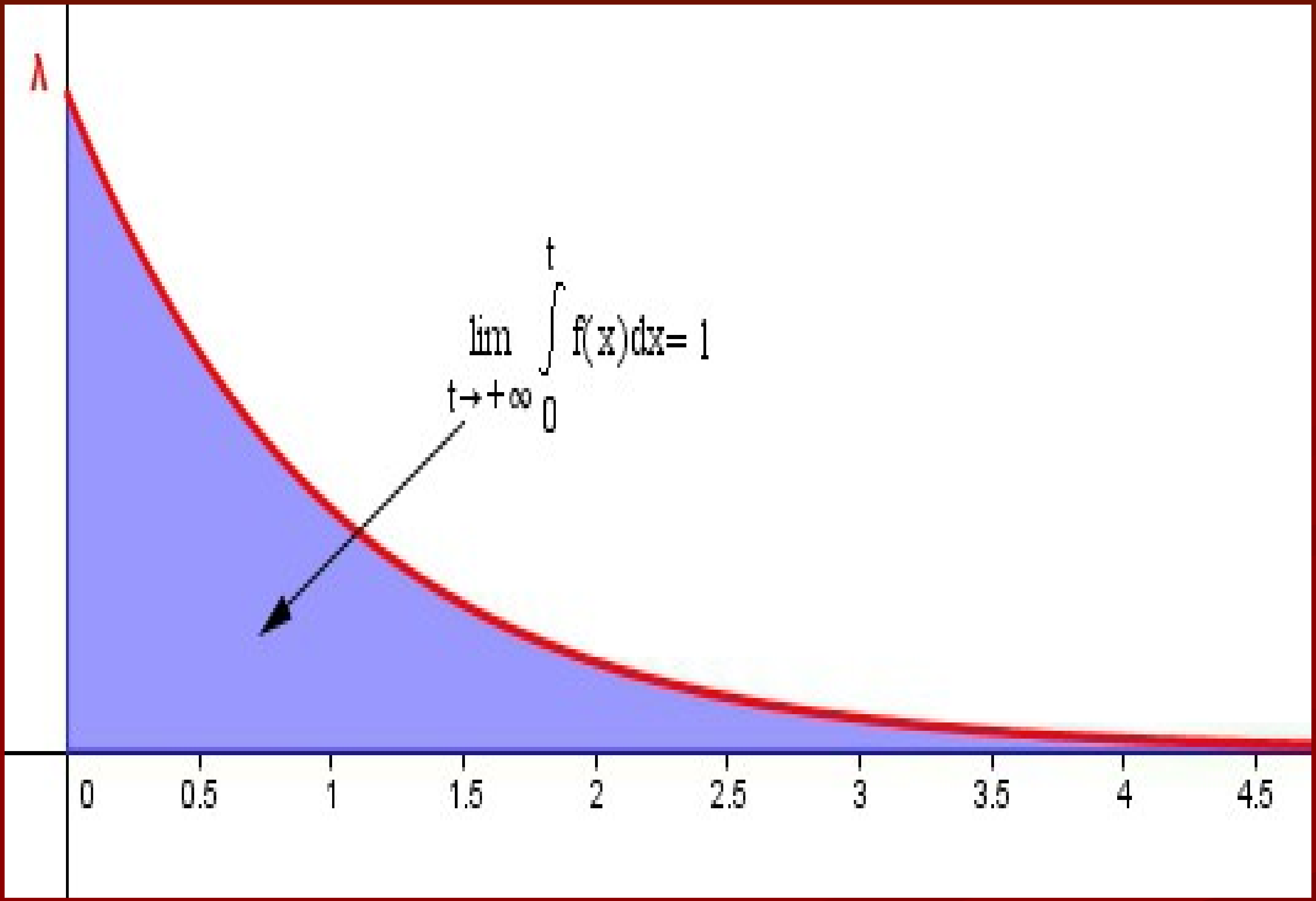
- **L'absence de mémoire se traduit par le fait qu'un phénomène a autant de chances de se produire sur un laps de temps donné après l'instant *t* qu'après l'instant *h*. La probabilité qu'il survienne aujourd'hui sachant qu'on l'attend depuis un siècle est la même que si on l'attendait depuis un jour.**

