Probabilités
Variables Aléatoires
Lois de Distribution

Pr Roch Giorgi

roch.giorgi@univ-amu.fr

SESSTIM, Faculté de Médecine, Aix-Marseille Université, Marseille, France
http://sesstim.univ-amu.fr

Introduction

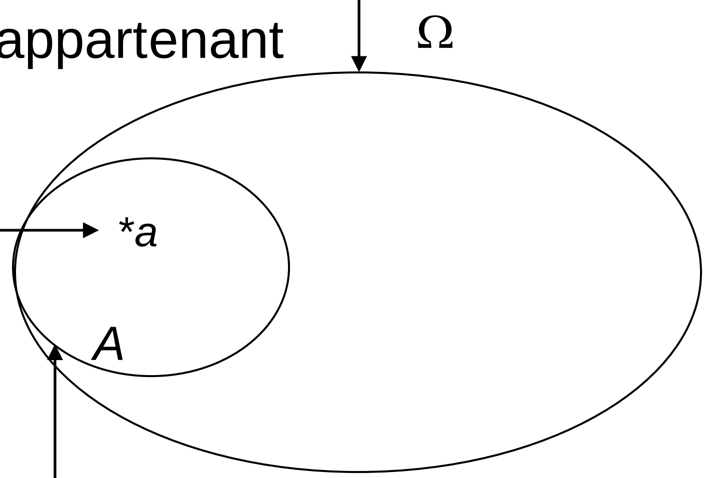
- Probabilité : modélise des phénomènes aléatoires dont les issues sont connues mais dont on ne peut en prédire la valeur car leur réalisation est incertaine

- Observation des issues d'un phénomène aléatoire sur des séries suffisamment grandes permet d'en déterminer leurs fréquences et par la suite la loi de distribution qui le dirige

Rappels sur les Ensembles

Ensemble : collection bien définie d'objets

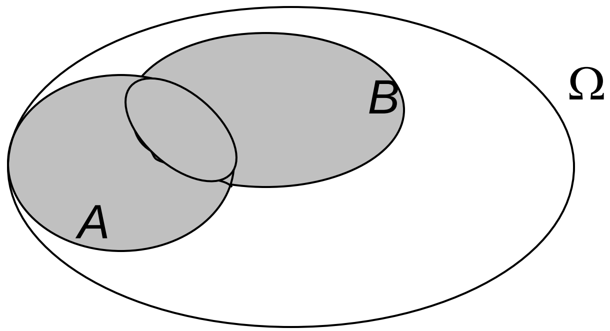
Elément : objet appartenant à l'ensemble



Sous-ensemble : *A* : ensemble d'éléments *a* appartenant tous à Ω

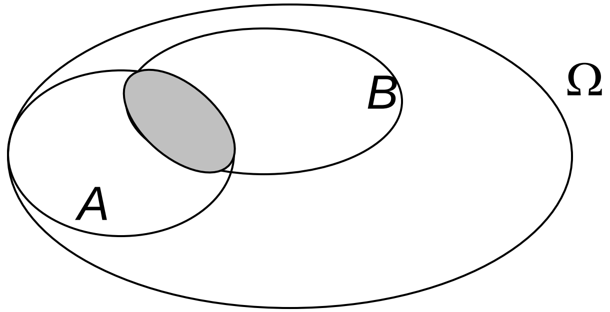
Rappels sur les Ensembles

<u>Réunion</u> : $A \cup B \Leftrightarrow A$ **ou** B

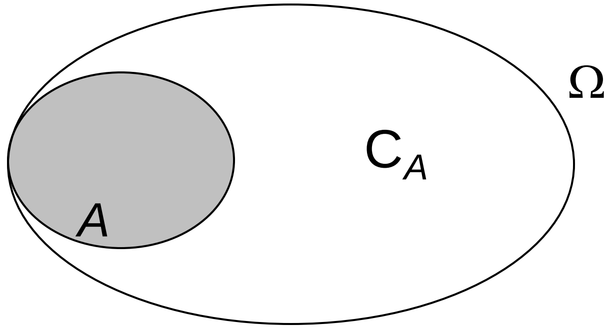


<u>Intersection</u> : $A \cap B \Leftrightarrow A$ **et** B

si $A \cap B = \varnothing$ alors A et B sont disjoints.



