

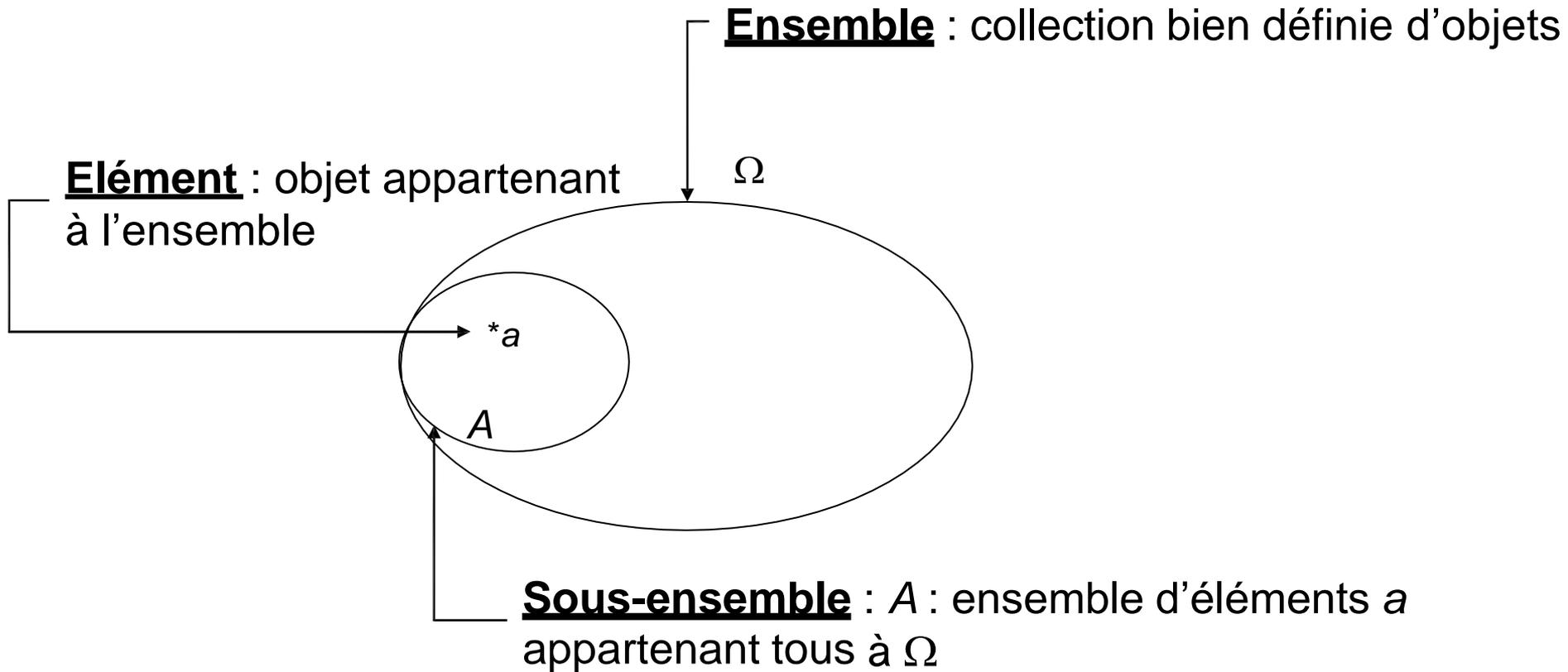
PARTIE 2

Probabilités, Variables Aléatoires lois de Distribution

Introduction

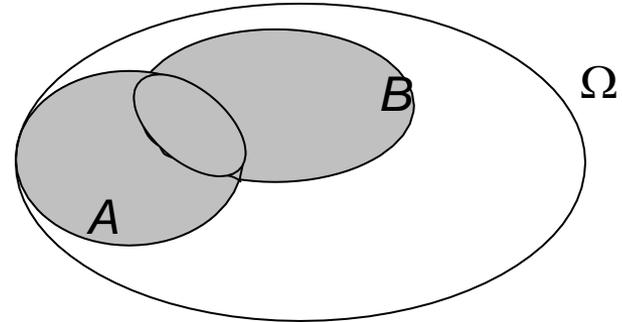
- **Probabilité** : modélise des phénomènes **aléatoires** dont les issues sont connues mais dont on ne peut en prédire la valeur car leur **réalisation** est **incertaine**
- Observation des issues d'un phénomène aléatoire sur des séries suffisamment grandes permet d'en déterminer leurs **fréquences** et par la suite la **loi de distribution** qui le dirige