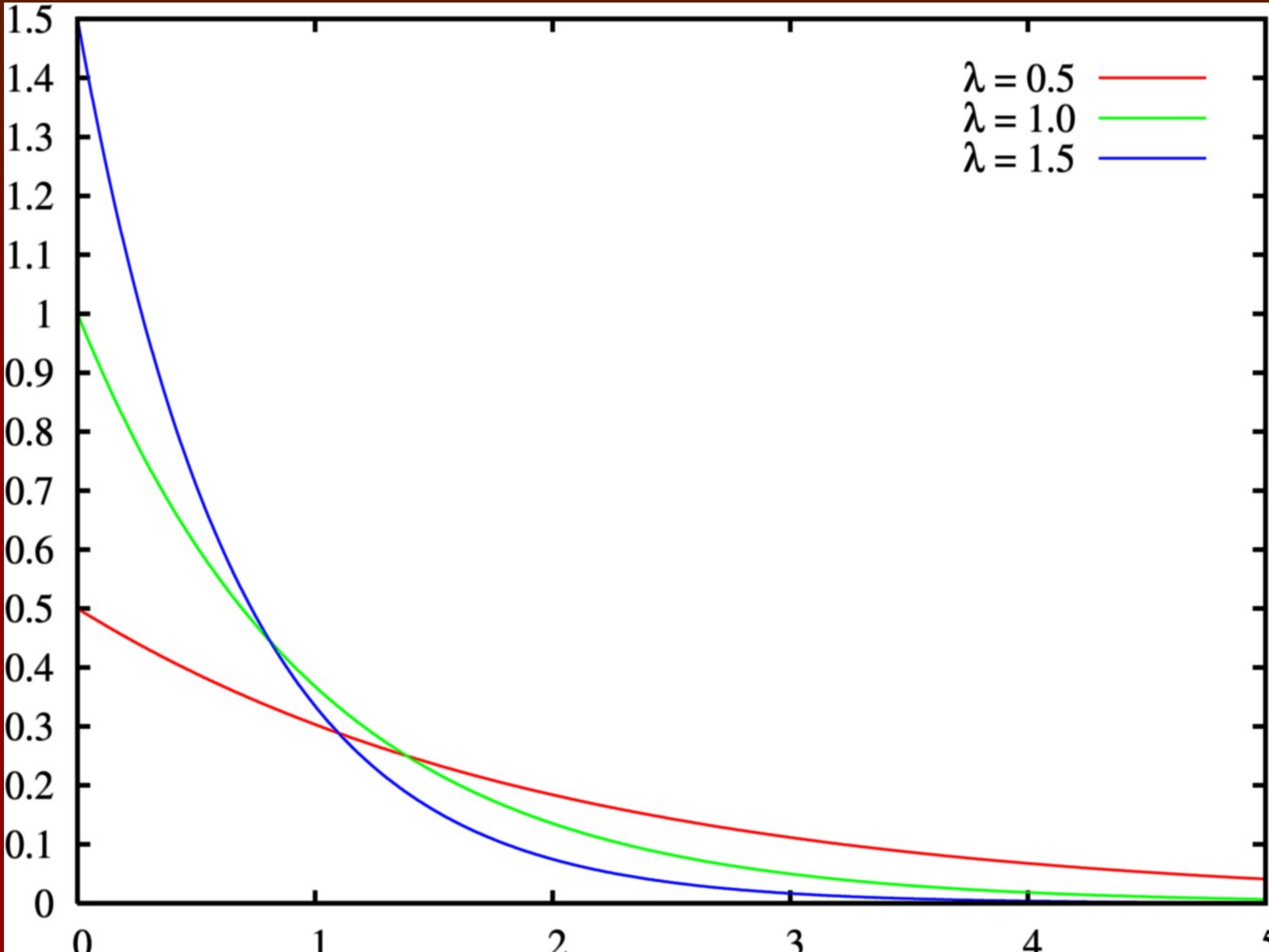
- **Définition**
  Soit λ une constante strictement positive.
  On dit qu'une variable aléatoire suit une loi exponentielle de paramètre λ sur , lorsque sa densité de probabilité associée est la fonction f(.) définie sur IR+ par:

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

et

Illustration




λ = 0.5
λ = 1.0
λ = 1.5
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5

Remarques

- $f(0) = \lambda$.