<u>Complémentarité</u> : C_A



Notion de Probabilité

- Probabilité : modélisation de phénomènes aléatoires

- Ensemble fondamental : Ω : ensemble des résultats possibles pour une expérience donnée (événements certains)

- Événement : c'est un sous-ensemble *A* de Ω, c'est-à-dire un ensemble de résultats. Un événement élémentaire est *a*

Exemple

Expérience aléatoire d'un jet de dé non pipé à 6 faces

Ensemble fondamental : $\Omega = \{f1, f2, f3, f4, f5, f6\}$

Événement A : face de nombre $\leq 2 = f1 \cup f2$
Événement B : face de nombre $\geq 5 = f5 \cup f6$
Événement C : face de nombre paire $\{2, 4, 6\} = f2 \cup f4 \cup f6$

$A \cup B = f1 \cup f2 \cup f5 \cup f6$, $A \cap B = \varnothing$
$A \cup C = f1 \cup f2 \cup f4 \cup f6$, $A \cap C \neq \varnothing$

Notion de Probabilité

Épreuve répétée n fois

	f1	f2	f3	f4	f5	f6	Total
Fréquences Absolues	n_1	n_2	n_3	n_4	n_5	n_6	n
Fréquences Relatives	n_1/n	n_2/n	n_3/n	n_4/n	n_5/n	n_6/n	1

$\text{Freq}(A)=(n_1+n_2)/n$

$\text{Freq}(A \cup B)=(n_1+n_2+n_5+n_6)/n=(n_1+n_2)/n+(n_5+n_6)/n=\text{Freq}(A)+\text{Freq}(B)$

$\text{Freq}(A \cup C)=(n_1+n_2+n_4+n_6)/n \neq \text{Freq}(A)+\text{Freq}(C)$

Lorsque $n \to \infty$ la fréquence relative d'un événement tend vers la probabilité de cet événement

Probabilités Élémentaires

- Soit Ω un ensemble fondamental, P la fonction de probabilité qui à tout événement A associe un nombre réel positif ou nul. $P(A)$ est appelée probabilité de l'événement A si :

  $P(A) \geq 0$

  $P(\Omega) = 1$

  si $A \cap B = \varnothing \Rightarrow P(A \cup B) = P(A) + P(B)$

  si $A_i \cap A_j = \varnothing \Rightarrow P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$

  on en déduit que :

  - ✓ $P(\varnothing) = 0$
  - ✓ $P(A) \leq 1$
  - ✓ $P(C_A) = 1 - P(A)$
  - ✓ si $A \subset B$, alors $P(A) \leq P(B)$
  - ✓ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Probabilités Conditionnelles

- Exemple :
  - ✓ On s'intéresse au test de l'Hémocult dans le cadre du dépistage du cancer colorectal.
  - ✓ La probabilité d'avoir un cancer colorectal sachant le test Hémocult positif est une probabilité conditionnelle

P(CCR \ Hémoc. Positif)

Probabilités Conditionnelles

- La probabilité de A sachant B est définie par

$$\boxed{\mathrm{P}(A \setminus B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}}$$

d'où $\mathrm{P}(A \cap B) = \mathrm{P}(A \setminus B)\mathrm{P}(B) = \mathrm{P}(B \setminus A)\mathrm{P}(A)$

et

$$\boxed{P(A \setminus B) = \frac{P(B \setminus A)P(A)}{P(B)}}$$

Théorème de Bayes

Indépendance en Probabilité

- A et B sont indépendants ssi

$$\boxed{\mathrm{P}(A \cap B) = \mathrm{P}(A).\mathrm{P}(B)}$$

- Si A et B sont indépendants et P(A) > 0, P(B) > 0, alors

$$\mathrm{P}(A \backslash B) = \mathrm{P}(A \cap B) \backslash \mathrm{P}(\mathrm{B}) = \mathrm{P}(A).\mathrm{P}(B) \backslash \mathrm{P}(\mathrm{B}) = \mathrm{P}(A)$$

- 2 événements disjoints de probabilités non nulles ne sont jamais indépendants
  - ✓ disjoints : $\mathrm{P}(\mathrm{A} \cap B) = 0$
  - ✓ indépendants : $\mathrm{P}(\mathrm{A} \cap B) = \mathrm{P}(\mathrm{A}).\mathrm{P}(\mathrm{B})$

