Première partie

Introduction/Préambule

La théorie des probabilités est l'étude des phénomènes **aléatoires**, c'est à dire soumis au hasard ou présentant une partie imprévisible. Les probabilités regroupent l'ensemble des connaissances **a priori** sur ces phénomènes ainsi que les **outils mathématiques** servant à les **modéliser**.
Par exemple, si on admet qu'un dé est équilibré, alors on sait qu'on a une chance sur six d'obtenir un six en le lançant **(a priori)**.

$\rightarrow$ Probabilités.

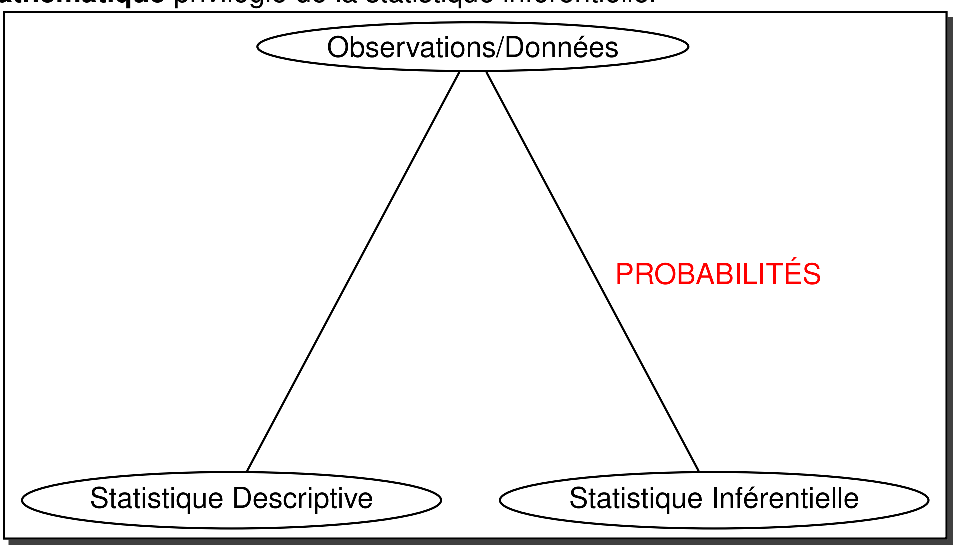
La statistique est au contraire une discipline ayant comme objectif le **traitement d'observations**, encore appelées données.

— La statistique descriptive est consacrée à l'étude, la représentation, la description des données (calcul d'une moyenne, d'un pourcentage, graphiques,...)

— La statistique inférentielle (ou inductive) consiste à tirer des conclusions **(a posteriori)** sur l'ensemble d'une population à partir des observations (échantillon prélevé). Par exemple on lance un dé 10 fois et on obtient 8 fois le chiffre 6 : le dé est-il truqué ?

$\rightarrow$ Statistique.

Les méthodes statistiques (de la statistique inférentielle) reposent essentiellement sur l'utilisation des modèles probabilistes. Les probabilités constituent donc **l'outil mathématique** privilégié de la statistique inférentielle.



— Probabilités : ensemble de résultats **théoriques**, connaissances **a priori**, outils de **modélisation**.
— Statistique : science basée sur les **observations** ou destinée à être appliquée à des observations.

Domaines d'application des probabilités :

Applications directes :
— Jeux de hasard (dés, cartes, roulette,...) : quelle est la probabilité de gagner à la roulette ? Comment répartir mes mises ? Combien puis-je espérer gagner en moyenne ?
— Économie, Finance (modèles économiques et financiers) : modélisation de l'évolution du cours d'une action, de données micro ou macro-économiques.

Applications par le biais des méthodes statistiques :
— Biologie, médecine (tests cliniques, études épidémiologiques) : un traitement contre une maladie est-il réellement plus efficace qu'un autre (test statistique) ?
— Finance, assurance (calcul de prime) : comment fixer le montant d'une prime en fonction des sinistres observés ?

Problème Probabiliste	Problème Statistique
Soit un test pour détecter une maladie	
On sait que la probabilité pour qu'une personne soit malade si le test est positif est de 0,8.	Dans une population de 100 000 personnes positives à un test, on en prélève un échantillon de 100. On observe que parmi ces 100 personnes, seules 80 sont malades.
Quelle est la probabilité pour qu'il y ait au moins 75 malades parmi 100 personnes ayant un test positif?	Que peut-on dire du pourcentage moyen de malades dans toute la population de personnes positives?
Quelle est le nombre moyen de malades sur 100 personnes positives?	Que peut-on dire de la probabilité qu'une personne tirée au hasard parmi les 100 000 positives soit malade?

Deuxième partie

Notions de Probabilités. Introduction aux Probabilités discrètes

1 Notions fondamentales, l'espace probabilisé

Au départ est une expérience aléatoire, c'est à dire dont on ne peut prédire le résultat à l'avance. Exemples :

— On lance un dé équilibré.

— On lance une pièce jusqu'à obtenir "‘face"’.

— On mesure la durée d'une réaction chimique (en micro-secondes).

1.1 L'espace des résultats possibles

Définition 1. On appelle espace des évènements élémentaires ou des éventualités, l'espace de tous les résultats possibles d'une expérience aléatoire. On le note généralement Ω. Exemples :

— $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$.

— $\Omega_2 = \{(F), (P, F), (P, P, F), \cdots, (P, \cdots, P, F), \cdots\} \cup \{(P, \cdots, P, \cdots)\}$.

— $\Omega_3 = \mathbb{R}^+$.

Ces trois ensembles présentent les caractéristiques suivantes :

— Ω_1 est fini.

— Ω_2 est infini dénombrable

— Ω_3 est (infini) non dénombrable

Définition 2. On dit qu'un ensemble E est dénombrable si on peut le compter, c'est à dire si on peut représenter tous ses éléments sous la forme d'une suite, autrement dit s'il existe une bijection de $\mathbb{N}$ dans E.

Exemples d'ensembles dénombrables : $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$. Les ensembles $\mathbb{R}, \mathbb{R}^+$ ainsi que tout intervalle de $\mathbb{R}$ ne sont pas dénombrables.

Définition 3. On appelle ensemble discret tout ensemble fini ou dénombrable. En probabilité, la distinction entre les ensembles discrets et les ensembles non-dénombrables est de la plus haute importance. En effet, on ne construit pas une probabilité de la même façon sur ces deux types d'ensembles.

1.2 Evènements

D'une manière intuitive, on appelle évènement toute propriété dont on peut dire si elle est vraie ou non une fois l'expérience aléatoire réalisée.
Plus formellement, un évènement est un sous-ensemble de l'espace Ω.
Mais il nous faut préciser cette notion.

Définition 4. (tribu) Soit un espace Ω. On appelle tribu sur Ω tout ensemble $\mathcal{F}$ de parties (sous-ensembles) de Ω tel que

1. $\Omega \in \mathcal{F}$.
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$.
3. Pour toute suite $(A_n)_{n\geq 1}$ d'éléments de $\mathcal{F}$, l'ensemble $\cup_{n=1}^{+\infty} A_n \in \mathcal{F}$.

Définition 5. Soit $(\Omega, \mathcal{F})$ un espace muni d'une tribu. On appelle évènement tout sous-ensemble de Ω appartenant à $\mathcal{F}$.
Remarque 1. Si Ω est discret, on verra plus tard qu'on peut construire une probabilité pour toute partie de Ω, ainsi dans ce cas, on peut prendre $\mathcal{F} = \mathcal{P}(\Omega)$ (ensembles des parties de Ω). Ainsi, dans le cas discret, tout sous-ensemble de Ω peut être considéré comme un évènement. **Exemples d'évènements :**

— Pour $\mathcal{E}_1$, lancer d'un dé équilibré :
$A_1 =$ " le résultat est pair " $= \{2, 4, 6\}$.
$A_2 =$ " le résultat est ≥ 5 " $= \{5, 6\}$.
$A_3 =$ " j'obtiens un résultat entre 1 et 6 " $= \Omega$.
$A_4 =$ " j'obtiens 7 " $= \emptyset$.

— Pour $\mathcal{E}_2$, lancer d'une pièce jusqu'à obtenir face :
$A =$ " on obtient face en 4 coups au plus " $= \{(F), (P, F), (P, P, F), (P, P, P, F)\}$.
$B =$ " il faut au moins 4 coups pour obtenir face " $= \{(P, P, P, F), (P, P, P, P, F), \ldots\}$ (dénombrable).

— Pour $\mathcal{E}_3$, mesure de la durée d'une réaction chimique en secondes :
$A =$ " la réaction dure moins d'1 minute " $=]0, 60]$ (non dénombrable).