Rappels sur les Ensembles

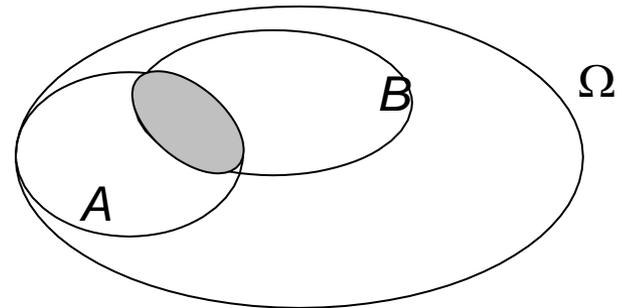


Rappels sur les Ensembles

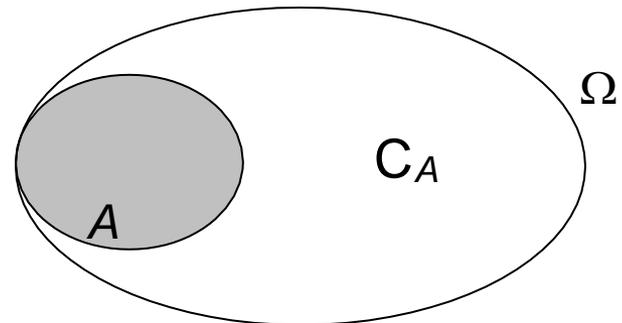
Réunion : $A \cup B \Leftrightarrow A \text{ ou } B$



Intersection : $A \cap B \Leftrightarrow A \text{ et } B$
si $A \cap B = \emptyset$ alors A et B
sont
disjoints.



Complémentarité : 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 : $\Omega = \{f_1, f_2, f_3, f_4, f_5, f_6\}$

Événement A : face de nombre $\leq 2 = f_1 \cup f_2$

Événement B : face de nombre $\geq 5 = f_5 \cup f_6$

Événement C : face de nombre paire $\{2, 4, 6\} = f_2 \cup f_4 \cup f_6$

$$A \cup B = f_1 \cup f_2 \cup f_5 \cup f_6, A \cap B = \emptyset$$

$$A \cup C = f_1 \cup f_2 \cup f_4 \cup f_6, A \cap C \neq \emptyset$$

Probabilités Élémentaires

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 $n \rightarrow \infty$ 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 :

$$P(A) \geq 0$$

$$P(\Omega) = 1$$

$$\text{si } A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$$

$$\text{si } A_i \cap A_j = \emptyset \Rightarrow P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

on en déduit que :

- ✓ $P(\emptyset) = 0$
- ✓ $P(A) \leq 1$
- ✓ $P(C_A) = 1 - P(A)$
- ✓ 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

$P(\text{CCR} \mid \text{Hémoc. Positif})$

Probabilités Conditionnelles

- La probabilité de A sachant B est définie par

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

d'où $P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$
et

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Théorème de Bayes

Indépendance en Probabilité

- A et B sont indépendants ssi

$$P(A \cap B) = P(A) \cdot P(B)$$

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

$$P(A \setminus B) = P(A \cap B) \setminus P(B) = P(A) \cdot P(B) \setminus P(B) = P(A)$$

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

Probabilités Conditionnelles

- A_1, \dots, A_n événements formant une **partition** de Ω
- B un événement quelconque

Alors

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

Et

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B \mid A_1)P(A_1) + \dots + P(B \mid A_n)P(A_n)}$$