- f est continue et positive sur IR.

- $E(X) = \dfrac{1}{\lambda}$

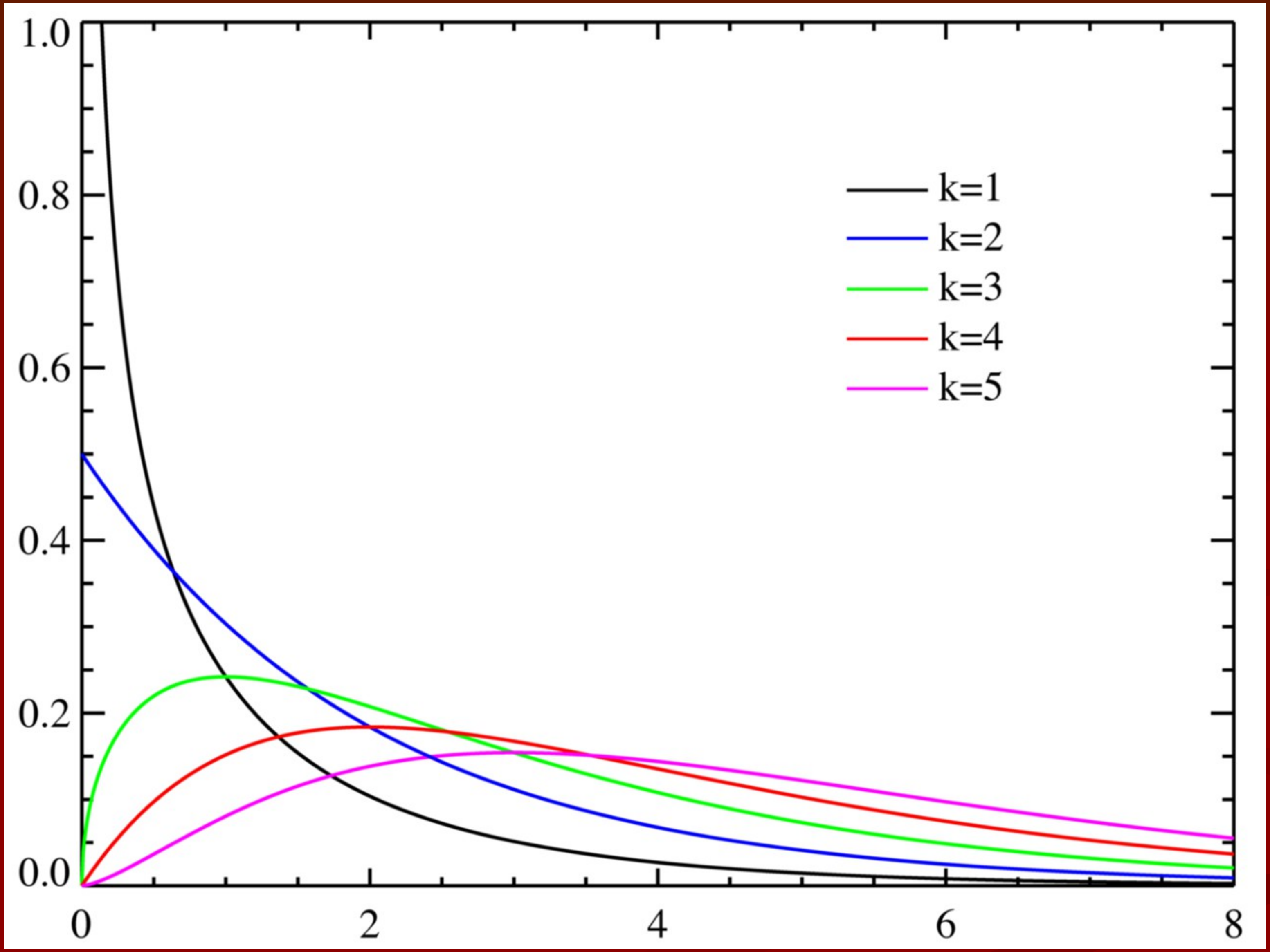
- $Var(X) = \dfrac{1}{\lambda^2}$

I.3. La loi de khi-carré

Soit X une v.a continue. On dit que X suit la loi de Khi-carré à k d.d.l notée χ_k^2 si sa fonction densité définie sur IR+ est :

$$f(x) = x_0 x^{k/2-1} e^{-\frac{1}{2}x} I_{[0,+\infty]}(x)$$

où x_0 est une constante positive.