REMARQUES :
- Un évènement constitué d'un unique élément (évènement élémentaire) est appelé **singleton**.
$\forall \omega \in \Omega$, on peut définir le singleton $\{\omega\}$.
- L'évènement certain est Ω.
- L'évènement impossible est $\emptyset$.

1.3 Evénements et logique ensembliste :

1.3.1 Opérations sur les ensembles

Soient A et B deux évènements.

Réunion :

L'évènement A ou (non exclusif) B est représenté par $A \cup B$:

$$A \cup B = \{\omega \in \Omega, \omega \in A \text{ ou } \omega \in B\}$$

Intersection :

L'évènement A et B est représenté par $A \cap B$:

$$A \cap B = \{\omega \in \Omega, \omega \in A \text{ et } \omega \in B\}$$

Complémentaire :

L'évènement contraire de A est représenté par le complémentaire de A noté $\overline{A}$ ou A^c ou $C_\Omega A$:

$$\overline{A} = \{\omega \in \Omega, \omega \notin A\}$$

Différence :

L'évènement A mais pas B est représenté par $A \backslash B$:

$$A \backslash B = \{\omega \in \Omega, \omega \in A \text{ et } \omega \notin B\}$$

Différence symétrique : L'évènement A ou (exclusif) B est représenté par $A \triangle B$:

$$A \triangle B = \{\omega \in \Omega, \omega \in A \text{ et } \omega \notin B \text{ ou } \omega \in B \text{ et } \omega \notin A\}$$

Remarque : $A \triangle B = (A \backslash B) \cup (B \backslash A) = (A \cup B) \backslash (A \cap B)$

1.3.2 Propriétés des opérations sur les ensembles

Distributivité

— Distributivité de $\cup$ par rapport à $\cap$:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

— Distributivité de $\cap$ par rapport à $\cup$:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Lois de Morgan

— Complémentaire d'une réunion : $\overline{A \cup B} = \overline{A} \cap \overline{B}$
— Complémentaire d'une intersection : $\overline{A \cap B} = \overline{A} \cup \overline{B}$
— Généralisation :

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A_i} \quad \text{et} \quad \overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i}$$

1.4 Probabilités

Le couple $(\Omega, \mathcal{F})$ est appelé espace probabilisable. Il ne reste plus qu'à lui adjoindre une probabilité pour obtenir un espace probabilisé.

Définition 6. On appelle probabilité sur $(\Omega, \mathcal{F})$ toute application notée le plus souvent P de $\mathcal{F}$ dans $[0, 1]$ et telle que
i)$P(\Omega) = 1$.
ii) Pour toute suite $(A_n)_{n \geq 1}$ d'éléments de $\mathcal{F}$ deux à deux disjoints, c'est à dire $A_i \cap A_j = \emptyset$, $\forall i \neq j$,

$$P(\cup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} P(A_n).$$

Définition 7. Le triplet $(\Omega, \mathcal{F}, P)$ est appelé espace probabilisé.

Probabilité sur un espace discret : Dans le cas où Ω est discret ($\Omega = \{\omega_1, \ldots, \omega_n, \ldots\}$), on définit une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ en définissant les valeurs de $P(\omega_i), \forall i$.
En effet, soit $A \in \mathcal{P}(\Omega)$, alors :

$$A = \bigcup_{i/\omega_i \in A} \{\omega_i\}$$

Or $\{\omega_i\} \cap \{\omega_j\} = \emptyset \quad \forall i \neq j \qquad \Rightarrow P(A) = \sum_{i/\omega_i \in A} P(\omega_i)$

est donc bien définie $\forall\ A \in \mathcal{P}(\Omega)$.

Cas particulier : Equiprobabilité

Définition 8. Soit $\Omega = \{\omega_1, \ldots, \omega_n\}$ un espace fini d'évènements élémentaires. On dit que les évènements sont **équiprobables** si :

$$P(\omega_1) = \cdots = P((\omega_n) = \frac{1}{n} = \frac{1}{\sharp\,\Omega}$$

Dans ce cas (d'équiprobabilité), $\forall A \subset \Omega : P(A) = \dfrac{\sharp\, A}{\sharp\,\Omega} = \dfrac{\sharp\, A}{n}$ **Exemples de probabilités**

— $\mathcal{E}_1 : \Omega_1 = \{1, 2, 3, 4, 5, 6\}$
  Le dé est équilibré, donc $P_1(1) = \cdots = P_1(6) = \dfrac{1}{6} = \dfrac{1}{\sharp\,\Omega_1}$
  On est dans un cas d'équiprobabilité (ou **probabilité uniforme**).

— $\mathcal{E}_2$: on lance une pièce jusqu'à obtenir face.
  $\Omega_7 = \{$(F),(P,F),(P,P,F), $\ldots$, (P,P,...,F),...$\} \cup \{$(P,...,P,...)$\}$
  $P_7((F)) = 1/2$, $P_7((P, F)) = 1/2^2, \ldots, P_7((P, \ldots, P, F)) = 1/2^n$ et $P_7((P, \ldots, P, \ldots)) = 0$.

Propriétés élémentaires des probabilités :

1. $P(\overline{A}) = 1 - P(A)$

2. $P(A \cup B) = P(A) + P(B \backslash A)$

3. Si $A \subset B$, alors $P(A) \leq P(B)$ et $P(B \backslash A) = P(B) - P(A)$

4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

5. Soit $A = \bigcup_{i=1}^{+\infty} A_i = A_1 \cup \cdots \cup A_n \cup \ldots$, alors

$$P(A) = \lim_{n \to +\infty} P(A_1 \cup \cdots \cup A_n)$$

6. Soit $A = \bigcap_{i=1}^{+\infty} A_i = A_1 \cap \cdots \cap A_n \cap \ldots$, alors

$$P(A) = \lim_{n \to +\infty} P(A_1 \cap \cdots \cap A_n)$$

1.5 Dénombrement

On note Ω l'ensemble des événements élémentaires, c'est à dire l'ensemble de toutes les réalisations possibles d'une expérience, de tous les "tirages" possibles. On note $\#A$ le nombre d'éléments (ou cardinal) d'un ensemble fini A . Si A est un événement, c'est-à-dire une partie de Ω, **et si les événements élémentaires sont supposés équiprobables (et seulement dans ce cas)**, alors sa probabilité est

$$P(A) = \frac{\#A}{\#\Omega}.$$

Modélisation

Pour "dénombrer" des cardinaux d'ensembles, on peut avoir recours à deux modèles équivalents :

• le modèle d'urne (U) : 1 urne contient n boules numérotées 1 à n. On en tire p,

–avec ou sans remise,

–en tenant compte de l'ordre des tirages ou non.

• le modèle de boîtes et de boules (B) : 1 boîte est divisée en n petites boîtes ou compartiments numéroté(e)s 1 à n. On place p boules dans ces n boîtes,

–sans limitation du nombre de boules dans chaque boîte ou (au contraire) avec au plus une boule par boîte,

–les boules sont numérotées ou non.

EXEMPLE : On choisit p boules dans une urne composée de n boules distinctes, numérotées de 1 à n. On veut déterminer le nombre d'issues possibles. Il faut distinguer 4 cas selon qu'on tire avec ou sans remise et selon qu'on tient compte ou non de l'ordre dans lequel les boules sont tirées.

Arrangements sans répétition : On tire sans remise et en tenant compte de l'ordre de tirage des boules⇔On place p boules numérotées dans n boites avec au plus une boule par boite.

Définition 9. Soient $E = \{e_1, \dots, e_n\}$ un ensemble à $n \geq 1$ éléments et $p \in \{1, \dots, n\}$.
Un *arrangement sans répétition* de p éléments de E ou p-*liste sans répétition d'éléments de* E est un p-uplet $(e_{i_1}, \dots, e_{i_p}) \in E^p$ où les e_{i_j} sont deux à deux distincts.

Proposition 1. *L*e nombre des arrangements sans répétition de p éléments pris parmi $n \geq 1$, $1 \leq p \leq n$ est :

$$A_n^p = n(n-1)\dots(n-p+1) = \frac{n!}{(n-p)!}$$

où on utilise la convention $0! = 1$.
Pour $n \in \mathbb{N}$, on pose aussi $A_n^0 = 1$ et $A_n^p = 0$ si $p > n$ ou $p < 0$.

Exemple. Il y a A_n^3 tiercés possibles dans l'ordre lorsqu'il y a n chevaux au départ d'une course.

Permutations : On tire toute les boules, donc $n = p$.

Définition 10. 2 : Une *permutation* (sans répétition) d'un ensemble E de cardinal $n \geq 1$ est un arrangement sans répétition des n éléments de E.

Proposition 2. *L*e nombre de permutations de n éléments est $n! = A_n^n$.

Exemple. Le nombre de rangements de n livres sur une étagère est $n!$.