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

$$P(\text{HIV}+) = 0,1$$

$$P(\text{T}+ \setminus \text{HIV}+) = 0,95$$

$$P(\text{T}+ \setminus \text{HIV}-) = 0,02$$

$$P(\text{HIV}+ \setminus \text{T}+) = \frac{P(\text{T}+ \setminus \text{HIV}+)P(\text{HIV}+)}{P(\text{T}+ \setminus \text{HIV}+)P(\text{HIV}+) + P(\text{T}+ \setminus \text{HIV}-)P(\text{HIV}-)}$$

$$P(\text{HIV}+ \setminus \text{T}+) = (0,95*0,1)/(0,95*0,1 + 0,02*0,9) = 0,84$$

Variable Aléatoire



Si Pile, A gagne 100 FCFA

Si Face, A perd 100 FCFA

- $\Omega : \{\text{Pile, Face}\}$
- $P(\text{Pile}) = P(\text{Face}) = 0,5$
- $G : \text{gain de A; } G = +100, \text{ si Pile; } G = -100, \text{ si Face}$
- $P(G = +100) = P(G = -100) = 0,5$
- Distribution de $G : \{(+100; 0,5), (-100; 0,5)\}$

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

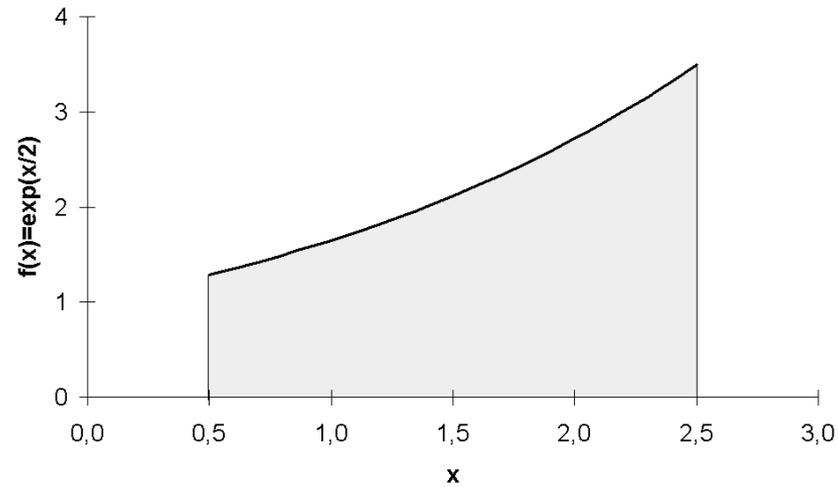
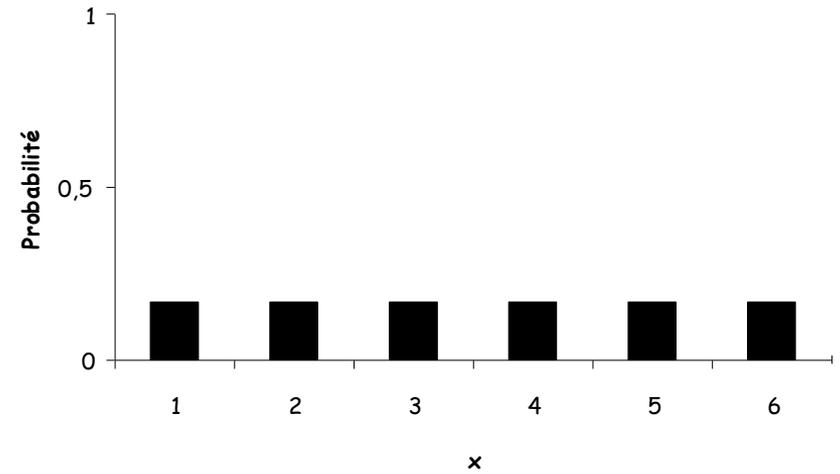
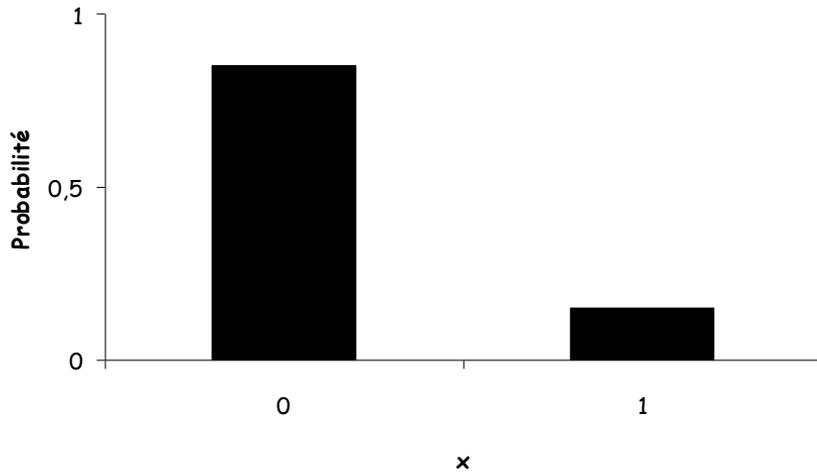
Variable Aléatoire : Exemple