Probabilités Conditionnelles

- $A_1, \ldots, A_n$ événements formant une partition de Ω
- B un événement quelconque

Alors

$$P(B) = P(B \cap A_1) \cup P(B \cap A_2) \cup \ldots \cup P(B \cap A_n)$$

Et

$$\boxed{P(A_i \backslash B) = \frac{P(B \backslash A_i)P(A_i)}{P(B \backslash A_1)P(A_1) + \ldots + P(B \backslash A_n)P(A_n)}}$$

Formule développée de Bayes

Probabilités Conditionnelles

- Exemple :
  - ✓ La prévalence du SIDA dans une population est de 10 %
  - ✓ On sait qu'un test diagnostique est positif chez 95% des HIV^+ et qu'il est négatif chez 98% des HIV^-
  - ✓ Qu'elle est la probabilité d'être HIV^+ si le test est positif

P(HIV+) = 0,1

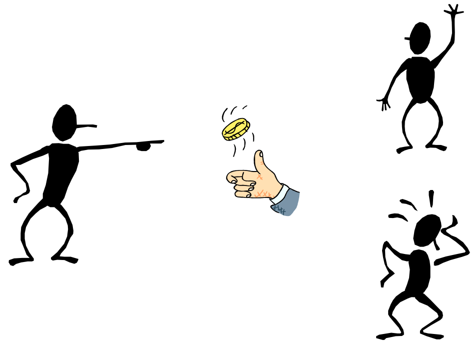
P(T+ \ HIV+) = 0,95

P(T+ \ HIV-) = 0,02

$$P(HIV+ \backslash T+) = \frac{P(T+ \backslash HIV+)\,P(HIV+)}{P(T+ \backslash HIV+)\,P(HIV+) + P(T+ \backslash HIV-)\,P(HIV-)}$$

P(HIV+\ T+) = (0,95*0,1)/(0,95*0,1 + 0,02*0,9) = 0,84

Variable Aléatoire



Si Pile, A gagne 1 €

Si Face, A perd 1 €

- Ω : {Pile, Face}
- P(Pile) = P(Face) = 0,5
- G : gain de A; G = +1, si Pile; G = -1, si Face
- P(G = +1) = P(G = -1) = 0,5
- Distribution de G : {(+1; 0,5), (-1; 0,5)}

G : variable aléatoire qui suit une certaine loi de probabilité

Variable Aléatoire : Définition

- Soit E un ensemble d'événements
- d'ensemble fondamental Ω fini, et
- *a* un événement élémentaire de E

Pour tout événement *a* appartenant à E on fait correspondre un nombre *x* (variable aléatoire) selon une loi bien définie

Variable Aléatoire : Exemple

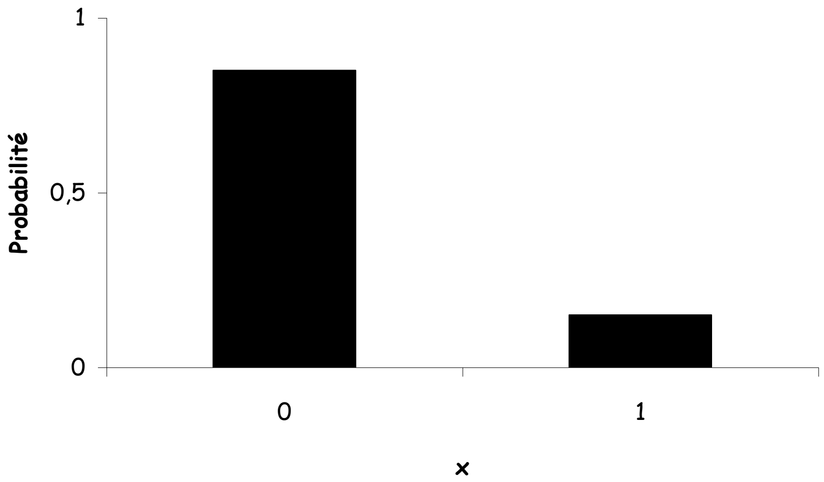
- Soit une maladie M pour laquelle il est nécessaire de débuter un TRT avant confirmation du diagnostic. Le médicament utilisé est cependant connu pour entraîner des effets indésirables.
- On sait que : $P(M^+)=5\%$; $P(EI^+ \backslash M+)=30\%$; $P(EI^- \backslash M^-)=85\%$