Combinaisons sans répétition : On tire sans remise et sans tenir compte de l'ordre de tirage⇔On place p boules non numérotées dans n boites avec au plus une boule par par boite..

Définition 11. Soient $E = \{e_1, \ldots, e_n\}$ un ensemble de cardinal $n \geq 1$ et $p \in \{0, \ldots, n\}$.
Une *combinaison sans répétition* de p éléments de E est une partie (sous-ensemble) $\{e_{i_1}, \ldots, e_{i_p}\}$ de E ayant p éléments.

Proposition 3. *L*e nombre de combinaisons de p éléments pris parmi $n \geq 0$, $1 \leq p \leq n$ est :

$$C_n^p = \begin{pmatrix} n \\ p \end{pmatrix} = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}$$

On pose aussi $C_n^p = 0$ si $p > n$ ou $p < 0$.

Exemple. Il y a C_{49}^6 résultats possibles los d'un tirage du loto.

Arrangements avec répétition : On tire avec remise et en tenant compte de l'ordre de tirage⇔On place p boules numérotées dans n boites sans limitation du nombre de boules par boite.

Définition 12. 4 : Soient $E = \{e_1, \ldots, e_n\}$ un ensemble à $n \geq 1$ éléments et $p \in \{1, \ldots, n\}$.
Un *arrangement avec répétition* de p éléments de E ou p*-liste avec répétition d'éléments de* E est un p-uplet $(e_{i_1}, \ldots, e_{i_p}) \in E^p$.

Proposition 4. *L*e nombre de p-listes avec répétition de n éléments est n^p. (Card(E^p)= Card $(E)^p$).

Exemple. Il y a $10^4 = 10000$ codes possibles pour les cartes banquaires.

Combinaisons avec répétition : On tire avec remise et sans tenir compte de l'ordre de tirage⇔ On place p boules non numérotées dans n boites sans limitation du nombre de boules par boite.

Proposition 5. *L*e nombre de combinaisons avec répétition de p éléments pris parmi n est :

$$C_{n+p-1}^p = C_{n+p-1}^{n-1} = \Gamma_n^p$$

Exemple. Il y a $\Gamma_6^2 = C_7^2 = 21$ configurations possibles en lançant 2 dés indiscernables.

Tableau résumé

	sans remise	avec remise
avec ordre	A_n^p (arrangements sans répétition)	n^p (arrangements avec répétition)
sans ordre	C_n^p (combinaisons sans répétition)	Γ_n^p (combinaisons avec répétition)

Permutations avec répétition

Définition 13. 5 : Soient $E = \{e_1, \dots, e_r\}$ de cardinal $r \geq 2$, $n \geq 2$ et $k_1, \dots, k_r \in \{0, \dots, n\}$ tels que $k_1 + k_2 + \dots + k_r = n$.
Une *permutation avec répétition* de n éléments pris dans E où figurent k_1 fois $e_1, \dots, k_r$ fois e_r est un n-uplet $(e_{i_1}, \dots, e_{i_n}) \in E^n$ où figurent k_1 fois $e_1, \dots, k_r$ fois e_r.

Proposition 6. *L*e nombre de permutations avec répétition de $n \geq 2$ éléments pris parmi $E = \{e_1, \dots, e_r\}$ où figurent k_1 fois $e_1, \dots, k_r$ fois e_r, avec $k_1 + k_2 + \dots + k_r = n$ est :

$$\begin{pmatrix} n \\ k_1 \dots k_r \end{pmatrix} = \frac{n!}{k_1!k_2!\dots k_r!}$$

Exemple. Le nombre de permutations des chiffres $1, 1, 1, 3, 3, 5, 6, 6, 6, 6$ est : $\begin{pmatrix} 10 \\ 3\,2\,1\,4 \end{pmatrix} = \frac{10!}{3!2!4!}$.

2 Indépendance et Probabilités conditionnelles

2.1 Indépendance

Indépendance de deux évènements

Intuitivement, deux évènements A et B sont indépendants pour la probabilité P, si la probabilité de réalisation de l'un n'est pas influencée par la réalisation de l'autre.

Définition 14. Deux évènements A et B sont <u>indépendants pour la probabilité P</u> si P(A∩B)=P(A)P(B).

Remarque : L'indépendance n'est pas une qualité intrinsèque des évènements, elle dépend de la probabilité considérée.
Exemple : Soit Ω = {1,2,3,4,5,6} muni de la tribu P(Ω). Soient A et B les évènements : A={1,2} et B={2,3}. Soient P_1 et P_2 les probabilités définies sur (Ω,P(Ω)) par :

i	1	2	3	4	5	6
$P_1(\{i\})$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$P_2(\{i\})$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

1. A et B sont-ils indépendants pour la probabilité P_1 ?
2. A et B sont-ils indépendants pour la probabilité P_2 ?

Proposition 7. *S*i A et B sont indépendants pour la probabilité P, A et $\bar{\text{B}}$, $\bar{\text{B}}$ et B, $\bar{\text{A}}$et $\bar{\text{B}}$ sont indépendants pour P.

Indépendance d'une famille d'évènements

Définition 15. On dit que n évènements $A_1, \ldots, A_n$ sont <u>(mutuellement) indépendants</u> (ou <u>indépendants dans leur ensemble</u>) <u>pour la probabilité P</u> si :

$$\forall \mathrm{k},\ 2 \leq \mathrm{k} \leq \mathrm{n},\quad \forall\ \mathrm{i}_1, \ldots, \mathrm{i}_\mathrm{k}\ \ :\ \mathrm{P}\left(A_{i_1} \bigcap \ldots \bigcap A_{i_k}\right) = P(A_{i_1}) \ldots P(A_{i_k})$$

Exemple : Trois évènements A, B, C sont indépendants si et seulement si :

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(A \cap C) &= P(A)P(C) \\ P(B \cap C) &= P(B)P(C) \\ P(A \cap B \cap C) &= P(A)P(B)P(C) \end{aligned}$$

Définition 16. On dit que n évènements $A_1, \ldots, A_n$ sont deux à deux indépendants pour la probabilité P si : $P(A_i \cap A_j) = P(A_i)P(A_j)$ pour tout (i, j) tel que i$\neq$j.
Remarque : Si $(A_i)i \in I$ est une famille d'évènements mutuellement indépendants, alors $(A_i)i \in I$ est une famille d'évènements deux à deux indépendants. La réciproque n'est pas vraie en général.
Exemple : On lance deux fois un dé cubique parfait. Soient les évènements :
A_1 : « le premier nombre obtenu est pair »
A_2 : « le deuxième nombre obtenu est impair »
A_3 : « la somme des deux nombres obtenus est paire » Les évènements A_1, A_2 , A_3 sont-

ils deux à deux indépendants ? Mutuellement indépendants ?

Propriété 1 - Si $(A_i)i \in I$ est une famille d'évènements mutuellement indépendants, alors (A_i')i∈I où A_i' désigne A_i ou $\overline{A_i}$ est également une famille d'évènements mutuellement indépendants.

Epreuves répétées

On postule souvent l'indépendance en probabilités, notamment dans le cas où une expérience aléatoire est reproduite plusieurs fois à la suite.
Exemple : On lance un dé équilibré 2 fois de suite. Alors, par définition, les évènements relatifs au premier dé sont indépendants de ceux relatifs au second.
Si A = {le premier dé donne un résultat pair} et B = { le 2ème dé donne 6}, alors A et B sont indépendants.

2.2 Probabilités conditionnelles

2.2.1 Introduction

On lance une fois un dé (non truqué).
Soit A l'évènement : « on obtient un nombre inférieur ou égal à 5 ».
Soit B l'évènement : « on obtient un nombre supérieur ou égal à 3 ».
Supposons que l'on sache que A est réalisé. Le résultat du lancer est donc un élément ω de {1,2,3,4,5} et la probabilité que B soit réalisé sachant que A est réalisé est $\frac{3}{5}$.
Or on a : P(A)=$\frac{5}{6}$, P(A∩B)=$\frac{3}{6}$ et $\frac{3}{5} = \frac{P(A \bigcap B)}{P(A)}$

Probabilité conditionnelle

Définition 17. Soit (Ω, F, P) un espace probabilisé et soit A un évènement de probabilité non nulle. Pour tout évènement B on pose

$$P_A(B) = P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

P_A est une nouvelle probabilité appelée <u>probabilité conditionnelle relative à A</u> ou <u>probabilité sachant A</u>.

1. $P(\bar{B}/A)$=1 – P(B/A)
2. $P(B_1 \cup B_2/A) = P(B_1/A) + P(B_2/A) - P(B_1 \cap B_2/A)$

Propriété 2 - Relation entre indépendance et probabilités conditionnelles
Si P(A)$>$0 et P(B)$>$0, alors A et B indépendants ⇔ P(B|A) = P(B) et P(A|B) = P(A).