1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
k=1
k=2
k=3
k=4
k=5

Résultat

$U_1, U_2, \ldots, U_p$ étant p variables indépendantes et distribuées selon la loi normales centrée et réduite.

Alors

$$\sum_{i}^{p} U_i^2 \approx \chi_p^2$$

Si X est distribuée selon la loi de χ_K^2 ,

$$E(X) = K \quad \text{et} \quad Var(X) = 2K$$

I.4. La loi de Student

Soit X une v.a continue. On dit que X suit la loi de Student à ν ($\nu \in IN$) degrés de liberté (d.d.l) si sa fonction densité est :

$$f(t) = \frac{t_0}{\left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}}$$

où t_0 est une constante positive.

On montre que:

E(X)=0 et Var(X)= $\frac{\nu}{\nu-2}$ si $\nu \geq 3$,

Résultat

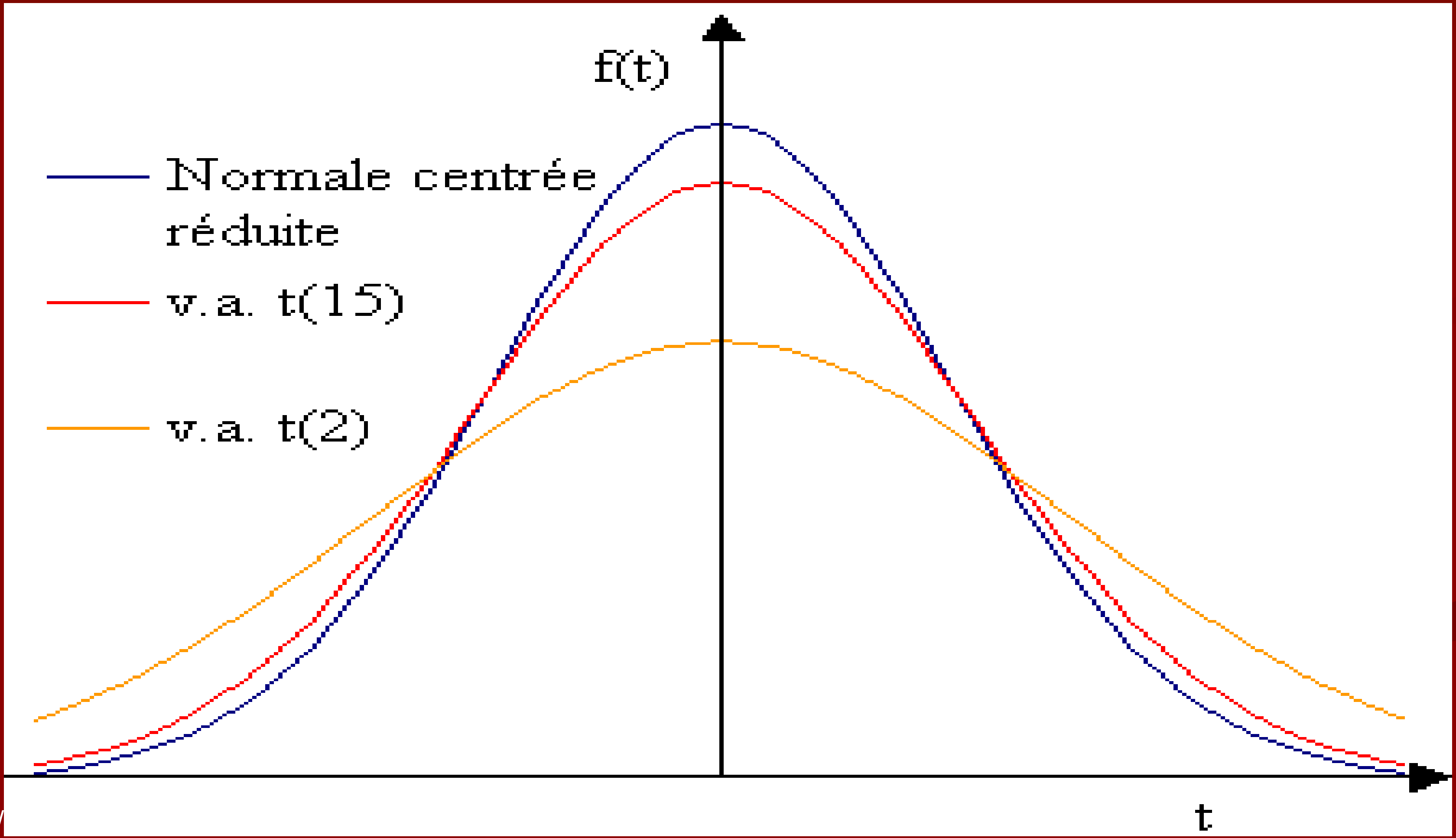
Soient U et X deux v.a <u>indépendantes</u> telles que