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

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

où X est une va indicatrice des EI.

Caractéristiques d'une Variable Aléatoire



Caractéristique de Position

Moyenne, Espérance

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

$$\mu = E(X) = \sum p_i x_i$$

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

$$\mu = E(X) = \int_a^b x f(x) dx$$

Caractéristique de Position

Moyenne, Espérance

- Exemple 1 : $\mu = (p \times 1) + (q \times 0) = p$
- Exemple 2 : $\mu = 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6 = 3,5$

Caractéristique de Dispersion

Variance, Écart-type

- Cas d'une variable discrète X :

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

- Cas d'une variable continue X :

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

$\sigma^2 =$ variance

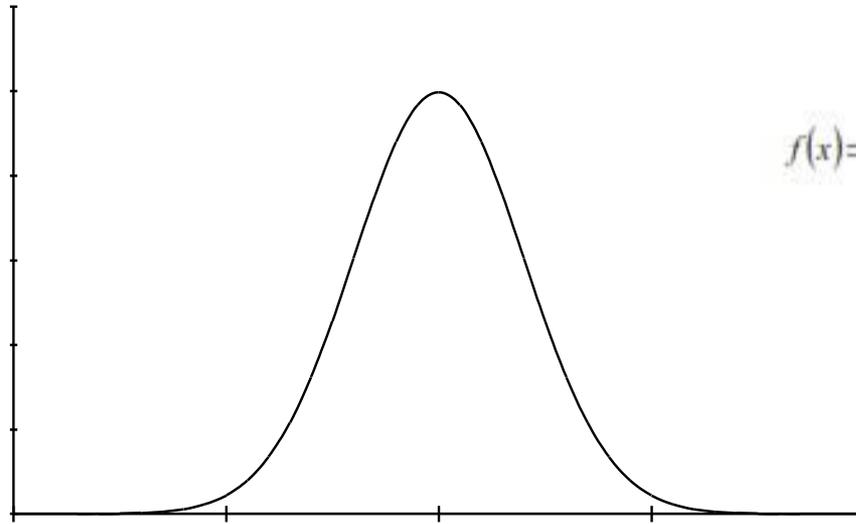
$\sigma =$ écart-type

Caractéristique de Dispersion

Variance, Écart-type

- Exemple 1 : $\sigma^2 = p \times (1 - p)^2 + q \times (0 - p)^2 = pq$
- Exemple 2 : $\sigma^2 = 1/6[(1 - 3,5)^2 + (2 - 3,5)^2 + \dots + (6 - 3,5)^2]$
 $= 2,9$

Loi Normale (Gauss)