Propriété 3 - Soit (A_i)1≤i≤n une famille d'évènements tels que $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \neq 0$.
Alors :

$$\mathrm{P}\left(\bigcap_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{A_i}\right) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2) \ldots P(A_n/A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Remarque : Toutes les probabilités conditionnelles écrites ont un sens car :
∀ j ∈ {1,..,n-1} : $(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \subset (A_1 \cap A_2 \cap \cdots \cap A_j)$ donc $P(A_1 \cap A_2 \cap \cdots \cap A_j) \neq 0$.

2.3 Système complet d'évènements, formule des probabilités totales, formule de Bayes

Définition 18. On dit que n évènements $A_1, \ldots, A_n$ forment un **système complet d'évènements** dans Ω si :

$$(1) \bigcup_{i=1}^{n} A_i = \Omega$$

(2) $\mathrm{A_i} \bigcap \mathrm{A_j} = \emptyset$ pour tout (i, j)tel que i $\neq$ j

Autrement dit, ils forment une partition de Ω.
Exemple : Si A $\subset \Omega$: {A, $\bar{\mathrm{A}}$} est un système complet d'évènements de Ω.

Formule des probabilités totales

Propriété 4 - Soit E $\in$ P (Ω) un évènement quelconque et soit $A_1, \ldots, A_n$ un système complet d'évènements de probabilités non nulles. Alors :
P(E) = $\sum_{i=1}^{n} P(E \bigcap A_i) = \sum_{i=1}^{n} P(E/A_i)P(A_i)$

Formule de Bayes

Propriété 5 - Soit E ∈ P (Ω) un évènement de probabilité non nulle et soit $A1, \ldots, An$ un système complet d'évènements de probabilités non nulles. Alors :
Pour tout i∈{1,..,n} , on a :

$$P(A_i|E) = \frac{P(E|A_i)P(A_i)}{\sum_{j=1}^{n} P(E|A_j)P(A_j)}.$$

Exemple fondamental :

On dispose de deux urnes :

1. L'urne 1 contient une proportion égale de boules blanches et de boules noires
2. L'urne 2 contient $\frac{3}{4}$ de boules blanches et $\frac{1}{4}$ de boules noires

Expérience : on sélectionne au hasard une des 2 urnes et on tire une boule.

1. (a) Quelle est la probabilité de tirer une boule blanche ?
   (b) Si la boule tirée est blanche, quelle est la probabilité qu'on ait choisi l'urne 2 ?

2.4 Indépendance conditionnelle

Définition 19. Soient A, B et E des évènements avec P(E)>0.
On dit que A et B sont indépendants conditionnellement à E si A et B sont indépendants pour la probabilité conditionnelle P_E :

$$P(A \cap B|E) = P(A|E)P(B|E)$$

Exemple : On reprend l'exemple des 2 urnes et on tire cette fois-ci 2 boules de l'urne sélectionnée, avec remise.

3 Variables aléatoires

Définition 20. Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et soit $(E, \mathcal{G})$ un espace probabilisable. On appelle variable aléatoire toute application de Ω dans E telle que $\forall B \in \mathcal{G},\ X^{-1}(B) \in \mathcal{F}$.

Une variable aléatoire est donc une fonction sur Ω. Si $E = \mathbb{R}$, on parle de variable aléatoire réelle. Si E (ou $X(\Omega)$) est discret, on parle de variable aléatoire discrète. Si la variable X peut prendre ses valeurs dans un ensemble non-dénombrable, on parlera de variable aléatoire continue (une définition rigoureuse sera donnée par la suite) Notons que E peut être également un ensemble de modalités, par exemple la variable aléatoire X=couleur des yeux avec E={vert,bleu,marron}. On peut parler ainsi de variables qualitatives. Faisons la correspondance avec les caractères statistiques :

caractère qualitatif	v.a. "qualitative"
caractère quantitatif discret	v.a. discrète
caractère quantitatif continu	v.a. continue

Définition 21. Loi de probabilité d'une variable aléatoire

Soit X une variable aléatoire $(\Omega, \mathcal{F}, P)$ dans $(E, \mathcal{G})$. On appelle loi de probabilité de la variable X la probabilité P_X sur $(E, \mathcal{G})$ telle que $\forall B \in \mathcal{G}$,

$$P_X(B) = P(X^{-1}(B)) = P(X \in B).$$

Remarque 2. Détermination de la loi de probabilité d'une v.a. discrète. Soit $X(\Omega) = \{x_1, \dots, x_n, \dots\}$ l'ensemble des valerus prises par la v.a. X. Alors pour déterminer entièrement la loi de probabilité de X, il suffit de déterminer $P(X = x_1), \dots, P(X = x_n), \dots$. Il en va de même pour une variable "'qualitative"'.

Exemple :

On lance un dé équilibré : $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Soit :

$$\begin{aligned} X : \ & \Omega \to \mathbb{R} \\ & \omega \to X(\omega) = \begin{cases} 0 \text{ si } \omega \text{ est pair} \\ 1 \text{ si } \omega \text{ est impair} \end{cases} \end{aligned}$$

Déterminons la loi de probabilité de X.

On a ici $P_X(0) = P(\{2, 4, 6\}) = 1/2$ et $P_X(1) = P(\{1, 3, 5\}) = 1/2$

Définition 22. Soit $A \in \mathcal{F}$ un évènement quelconque. On appelle **fonction indicatrice de** A (ou indicatrice de A) la v.a. notée $\mathbb{1}_A$ et définie par :

$$\begin{aligned} \mathbb{1}_A : \ & \Omega \to \mathbb{R} \\ & \omega \to \mathbb{1}_A(\omega) = \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \notin A \end{cases} \end{aligned}$$

Les lois de probabilité des indicatrices sont appelées **lois de Bernoulli**.

Autre exemple : On lance 2 dés équilibrés (discernables).

$\Omega = \{1, 2, 3, 4, 5, 6\}^2$.

Soit

$$\begin{aligned} X_1 : \quad & \Omega \to \mathbb{R} \\ & \omega = (\omega_1, \omega_2) \to X_1(\omega) = \omega_1 \end{aligned}$$

X_1 est la v.a. “résultat du premier dé”.

On peut de même définir :

$$\begin{aligned} X_2 : \quad & \Omega \to \mathbb{R} \\ & \omega = (\omega_1, \omega_2) \to X_2(\omega) = \omega_2 \end{aligned}$$

X_2 est la v.a. “résultat du deuxième dé”.

On peut définir :

$$\begin{array}{rcl} S: & \Omega & \to \mathbb{R} \\ & \omega = (\omega_1, \omega_2) & \to S(\omega) = \omega_1 + \omega_2 \end{array}$$

c'est-à-dire $S = X_1 + X_2 =$ somme des marques des deux dés.
On peut également définir $Z = X_1 X_2 =$ produit des marques des deux dés etc...

3.1 Fonction de répartition - Indépendance de variables aléatoires

3.1.1 Fonction de répartition

Définition 23. Soit X une v.a réelle. On appelle **fonction de répartition** de la loi de probabilité de X la fonction définie par :

$$\begin{array}{rcl} F = F_X : & \mathbb{R} & \to [0, 1] \\ & x \in \mathbb{R} & \to F(x) = P(X \leq x) \end{array}$$

Conséquences :

— Si $X(\Omega) = \{x_1, \ldots, x_n, \ldots\}$ dénombrable avec $x_1 < x_2 < \cdots < x_n < \ldots$, alors si on pose $p_j = P(X = x_j)$, on a :

$$F_X(x) = F(x) = P(X \leq x) = \begin{cases} 0 & \text{si } x < x_1 \\ \sum_{j=1}^{i} p_j = P(X \leq x_i) & \text{si } x_i \leq x < x_{i+1} \end{cases}$$

— Si $X(\Omega) = \{x_1, \ldots, x_n\}$ (fini) avec $x_1 < x_2 < \cdots < x_n$, alors on a de plus : $F(x) = 1 \ \ \forall x \geq x_n$.

Propriétés :

1. La connaissance de la loi de probabilité P_X de la v.a. X est équivalente à celle de sa fonction de répartition F_X.
   Dans le cas où $X(\Omega) = \{x_1, \ldots, x_n, \ldots\}$ avec $x_1 < x_2 < \cdots < x_n < \ldots$, alors :
   $$P(X = x_i) = F_X(x_i) - F_X(x_{i-1})$$
2. F est croissante sur $\mathbb{R}$: si $x < y$ alors $F(x) \leq F(y)$
3. F est continue à droite en tout point : $\lim\limits_{\varepsilon \to 0} F(x + \varepsilon) = F(x)$
4. $\lim\limits_{x \to -\infty} F(x) = 0$ et $\lim\limits_{x \to +\infty} F(x) = 1$

3.1.2 Indépendance de variables aléatoires

Notation : Soient X et Y deux variables aléatoires et A et B deux évènements. L'évènement $(X \in A) \cap (Y \in B)$ sera noté plus simplement

$$(X \in A, Y \in B)$$

.