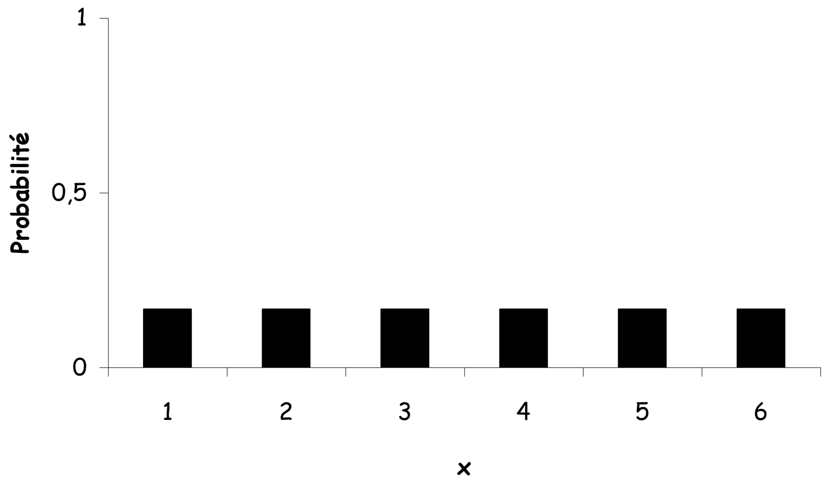
	M^+	M^-
EI^+	$P(EI^+ \cap M^+) = 0{,}3 x 0{,}05$ $= 1{,}5\%$ $X = 1$	$P(EI^+ \cap M^-) = (1-0{,}85) x (1-0{,}05)$ $= 14{,}3\%$ $X = 1$
EI^-	$P(EI^- \cap M^+) = (1-0{,}3) x 0{,}05$ $= 3{,}5\%$ $X = 0$	$P(EI^- \cap M^-) = 0{,}85 x (1-0{,}05)$ $= 80{,}8\%$ $X = 0$

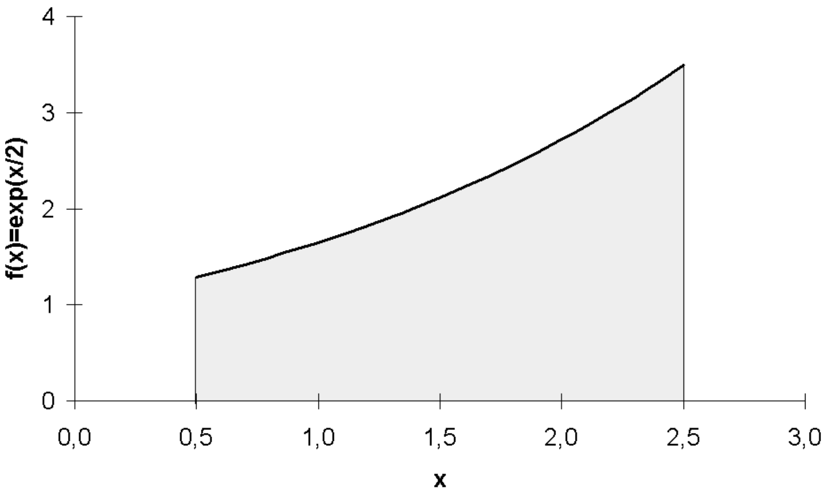
où X est une va indicatrice des EI.

La distribution de X est : {(0; 0,84), (1; 0,16)}

Caractéristiques d'une Variable Aléatoire







Caractéristique de Position : Moyenne, Espérance

- Variable discrète X
  - ✓ soit X une va prenant les valeurs $x_1, x_2, \ldots, x_n$ avec les probabilités $p_1, p_2, \ldots, p_n$ et $\Sigma p_i = 1$, i = 1, …, n :

$$\boxed{\mu = E(X) = \sum p_i x_i}$$

- Cas d'une variable continue X
  - ✓ définie par une loi de densité $f(x)$

$$\boxed{\mu = E(X) = \int_a^b x f(x)\,dx}$$

Caractéristique de Position :