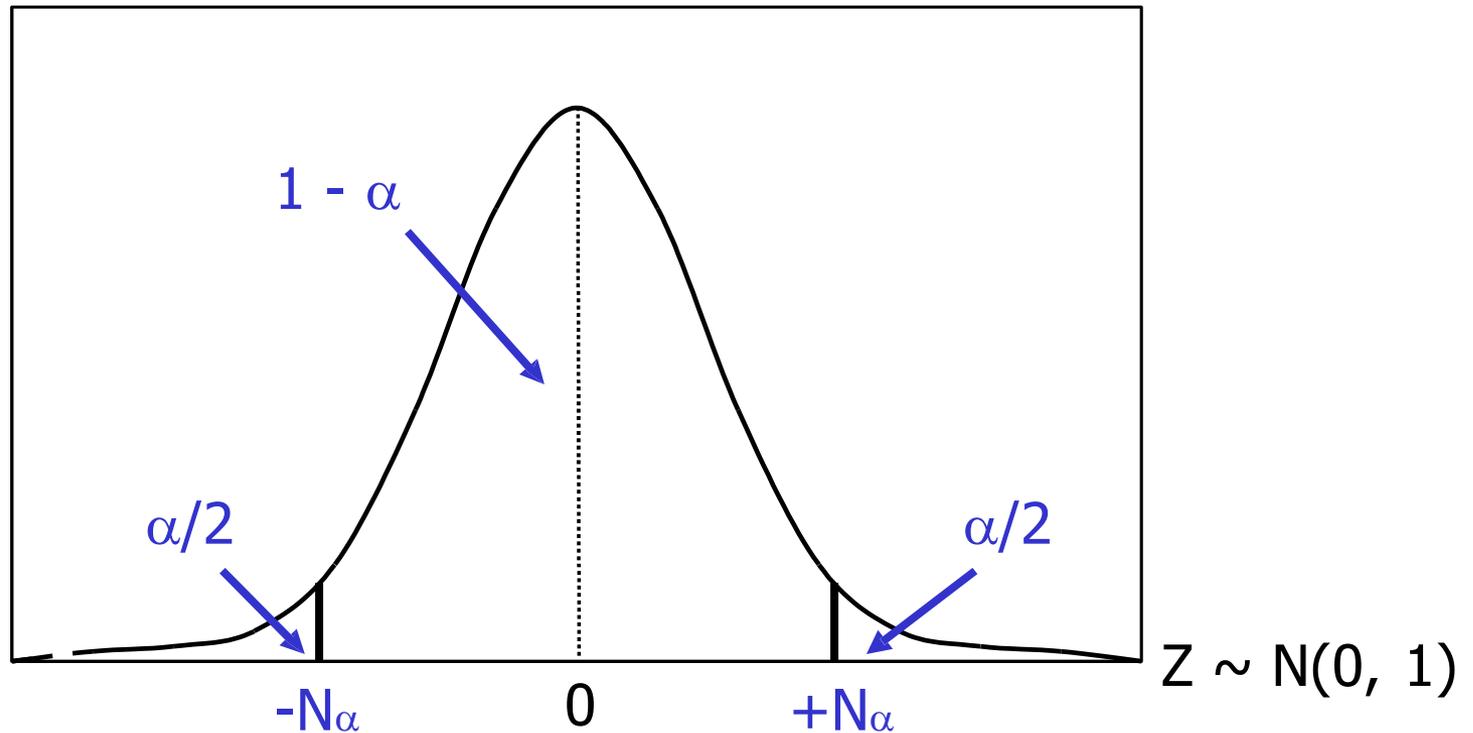


$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

- 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 $N(0, 1)$

| | $\alpha = 0,10$ | $\alpha = 0,05$ | $\alpha = 0,01$ |
|------------|-----------------|-----------------|-----------------|
| 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 $\nu \rightarrow \infty$