Définition 24. Soient X et Y deux v.a. réelles. On dit que X et Y sont **indépendantes** si :

$$\forall (x, y) \in \mathbb{R}^2 : P(X \leq x, Y \leq y) = P(X \leq x) \times P(Y \leq y)$$

$$\Leftrightarrow \forall (x, y) \in \mathbb{R}^2 : F_{X,Y}(x, y) = F_X(x) \times F_Y(y)$$

Si X et Y sont discrètes, alors X et Y sont indépendantes si :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega) : P(X = x, Y = y) = P(X = x) \times P(Y = y)$$

Définition 25. Soient $X_1, \ldots, X_n$ n variables aléatoires réelles. On dit que $X_1, \ldots, X_n$ sont **(mutuellement) indépendantes** si :

$$\forall (x_1, \ldots, x_n) \in \mathbb{R}^n : P(X_1 \leq x_1, \ldots X_n \leq x_n) = P(X_1 \leq x_1) \times \cdots \times P(X_n \leq x_n)$$

Si $X_1, \ldots, X_n$ sont discrètes, alors $X_1, \ldots, X_n$ sont (mutuellement) indépendantes si :

$$\forall (x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega) : P(X_1 = x_1, \ldots X_n = x_n) =$$

$$P(X_1 = x_1) \times \cdots \times P(X_n = x_n)$$

Définition 26. On dit que n v.a. sont **indépendantes 2 à 2** si $\forall i \neq j \;\; X_i$ et X_j sont indépendantes.

Propriétés :

1. Si $X_1, \ldots, X_n$ sont (mutuellement) indépendantes, alors $\forall \; i_1, \ldots, i_k \in \{1, \ldots, n\}$, $X_{i_1}, \ldots, X_{i_k}$ sont (mutuellement) indépendantes.
2. Soit X et Y deux v.a. discrètes indépendantes. Alors :

$$\forall A \in \mathcal{P}(X(\Omega)), \forall B \in \mathcal{P}(Y(\Omega)) :$$

$$P(X \in A, Y \in B) = P(X \in A) \times P(Y \in B)$$

3. Soient A et B deux évènements.
   A et B indépendants $\Leftrightarrow$ $\mathbb{1}_A$ et $\mathbb{1}_B$ sont indépendantes.
   De même, soient $A_1, \ldots, A_n$ n évènements :
   $A_1, \ldots, A_n$ mutuellement indépendants $\Leftrightarrow$ $\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_n}$ sont mutuellement indépendantes.

Théorème 1 $-$ *S*oient X et Y deux v.a. indépendantes et f et g deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Alors $f(X)$ et $g(Y)$ sont indépendantes.
Exemple : On dispose d'une urne contenant 2 boules blanches et 2 boules noires. On en tire 2.

(a) avec remise

(b) sans remise

Soient

$$X_1 = \begin{cases} 1 & \text{si la 1ère boule tirée est blanche} \\ 0 & \text{si la 1ère boule tirée est noire} \end{cases}$$

et

$$X_2 = \begin{cases} 1 & \text{si la 2ème boule tirée est blanche} \\ 0 & \text{si la 2ème boule tirée est noire} \end{cases}$$

Théorème 2 $-$ *S*oit une urne contenant M boules blanches et $N - M$ boules noires (soit au total N boules). On effectue des tirages sans remise.
Soit $\forall i, 1 \leq i \leq N : X_i = \begin{cases} 1 & \text{si la } i\text{ème boule tirée est blanche} \\ 0 & \text{si la } i\text{ème boule tirée est noire} \end{cases}$ Alors les X_i sont identiquement distribuées (c'est-à-dire de même loi de probabilité), de loi de Bernoulli $\mathcal{B}(1, p = \frac{M}{N})$ (mais les X_i ne sont pas indépendantes).

4 Espérance mathématique, variance, moments, inégalités probabilistes

4.1 Espérance mathématique : définition et propriétés

Définition 27. Soit X une v.a. discrète réelle et soit $X(\Omega)$ l'ensemble des valeurs prises par X. On dit que X est d'espérance finie si

$$\sum_{x_i \in X(\Omega)} |x_i| P(X = x_i) < +\infty$$

et on appelle alors **espérance de** X le nombre défini par :

$$\mathbb{E}(X) = \sum_{x_i \in X(\Omega)} x_i P(X = x_i)$$

REMARQUE : On voit que $\mathbb{E}(X)$ n'est autre que la moyenne des valeurs prises par X pondérées par les probabilités correspondantes. Donc $\mathbb{E}(X)$ représente la valeur moyenne (a priori) de la v.a. X, autrement dit le résultat moyen qu'on peut "espérer" obtenir en observant une réalisation de la v.a. X.

Exemples :

1. Si $X \equiv a$ constante, alors $\mathbb{E}(X) = a$.
2. Soit $A \in \mathcal{P}(\Omega)$ et $\mathbb{1}_A$ l'indicatrice de A : $\mathbb{1}_A(\omega) = \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ sinon} \end{cases}$

   Alors $E(X) = P(A)$.
3. Soient $X =$ résultat du lancer d'un dé équilibré. Alors

$$E(X) = \sum_{i=1}^{6} iP(X = i) = \sum_{i=1}^{6} i\frac{1}{6} = \frac{1}{6}\sum_{i=1}^{6} i = 3,5.$$

Propriétés :

1. Si $X \geq 0$, alors $\mathbb{E}(X) \geq 0$.
2. Soit $X \geq 0$. Alors : $\mathbb{E}(X) = 0 \Leftrightarrow P(X = 0) = 1$

Théorème 3 $-$ *Théorème du transfert.*
Soient $X_1, \ldots, X_n$ n v.a. discrètes réelles et soit $f : \mathbb{R}^n \to \mathbb{R}$. On suppose que la v.a. réelle $Z = f(X_1, \ldots, X_n)$ est d'espérance finie. Alors :

$$\mathbb{E}(Z) = \sum_{(x_1,\ldots,x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)} f(x_1, \ldots, x_n) P(X_1 = x_1, \ldots, X_n = x_n)$$

Rappel : Par définition, $\mathbb{E}(Z) = \sum_{z_i \in Z(\Omega)} z_i.P(Z = z_i)$
La difficulté réside dans le calcul de $P(Z = z_i)$, $\forall i$.

Exemple : $(n = 1)$
Soit $Z = f(X)$. Alors $\mathbb{E}(Z) = \sum_{x_i \in X(\Omega)} f(x_i).P(X = x_i)$

Applications du théorème du transfert :

(a) Soit X d'espérance finie et $\lambda \in \mathbb{R}$ alors λX est d'espérance finie et :

$$\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$$

(b) Soient X et Y d'espérance finie, alors $X + Y$ est d'espérance finie et :

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

Conséquence : Si $X \leq Y$ alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$

Généralisation : Soient $X_1, \dots, X_n$ n v.a. réelles d'espérance finie alors $X_1 + \dots + X_n$ est d'espérance finie et :

$$\mathbb{E}(X_1 + \dots + X_n) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_n)$$

Exemple : On lance 2 dés équilibrés discernables.
Soient $X_1 =$ résultat du 1er dé, $X_2 =$ résultat du 2ème dé et $S = X_1 + X_2$ la somme des 2 dés.
Alors $E(S) = E(X_1) + E(X_2) = 3,5 + 3,5 = 7$.
On a aussi :

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$$
$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$$

(c) Soient X et Y deux v.a. réelles **INDÉPENDANTES** d'espérance finie, alors XY est d'espérance finie et :

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Généralisation : Soient $X_1, \dots, X_n$ n v.a. réelles (mutuellement) **indépendantes** d'espérance finie alors $X_1 \times \dots \times X_n$ est d'espérance finie et :

$$\mathbb{E}(X_1 \times \dots \times X_n) = \mathbb{E}(X_1) \times \dots \times \mathbb{E}(X_n)$$

4.2 Moments d'ordre supérieur - Variance

Définition 28. On dit qu'une v.a. discrète réelle X admet un **moment non centré d'ordre** k si X^k est d'espérance finie, c'est-à-dire si

$$E(|X^k|) = \sum_{x_i \in X(\Omega)} |x_i|^k P(X = x_i) < +\infty$$

et on le définit par :

$$\mathbb{E}(X^k) = \sum_{x_i \in X(\Omega)} x_i^k P(X = x_i)$$

REMARQUE : $\mathbb{E}(X)$ est le moment non centré d'ordre 1.