Moyenne, Espérance

- Exemple 1 : μ = (p x 1) + (q x 0) = p

- Exemple 2 : μ = 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6 = 3,5

- Exemple 3 : $\mu = E(X) = \int_{0,5}^{2,5} f(x)dx = \int_{0,5}^{2,5} \exp\left(x/2\right)dx = \left[\exp\left(x/2\right)\right]_{0,5}^{2,5} = 2,21$

Caractéristique de Dispersion : Variance, Écart-type

- Cas d'une variable discrète X :

$$\boxed{\begin{aligned}\sigma^2 &= \sum p_i [x_i - \mu]^2 \\ &= E\left((X-\mu)^2\right) = E\left(X^2\right) - [E(X)]^2\end{aligned}}$$

- Cas d'une variable continue X :

$$\boxed{\sigma^2 = \int_a^b (x-\mu)^2 f(x)dx}$$

σ^2 = variance

σ = écart-type

Caractéristique de Dispersion :

Variance, Écart-type

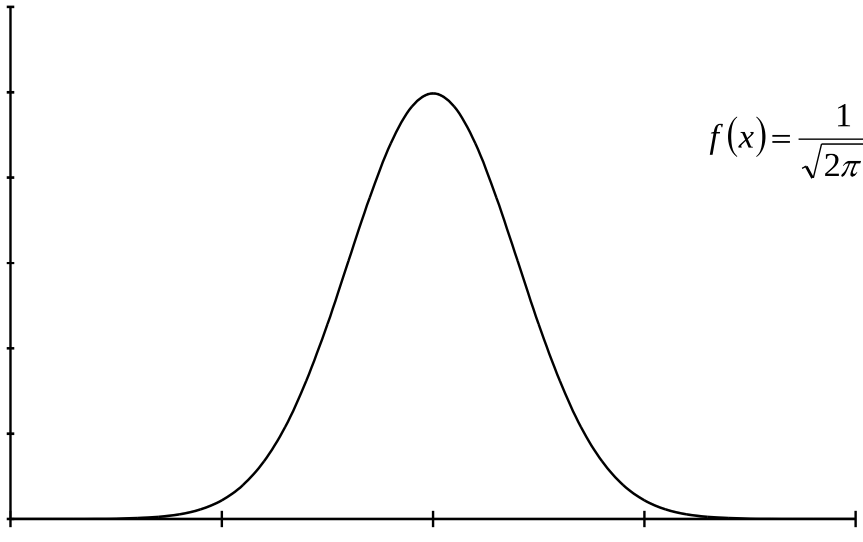
- Exemple 1 : $\sigma^2 = p \times (1 - p)^2 + q \times (0 - p)^2 = pq$

- Exemple 2 : $\sigma^2 = 1/6[(1 - 3,5)^2 + (2 - 3,5)^2 + ... + (6 - 3,5)^2]$

  $= 2,9$

- Exemple 3 : $\sigma 2 = \int_{0.5}^{2,5} (x - 2,21)^2 \exp\left(\frac{x}{2}\right) dx = \int_{0.5}^{2,5} x^2 \exp\left(\frac{x}{2}\right) dx - 2,21^2 = 0,68$