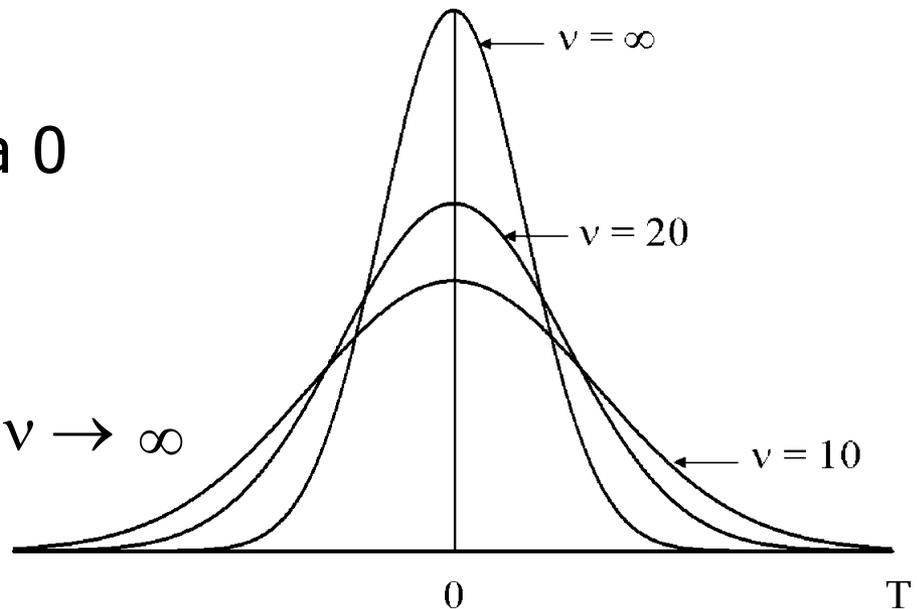


Table de la Loi de Student

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

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

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

Exemple :

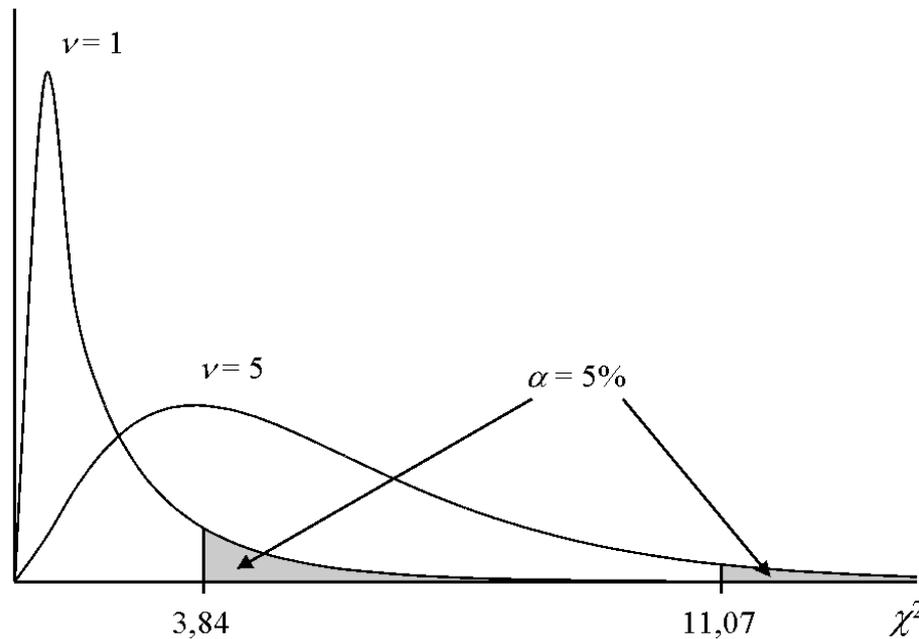
si $\alpha = 5 \%$ et $\nu = 10$ alors $T_{0,05; 10} = 2,228$

si $\alpha = 10 \%$ et $\nu = 8$ alors $T_{0,10; 8} = 1,860$

si $\alpha = 5 \%$ et $\nu \geq 30$ alors $T_{0,05; \geq 30} = 1,96$

Table de la Loi du Chi-deux

- 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 $\alpha = 5 \%$ et $\nu = 1$ alors $\chi_{0,05, 1}^2 = 3,84$

si $\alpha = 5 \%$ et $\nu = 5$ alors $\chi^2 =$

REFERENCES

1. Biostatistique connaissances de base , R. Giorgi
2. Biostatistiques pour clinicien, Michel Huguier
3. STA UNIV , SESSTIM, R. Giorgi