Définition 29. On dit qu'une v.a. discrète réelle X admet un **moment centré d'ordre** k si la v.a. $(X - \mathbb{E}(X))^k$ est d'espérance finie et on le définit par :

$$\mathbb{E}[(X - \mathbb{E}(X))^k] = \sum_{x_i \in X(\Omega)} (x_i - E(X))^k P(X = x_i)$$

Définition 30. On dit qu'une v.a. discrète réelle X est de **variance finie** si elle admet un moment centré d'ordre 2 et on appelle alors variance de X ce moment centré d'ordre 2, c'est-à-dire le nombre :

$$\mathbf{Var}X = \sigma_X^2 = \mathbb{E}[(X - \mathbb{E}(X))^2] = \sum_{x_i \in X(\Omega)} (x_i - E(X))^2 P(X = x_i)$$

REMARQUE : La variance est donc l'espérance (la moyenne) des écarts à la moyenne élevés au carré.

Définition 31. On appelle **écart-type** de X la quantité :

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{\mathbf{Var}X}$$

Propriétés :

1. $\mathbf{Var}X \geq 0$ et $\mathbf{Var}X = 0 \Leftrightarrow \exists b \in \mathbb{R}$ t.q $P(X = b) = 1$

2. Soit X de variance finie et soient a et $b \in \mathbb{R}$. Alors :

$$\mathbf{Var}(aX + b) = a^2 \mathbf{Var}X$$

ou ce qui est équivalent : $\sigma_{aX+b} = |a| \sigma_X$

3. **Formule de calcul de la variance : formule de Koenig**
   Soit X une v.a. de variance finie, alors :
   $$\sigma_X^2 = \mathbf{Var}X = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$
   REMARQUE : $\mathbb{E}(X^2) = \sum_{x_i \in X(\Omega)} x_i^2 P(X = x_i)$
4. Soient X et Y 2 v.a. **indépendantes** de variance finie. Alors $X + Y$ est de variance finie et :
   $$\mathbf{Var}(X+Y) = \mathbf{Var}X + \mathbf{Var}Y$$
5. Généralisation : Si $X_1, \dots, X_n$ sont des v.a. réelles **2 à 2 indépendantes** et de variance finie, alors $X_1 + \cdots + X_n$ est de variance finie et :
   $$\mathbf{Var}(X_1 + \cdots + X_n) = \mathbf{Var}(X_1) + \cdots + \mathbf{Var}(X_n)$$
6. Si X admet un moment non centré d'ordre k, alors X admet un moment non centré d'ordre $k-1$ (et donc d'ordre $k-2, \dots, 1$)
7. X admet un moment non centré d'ordre $k \Leftrightarrow \forall a \in \mathbb{R},\ \mathbb{E}(|X+a|^k) < +\infty$ (i.e. $\forall a \in \mathbb{R},\ X + a$ admet un moment non centré d'ordre k).

Conséquence : X admet un moment centré d'ordre $k \Leftrightarrow X$ admet un moment non centré d'ordre k.

4.3 Inégalités probabilistes

(a) **Inégalité de Markov**
Soit X une v.a. d'espérance finie, alors :
$$\forall \lambda > 0,\ P(|X| \geq \lambda) \leq \frac{\mathbb{E}(|X|)}{\lambda}$$

(b) **Inégalité de Bienaymé Tchebychev :**
Soit X une v.a. de variance finie, alors :
$$\forall \varepsilon > 0,\ P(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbf{Var}X}{\varepsilon^2}$$
$$\Leftrightarrow \forall \varepsilon > 0,\ P(|X - \mathbb{E}(X)| < \varepsilon) \geq 1 - \frac{\mathbf{Var}X}{\varepsilon^2}$$
$$\Leftrightarrow \forall \varepsilon > 0,\ P(\mathbb{E}(X) - \varepsilon < X < \mathbb{E}(X) + \varepsilon) \geq 1 - \frac{\mathbf{Var}X}{\varepsilon^2}$$

L'inégalité de Bienaymé Tchebychev permet d'obtenir des ordres de grandeur des probabilités des écarts à la moyenne sans autre connaissance que celle de l'espérance et de la variance de la loi de X.

Exemple : Soit X une v.a. de moyenne (espérance) 100 et de variance 16. Posons $\varepsilon = 8$. Alors

$$P(|X - 100| \geq 8) \leq \frac{16}{64} = 0,25$$

$$\Leftrightarrow P(92 < X < 108) \geq 0,75.$$

5 Lois de probabilités discrètes

5.1 Loi uniforme discrète

On dit que la v.a. X suit une loi de probabilité uniforme discrète sur $\{1, \ldots, n\}$ si

$$\forall\, 1 \leq i \leq n,\ P_X(i) = P(X = i) = \frac{1}{n}$$

On note $X \sim \mathcal{U}(\{1, \ldots, n\})$.

$$\mathbb{E}(X) = \frac{n+1}{2} \qquad\qquad \mathbf{Var}X = \frac{n^2-1}{12}$$

5.2 Loi de Bernoulli

On dit que la v.a. X suit une loi de probabilité de Bernoulli de paramètre p si :

$$P_X(1) = P(X = 1) = p \qquad \text{et} \qquad P_X(0) = P(X = 0) = 1 - p$$

On note $X \sim \mathcal{B}(1, p)$ ou bien $X \sim \mathcal{B}(p)$.

$$\mathbb{E}(X) = p \qquad\qquad \mathbf{Var}X = p(1-p)$$

5.3 Loi binomiale

On dit que la v.a. X suit une loi de probabilité binomiale de paramètres n et p si :

$$\forall\, 0 \leq k \leq n : P_X(k) = P(X = k) = C_n^k p^k (1-p)^{n-k}$$

On note $X \sim \mathcal{B}(n, p)$.

$$\mathbb{E}(X) = np \qquad\qquad \mathbf{Var}X = np(1-p)$$

Lien avec le modèle d'urne :
Soit une urne contenant une proportion p de boules blanches (et une proportion $1 - p$ de boules noires). On en tire n avec remise. On pose $\forall\, 1 \leq i \leq n$:

$$X_i = \begin{cases} 1 \text{ si la } i\text{ème boule tirée est blanche} \\ 0 \text{ sinon} \end{cases}$$

Propriété : Les v.a. $X_1, \ldots, X_n$ sont indépendantes et identiquement distribuées (= i.i.d = indépendantes et de même loi) de loi $\mathcal{B}(1, p)$.

Soit $S_n = X_1 + \cdots + X_n$. S_n est le nombre de boules blanches tirées parmi les n et :

$$P(S_n = k) = C_n^k p^k (1-p)^{n-k} \ \ \forall k \in \{1, \ldots, n\}$$

$S_n \sim \mathcal{B}(n, p)$.

Propriété : Soient X et Y deux v.a. **indépendantes** telles que $X \sim \mathcal{B}(n, p)$ et $Y \sim \mathcal{B}(m, p)$.

Alors $X + Y \sim \mathcal{B}(n + m, p)$.

5.4 Loi hypergéométrique

Modèle d'urne :

Soit une urne comportant M boules blanches et $N - M$ boules noires (soit au total N boules). On pose $p = \frac{M}{N}$ (= proportion de boules blanches au départ). On en tire n sans remise.

On pose $\forall 1 \leq i \leq n$:

$$X_i = \begin{cases} 1 \text{ si la } i\text{ème boule tirée est blanche} \\ 0 \text{ sinon} \end{cases}$$

et $S_n = X_1 + \cdots + X_n$ le nombre de boules blanches tirées. Alors :

1. On a $S_n \geq 0$ et $S_n \geq n - (N - M)$ car $n - S_n \leq N - M$, donc $S_n \geq \sup(0, n - N + M)$.
   On a $S_n \leq n$ et $S_n \leq M$ donc $S_n \leq \inf(n, M)$.
2. $\forall k, \ \sup(0, n - N + M) \leq k \leq \inf(n, M)$:

   $$P_{S_n}(k) = P(S_n = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

   On dit que S_n suit une loi hypergéométrique de paramètres N, n et $p = \frac{M}{N}$.
   On note $S_n \sim \mathcal{H}(N, n, p = \frac{M}{N})$

   $$\mathbb{E}(S_n) = np = n\frac{M}{N} \qquad\qquad \mathbf{Var}X = n\frac{M}{N}\frac{N-M}{N}\frac{N-n}{N-1}$$

   Propriété : Les v.a. $X_1, \ldots, X_n$ sont identiquement distribuées mais elles ne sont pas indépendantes.
   Remarque : $\frac{N-n}{N-1} \leq 1$ et $\frac{N-n}{N-1} < 1$ pour $n \geq 2$.
   D'où :

   $$\mathbf{Var}(\mathcal{H}(N, n, p)) \leq \mathbf{Var}(\mathcal{B}(n, p))$$

   et $\mathbf{Var}(\mathcal{H}(N, n, p)) < \mathbf{Var}(\mathcal{B}(n, p))$ dès que $n \geq 2$.