$$U \approx N(0,1) \quad \text{et} \quad X \approx \chi_p^2$$

alors

$$\frac{U}{\sqrt{\frac{X}{p}}} \approx t_p$$

Remarque: La loi de Student tend vers la loi normale.

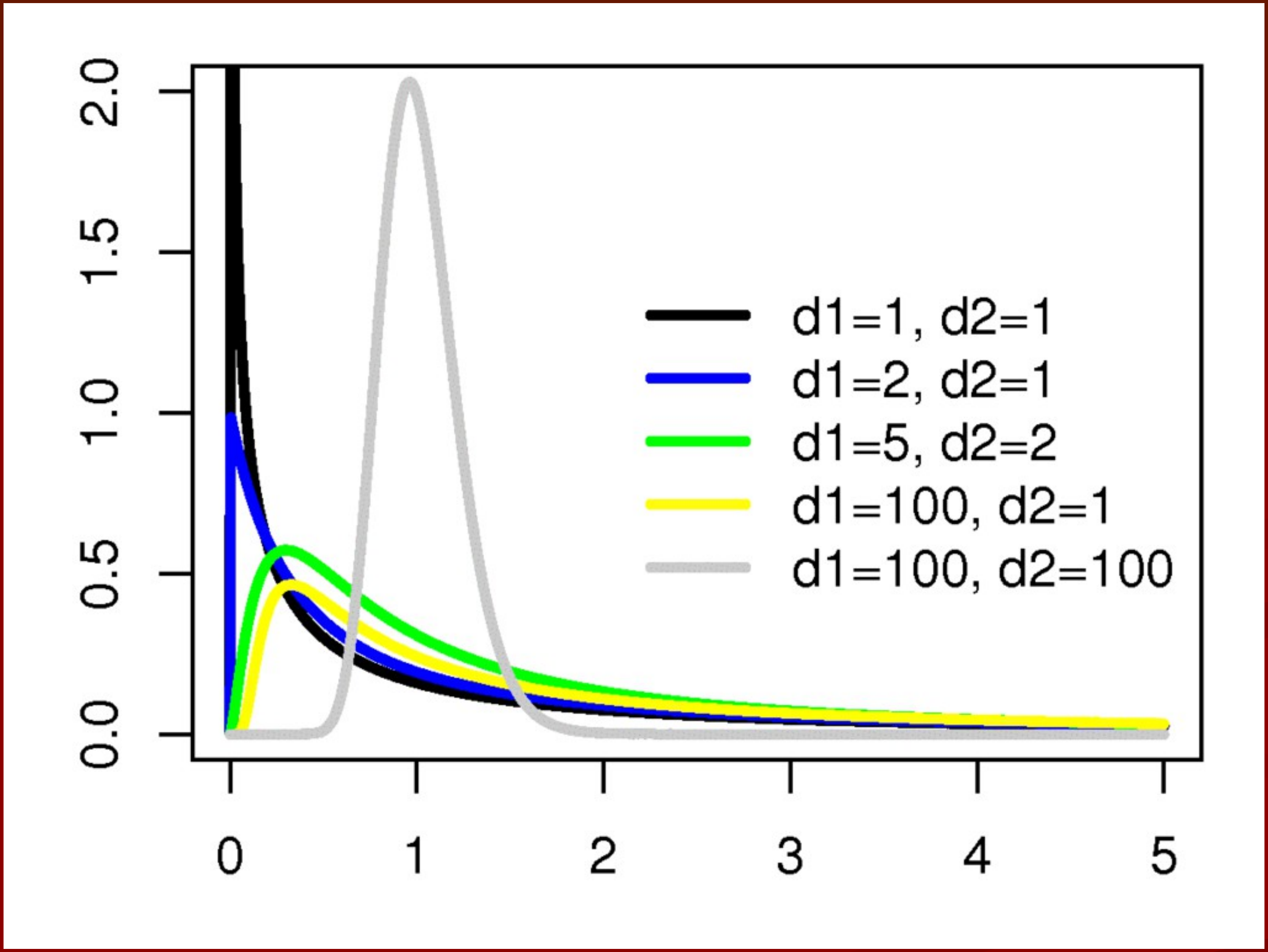


I.5. La loi de Fisher

Soit X une v.a continue. On dit que X suit la loi de Fisher à n_1 et n_2 d.d.l notée F_{n_1,n_2} si sa fonction densité définie sur IR+ est :

$$f(x) = \frac{Cx^{\frac{n_1}{2}-1}}{(n_1x+n_2)^{\frac{n_1+n_2}{2}}}$$

où **C** est une constante positive.



d1=1, d2=1
d1=2, d2=1
d1=5, d2=2
d1=100, d2=1
d1=100, d2=100
0.0
0.5
1.0
1.5
2.0
0
1
2
3
4
5

Résultat

Soient X et Y deux v.a indépendantes telles que

$X \approx \chi_p^2$ et $Y \approx \chi_q^2$

Alors

$$\frac{X/p}{Y/q} \approx F_{p,q}$$

II. Quelques loi discrètes

II.1 Loi de Bernoulli

On appelle **variable de Bernoulli** ou variable ***indicatrice***, la variable aléatoire *X* telle que :

$$X : \Omega \to R / \ \Omega = \{\text{échec, succès}\} \ \text{et} \ X(\Omega) = \{0,1\}$$

Loi de probabilité: $P(X=0)=1-p$ et $P(X=1)=p$
où p est compris entre 0 et 1,
et $E(X) = p$ et $Var(X) = p(1-p)$

II.2 Loi Binomiale

- On appelle **variable Binomiale** une variable aléatoire X correspondant à la somme de n variables de Bernoulli :