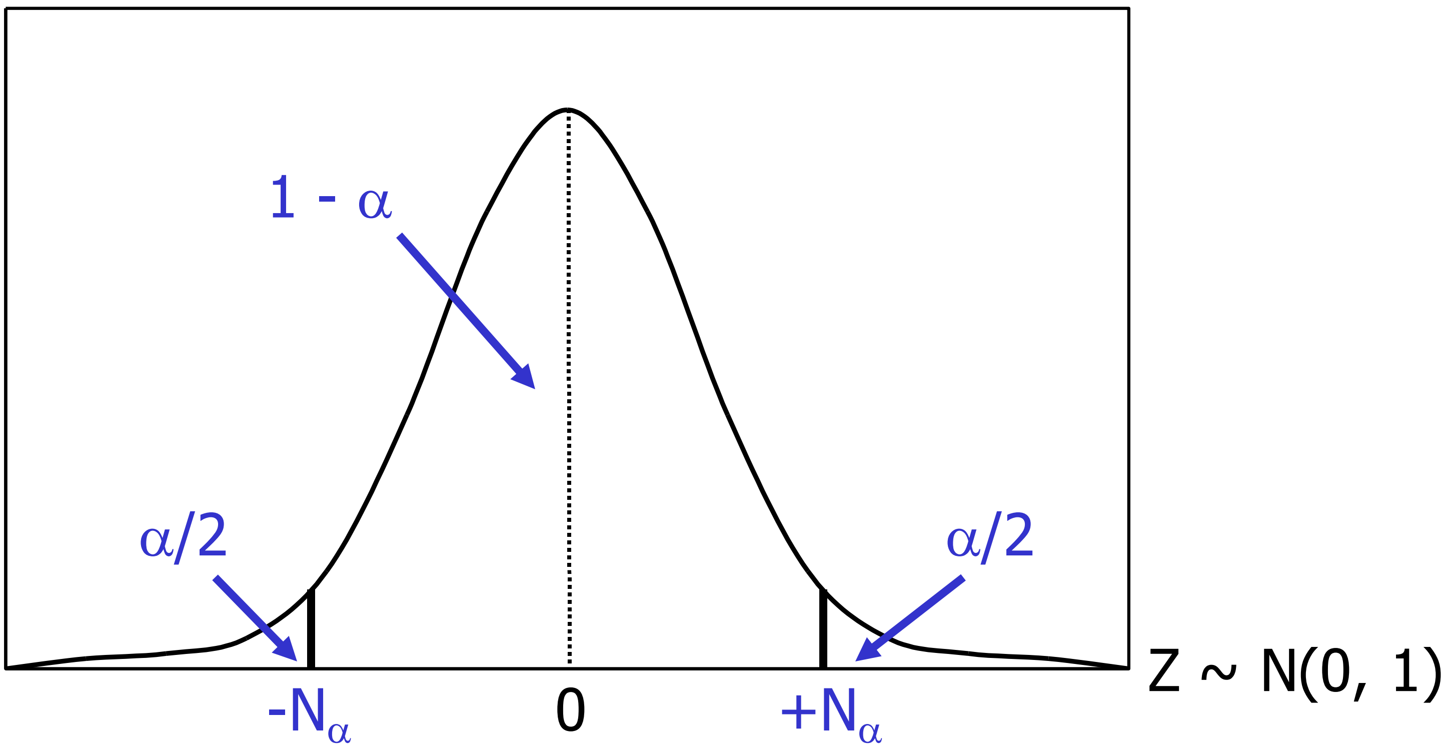
Loi Normale (Gauss)



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

- Propriétés
  - ✓ Fonction de densité continue, déterminée par μ et σ
  - ✓ Fonction de densité symétrique par rapport à μ
  - ✓ Fonction de densité passe par un maximum pour $x = \mu$ (mode = μ)
  - ✓ Médiane = μ

Loi Normale Centrée Réduite : N(0, 1)



$$\alpha = \text{Prob}(Z \leq -N_\alpha \text{ ou } Z \geq +N_\alpha) = \text{Prob}(|Z| \geq N_\alpha)$$

Table de la loi N(0, 1)

	α = 0,10	α = 0,05	α = 0,01
N_α	1,64	1,96	2,58

Loi de Student

- ν : Nombre de degré de liberté (nombre de données indépendantes)
- Une famille de distributions de Student pour chaque ddl
- Propriétés
  - ✓ symétrique par rapport à 0
  - ✓ mode = 0
  - ✓ s'aplatie quand ν petit
  - ✓ Tend vers N(0,1) quand $\nu \to \infty$