5.5 Loi géométrique

On dit que la v.a. X suit une loi de géométrique de paramètre $p \in]0,1]$ si

$$\forall\, k \geq 1,\; P_X(k) = P(X = k) = (1-p)^{k-1} p$$

On note $X \sim \mathcal{G}(p)$.

$$\mathbb{E}(X) = \frac{1}{p} \qquad\qquad \mathbf{Var} X = \frac{1-p}{p^2}$$

REMARQUE : Ici $X(\Omega) = \mathbb{N}^*$ (infini dénombrable)

Modèle d'urne :
Une urne contient une propotion $p > 0$ de boules blanches. On effectue des tirages **avec remise** jusqu'à obtenir une boule blanche. Soit X le numéro du tirage où on tire une boule blanche pour la $1^{\text{ère}}$ fois.
Alors $X \sim \mathcal{G}(p)$.

Exemple : On lance un dé équilibré jusqu'à obtenir un 6. Soit X le numéro du lancer où on obtient 6.

5.6 Loi de Poisson

On dit qu'une v.a. X à valeurs dans $\mathbb{N}$ suit une loi de Poisson de paramètre $\lambda > 0$ si :

$$\forall\, k \geq 0 : P_X(k) = P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

On note $X \sim \mathcal{P}(\lambda)$.

$$\mathbb{E}(X) = \lambda \qquad\qquad \mathbf{Var} X = \lambda$$

Cette loi sert le plus souvent à modéliser un nombre (aléatoire) d'évènements se produisant pendant un intervalle de temps donné.
Par exemple, $X =$ nombre de personnes entrant dans un magasin pendant 10 minutes.
On peut modéliser ce phénomène en supposant que $X \sim \mathcal{P}(\lambda)$ pour un certain $\lambda > 0$.

REMARQUE : $\mathbb{E}(X) = \lambda$, donc λ représente la valeur moyenne de la variable X. Dans notre exemple, λ représente le nombre moyens de clients entrant dans le magasin en 10 minutes.

Propriété : Soit $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ deux variables **indépendantes**. Alors $X + Y \sim \mathcal{P}(\lambda + \mu)$.

5.7 Loi de Pascal (ou binomiale négative)

Modèle d'urne : Soit une urne comprtant une proportion $p > 0$ de boules blanches. On effectue des tirages avec remise jusqu'à obtenir une boule blanche pour la $r^{\text{ième}}$ fois. Soit X le numéro du tirage où on tire la $r^{\text{ième}}$ boule blanche. Alors :

$$\forall n \geq r : P_X(n) = P(X = n) = C_{n-1}^{r-1} p^r (1-p)^{n-r}$$

On dit que X suit une loi de Pascal (ou binomiale négative) de paramètres r et p. On note $X \sim Pa(r, p)$

$$\mathbb{E}(X) = \frac{r}{p} \qquad\qquad \mathbf{Var}X = \frac{r(1-p)}{p^2}$$

5.8 Approximation des lois hypergéométrique et binomiale

5.8.1 Approximation de la loi hypergéométrique par la loi binomiale

THÉORÈME 1 *Soit $X_N \sim \mathcal{H}(N, n, p = \frac{M}{N})$. On suppose qu'on fait tendre N vers l'infini mais que n et p restent constants. Alors :*

$$\lim_{N \to +\infty} P(X_N = k) = C_n^k p^k (1-p)^{n-k} = P\left(\mathcal{B}(n, p) = k\right)$$

Démonstration : ...
Interprétation : Supposons que le nombre de tirages n est très petit par rapport au nombre N de boules dans l'urne. Alors, à chaque tirage, la proportion de boules blanches est pratiquement inchangée. Donc ceci revient pratiquement à effectuer des tirages **avec remise**.
En pratique, on utilise l'approximation $P(\mathcal{H}(N, n, p) = k) \simeq P(\mathcal{B}(n, p) = k)$ dès que $\frac{n}{N} \leq 0,1$.

5.8.2 Approximation de la loi binomiale par la loi de Poisson

THÉORÈME 2 *Soit $X_n \sim \mathcal{B}(n, p)$. On fait tendre n vers l'infini mais que on suppose que $\lambda = np$ reste constant (donc $p = \frac{\lambda}{n} \to 0$). Alors :*

$$\forall k \geq 1, \lim_{n \to +\infty} P(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Autrement dit : $$\lim_{n \to +\infty} C_n^k p^k (1-p)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!} = P(\mathcal{P}(\lambda) = k)$$

Démonstration :...

En pratique, on peut dire que $P(\mathcal{B}(n,p)=k) \simeq P(\mathcal{P}(\lambda)=k)$ avec $\lambda = np$ dès que : $\begin{cases} n \geq 50 \\ p \leq 0,1 \\ np \leq 20 \end{cases}$

Troisième partie

Lois de probabilités continues

Rappels sur l'intégration.
On rappelle qu'étant donnée une fonction positive f,

$$\int_a^{+\infty} f(x)dx = \lim_{b\to+\infty} \int_a^b f(x)dx,$$

$$\int_{-\infty}^b f(x)dx = \lim_{a\to-\infty} \int_a^b f(x)dx,$$

et

$$\int_{-\infty}^{+\infty} f(x)dx = \lim_{a\to-\infty, b\to+\infty} \int_a^b f(x)dx.$$

6 Variables aléatoires continues.

Définition 32. *On dit qu'une variable aléatoire réelle X sur un espace probabilisé $(\Omega, \mathcal{F}, \mathrm{P})$ est continue (ou absolument continue) s'il existe une fonction positive f telle que $\forall a < b$,*

$$\mathrm{P}_X(]a, b]) = \mathrm{P}(X \in]a, b]) = \int_a^b f(x)dx.$$

La fonction f est appelée densité de probabilité de la variable X.
Propriété 6 — Une variable aléatoire réelle X est continue (ou absolument continue) s'il existe une fonction positive f telle que $\forall t \in \mathbb{R}$,

$$\mathrm{P}_X(]-\infty, t]) = \mathrm{P}(X \in]-\infty, t]) = \mathrm{P}(X \leq t) = \int_{-\infty}^t f(x)dx.$$

On remarque que la probabilité $\mathrm{P}_X(]a, b]) = \mathrm{P}(X \in]a, b])$ est donnée par la valeur de l'aire délimitée par l'axe Ox en bas, la courbe de la fonction $f(x)$ en haut, et les droites verticales $x = a$ et $x = b$ à gauche et à droite.
Propriété 7 — Pour tout $x \in \mathbb{R}$, $\mathrm{P}(X = x) = 0$. En effet

$$\mathrm{P}(X = x) = \int_x^x f(t)dt = 0.$$

Il en résulte que $\mathrm{P}(X \leq t) = \mathrm{P}(X < t)$ et $\mathrm{P}(X \geq t) = \mathrm{P}(X > t)$.
Propriété 8 $-$ On a

$$\int_{-\infty}^{+\infty} f(x)dx = \mathrm{P}(X \in]-\infty, +\infty[) = 1.$$

Premier exemple
Définition 33. *On dit qu'une variable aléatoire X suit une loi uniforme sur l'intervalle $[0,1]$ si elle est (absolument) continue de densité donnée par la fonction*

$$f(x) = \mathbb{1}_{[0,1]}(x).$$

On note $X \sim \mathcal{U}([0,1])$.

6.1 Fonction de répartition.

Définition 34. *Soit une variable aléatoire réelle X absolument continue de densité f. On appelle fonction de répartition de X (ou de la loi de X) la fonction notée F_X et définie $\forall t \in \mathbb{R}$ par*

$$F(t) = F_X(t) = \mathrm{P}(X \leq t) = \int_{-\infty}^{t} f(x)dx.$$

Propriété 9 $-$ Si X est absolument continue de densité f, alors sa fonction de répartition F est continue et $F' = f$.

6.2 Espérance, moments, variance d'une variable aléatoire continue.

Définition 35. *Soit une variable aléatoire réelle X absolument continue de densité f. On dit que X est d'espérance finie si*

$$\int_{-\infty}^{+\infty} |x| f(x)dx < +\infty,$$

et on appelle alors espérance de X le nombre

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} x f(x)dx.$$

Exemple : Soit X une variable aléatoire uniforme sur $[0,1]$. L'espérance de X vaut

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^{+\infty} x \mathbb{1}_{[0,1]}(x) dx \\ &= \int_0^1 x dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}. \end{aligned}$$

Théorème du transfert.