$$X = \sum_{i=1}^{n} X_i : \Omega^n \to IN \quad \text{telle que}$$

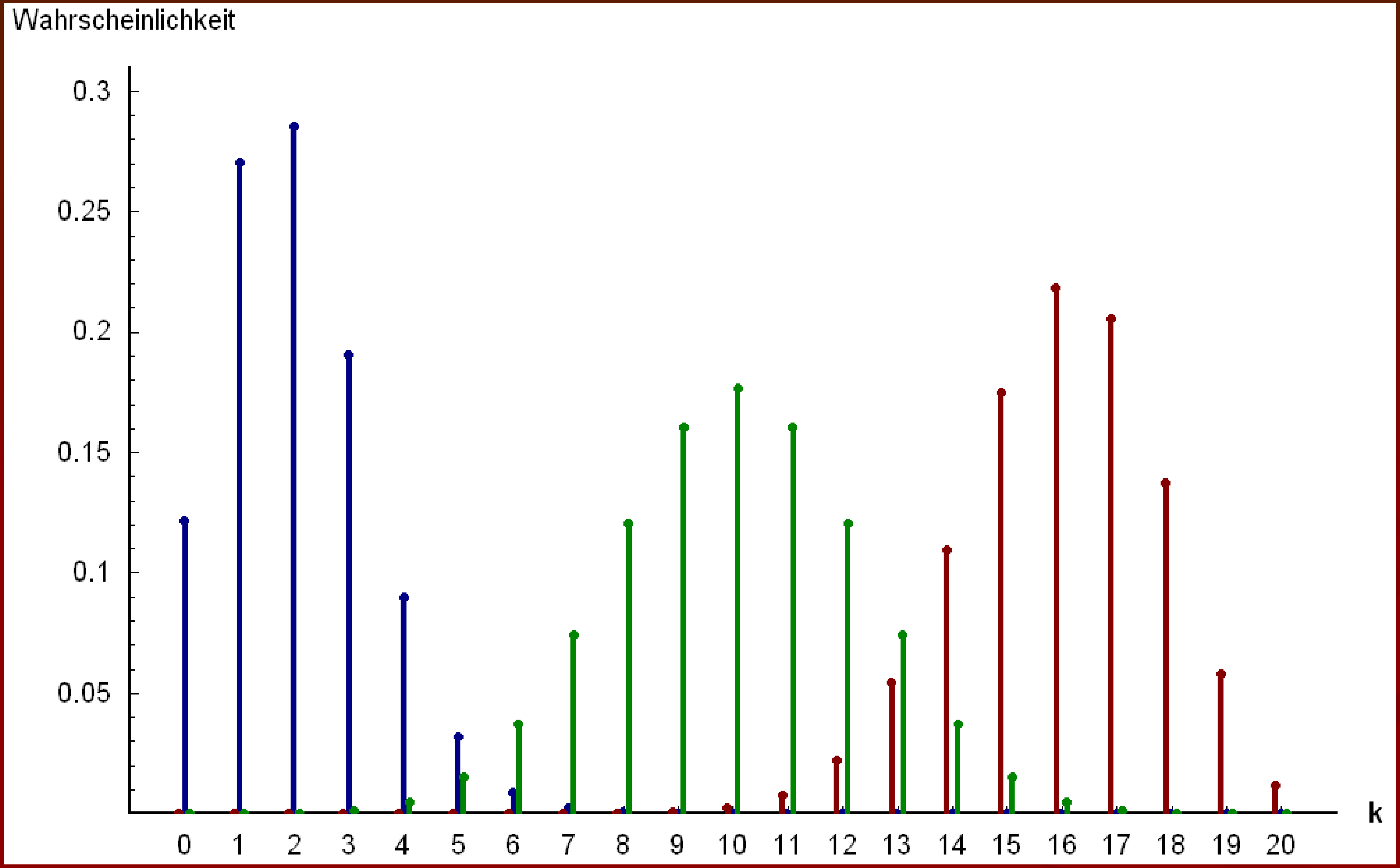
$$X(\Omega^n) = \{0,1, \ldots.n\}$$

Loi de probabilité: $p(X = k) = C_n^k \, p^k \, q^{n-k} \quad \text{avec} \quad C_n^k = \dfrac{n!}{k!(n-k)!}$

Elle est appelée **loi Binomiale**, notée **B(n,p) et on a,**

$$E(X) = np$$

$$V(X) = npq$$



Diagrammes en bâtons de trois fonctions de masse de lois binomiales. Les paramètres sont $n = 20$ et $p = 0,1$ (en bleu), $p = 0,5$ (en vert) et $p = 0,8$ (en rouge).

II.3 Loi de Poisson

Le nombre moyen d'occurrence d'un événement X dans un temps T est k

- X={0,1,...} :

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

On note $X \approx P(\lambda)$

$$E(X) = \lambda \text{ et } V(X) = \lambda$$

λ = nombre moyen d'événement par unité de temps.

- Relation avec la loi binomiale :
  - Si p<0.1 et n>50 : $\mathcal{B}(n,p) \approx P(np)$

II.4 Loi géométrique (loi de Pascal)

- Une variable géométrique correspond au nombre d'épreuves de Bernoulli nécessaires pour obtenir 1 su

$$X: \Omega \to IN^* \text{ telle que}$$
$$X(\Omega) = \{1, 2 \ldots, n \ldots\}$$

Loi de probabilité: $$p(X = i) = q^{i-1}\, p$$

$$E(X) = \frac{1}{p}$$

$$V(X) = \frac{q}{p^2}$$

II.5 Loi Hypergéométrique

Une variable hypergéométrique correspond à l'observation de k succè

lors de n épreuves sans remise, dans une population de dimension N:

$$X : \Omega^n \rightarrow IN \quad \text{telle que}$$
$$X(\Omega^n) = \{0, 1, ...n\}$$

Loi de probabilité:

$$p(X = k) = \frac{C_{Np}^{k} \; C_{N(1-p)}^{n-k}}{C_N^n}$$

p étant la probabilité qu'un tirage se réalise avec succès.

$$E(X) = np$$

$$V(X) = \frac{N - n}{N - 1} n\,p(1 - p)$$