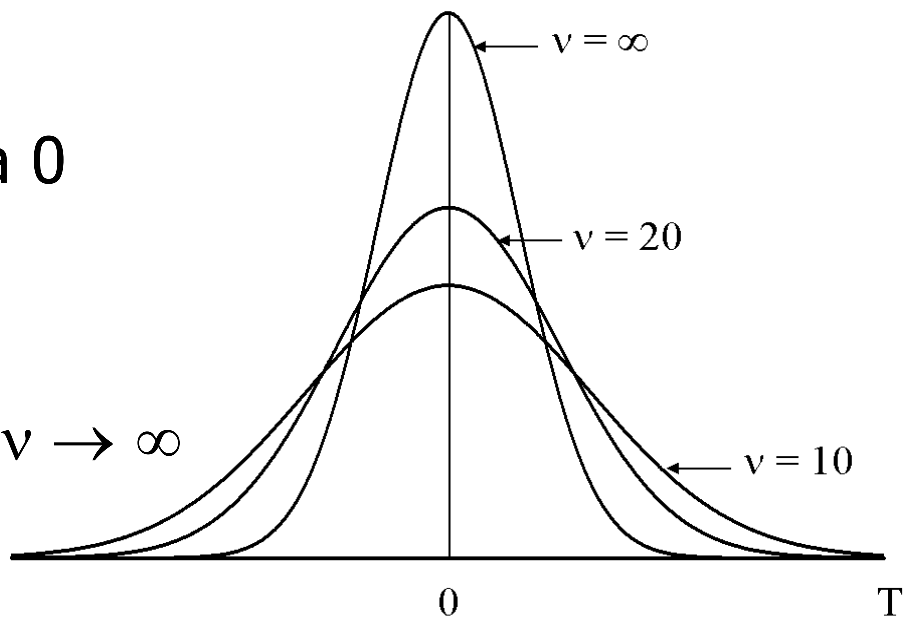


Table de la Loi de Student

- Valeur de $T_{\alpha,\nu}$ en fonction de certaines valeurs α

$$\alpha = \text{Proba}\left(\left(T \leq -T_{\alpha,\nu}\right) \text{ ou } \left(T \geq +T_{\alpha,\nu}\right)\right) = \text{Proba}\left(\left|T\right| \geq T_{\alpha,\nu}\right)$$

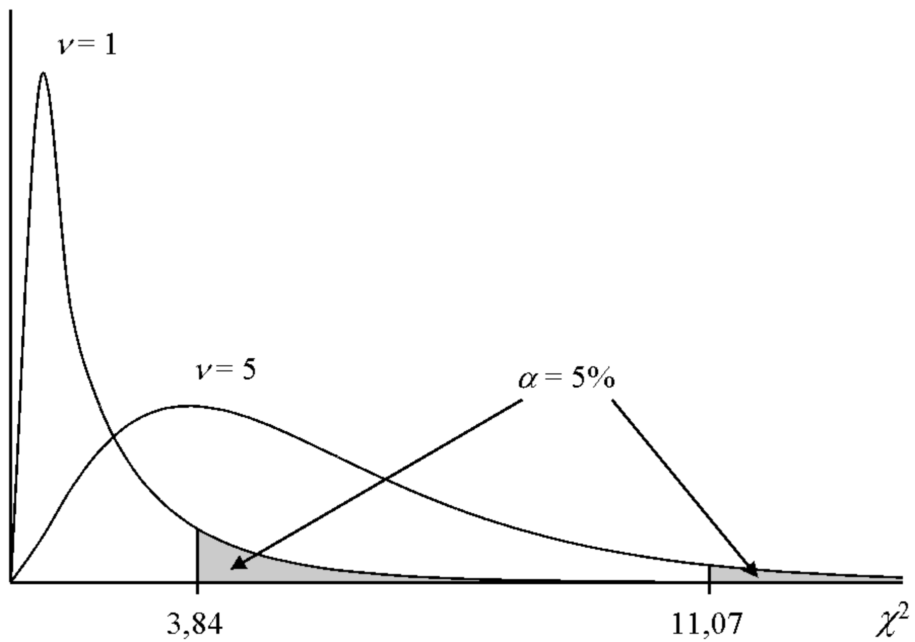
- Pour ν supérieur ou égal à 30, $T_{\alpha,\nu}$ est arrondi à N_{α}

Exemple :

si α = 5 % et ν = 10 alors $T_{0,05\,;\,10} = 2,228$
si α = 10 % et ν = 8 alors $T_{0,10\,;\,8} = 1,860$
si α = 5 % et $\nu \geq 30$ alors $T_{0,05\,;\,\geq 30} = 1,96$

Loi du Chi-deux (χ^2)

- Une famille de distributions de χ^2 pour chaque ddl



- Propriétés
  - ✓ Asymétrique pour ν petit

Table de la Loi du Chi-deux

- Donne la probabilité α que χ^2 soit supérieur ou égal à une valeur donnée pour chaque degré de liberté

Exemple :

si α = *5 % et* ν = 1 alors $\chi^2_{0,05,\,1} = 3,84$

si α = *5 % et* ν = 5 alors $\chi^2_{0,05,\,5} = 11,07$