Théorème 4 - *Soit une variable aléatoire réelle X absolument continue de densité f d'espérance finie, et soit g une fonction de $\mathbb{R}$ dans $\mathbb{R}$, telle que $g(X)$ soit d'espérance finie. Alors*

$$\mathrm{E}(g(X)) = \int_{-\infty}^{+\infty} g(x) f(x) dx.$$

Définition 36. *Soit une variable aléatoire réelle X absolument continue de densité f. On dit que X admet un moment non centré d'ordre k si*

$$\int_{-\infty}^{+\infty} |x^k| f(x) dx < +\infty,$$

et on appelle alors moment non centré d'ordre k de X le nombre

$$\mathrm{E}(X^k) = \int_{-\infty}^{+\infty} x^k f(x) dx.$$

Définition 37. *Soit une variable aléatoire réelle X absolument continue de densité f. On dit que X admet un moment centré d'ordre k si*

$$\int_{-\infty}^{+\infty} |(x - \mathrm{E}(X))^k| f(x) dx < +\infty,$$

et on appelle alors moment centré d'ordre k de X le nombre

$$\mathrm{E}(X - \mathrm{E}(X))^k = \int_{-\infty}^{+\infty} (x - \mathrm{E}(X))^k f(x) dx.$$

Définition 38. *On appelle variance de X le moment centré d'ordre $k = 2$, noté σ_X^2 ou* $\mathrm{Var}(X)$.

Propriété 10 — On a la formule de Koenig de calcul de la variance :

$$\sigma_X^2 = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = \int_{-\infty}^{+\infty} x^2 f(x)dx - \left(\int_{-\infty}^{+\infty} xf(x)dx\right)^2.$$

Exemple : Soit X une variable aléatoire uniforme sur $[0, 1]$. La variance de X vaut

$$\sigma_X^2 = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = \int_{-\infty}^{+\infty} x^2 f(x)dx - \left(\int_{-\infty}^{+\infty} xf(x)dx\right)^2$$

$$= \left[\frac{x^3}{3}\right]_0^1 - \frac{1}{4} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

7 La loi normale ou loi de Gauss.

7.1 Définition et propriétés.

Définition 39. *On dit qu'une variable aléatoire X suit une loi normale centrée réduite si elle est (absolument) continue de densité donnée par la fonction*

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

On note $X \sim \mathcal{N}(0, 1)$.

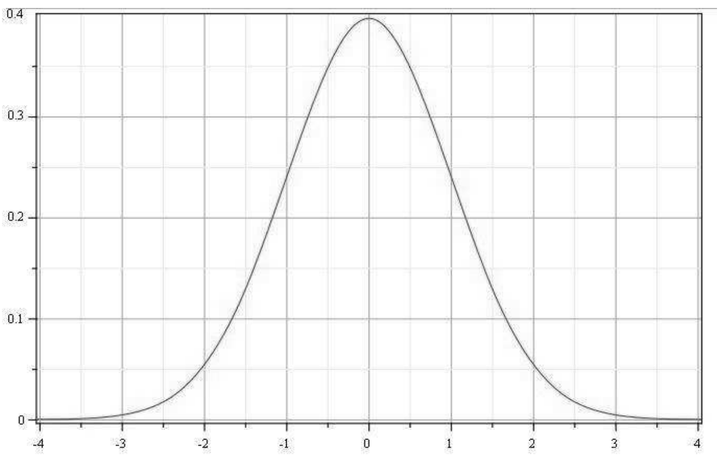
Représentation graphique de la densité de la loi normale centrée réduite

Propriétés de symétrie On voit graphiquement que

$$\forall x \in \mathbb{R},\ \mathrm{P}(X \geq x) = \mathrm{P}(X \leq -x),\ \mathrm{P}(X \leq x) = \mathrm{P}(X \geq -x),$$

ou encore que

$$\forall x > 0,\ \mathrm{P}(0 \leq X \leq x) = \mathrm{P}(-x \leq X \leq 0).$$



Définition 40. *On dit qu'une variable aléatoire X suit une loi normale de moyenne μ et de variance σ^2 si la variable aléatoire $\frac{X-\mu}{\sigma}$ suit une loi normale centrée réduite $\mathcal{N}(0,1)$. On note $X \sim \mathcal{N}(\mu, \sigma^2)$.*

Propriété 11 $-$ Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Alors la densité f de X vaut

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

Remarque 3.

— On ne peut déterminer de primitive de la fonction $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, c'est pourquoi on ne peut déterminer d'expression fonctionnelle de la fonction de répartition

$$F(x) = \mathrm{P}(\mathcal{N}(0,1) \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Pour calculer les valeurs de $\mathrm{P}(\mathcal{N}(0,1) \leq x)$, on aura recours à des tables.

— Dans le cas où $X \sim \mathcal{N}(\mu, \sigma^2)$, on peut toujours se ramener à la table de la loi $\mathcal{N}(0,1)$, puisque

$$\mathrm{P}(X \leq z) = \mathrm{P}\left(\frac{X-\mu}{\sigma} \leq \frac{z-\mu}{\sigma}\right) = \mathrm{P}(\mathcal{N}(0,1) \leq x),$$

avec $x = \frac{z-\mu}{\sigma}$.

Espérance et variance d'une variable aléatoire normale.

$$\mathrm{E}(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \left[-e^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} = 0 - 0 = 0.$$

$$\mathrm{E}(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x x e^{-\frac{x^2}{2}} dx.$$

Intégration par parties $u(x) = x \Rightarrow u'(x) = 1$ et $v'(x) = xe^{-\frac{x^2}{2}} \Rightarrow v(x) = -e^{-\frac{x^2}{2}}$, d'où

$$\mathrm{E}(X^2) = \frac{1}{\sqrt{2\pi}} \left\{ \left[-xe^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} -e^{-\frac{x^2}{2}} dx \right\}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 1.$$

D'où finalement

$$\mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = 1 - 0^2 = 1.$$

Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Alors

$$U = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1) \Leftrightarrow X = \sigma U + \mu.$$

D'où
$$\mathrm{E}(X) = \sigma \mathrm{E}(U) + \mu = \sigma.0 + \mu = \mu,$$
et
$$\mathrm{Var}(X) = \sigma^2 \mathrm{Var}(U) = \sigma^2.1 = \sigma^2.$$
Donc μ est bien l'espérance (la moyenne) et σ^2 la variance d'une loi $X \sim \mathcal{N}(\mu, \sigma^2)$.

Propriété 12 – Si $X \sim \mathcal{N}(0,1)$, alors $-X \sim \mathcal{N}(0,1)$. **Preuve :** $\mathrm{P}(-X \leq x) = \mathrm{P}(X \geq -x) = \mathrm{P}(X \leq x)$ par symétrie.
Propriété 13 – Soit $a \neq 0,\ b \in \mathbb{R}$, et $X \sim \mathcal{N}(\mu, \sigma^2)$. Alors
$$aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2).$$

Preuve : On peut écrire
$$\frac{aX + b - (a\mu + b)}{|a|\sigma} = \frac{a(X-\mu)}{|a|\sigma} = +/- \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1).$$

Somme de deux variables aléatoires normales indépendantes
Théorème 5 – *Soit* $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ deux variables aléatoires **indépendantes**. Alors
$$X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Propriété 14 – Soit $X_1, \ldots, X_n$ n variables aléatoires indépendantes de loi $\mathcal{N}(\mu, \sigma^2)$. Posons :
$$S_n = X_1 + \cdots + X_n \text{ et } \bar{X}_n = \frac{S_n}{n}.$$
Alors
$$S_n \sim \mathcal{N}(n\mu, n\sigma^2) \text{ et } \bar{X}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n}).$$

7.2 Théorème de la limite centrale.

Remarque 4. On dit aussi "Théorème Central Limite" ou en abrégé T.C.L.
Théorème 6 – *Soit* $X_1, \ldots, X_n, \ldots$ une suite de variables aléatoires indépendantes de loi quelconque d'espérance μ et de variance σ^2. Posons $\forall n \geq 1$:
$$S_n = X_1 + \cdots + X_n \text{ et } \bar{X}_n = \frac{S_n}{n}.$$

Alors $\forall x \in \mathbb{R}$,

$$\lim_{n\to+\infty} \mathrm{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) = \mathrm{P}(\mathcal{N}(0,1) \le x),$$

et

$$\lim_{n\to+\infty} \mathrm{P}\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \le x\right) = \mathrm{P}(\mathcal{N}(0,1) \le x).$$

Application : approximation des lois binomiales et de Poisson par la loi normale.
Théorème 7 $-$ *O*n peut écrire $\forall x \in \mathbb{R}$,

$$(1) \quad \lim_{n\to+\infty} \mathrm{P}\left(\frac{\mathcal{B}(n,p) - np}{\sqrt{np(1-p)}} \le x\right) = \mathrm{P}(\mathcal{N}(0,1) \le x),$$

et

$$(2) \quad \lim_{n\to+\infty} \mathrm{P}\left(\frac{\mathcal{P}(n\lambda) - n\lambda}{\sqrt{n\lambda}} \le x\right) = \mathrm{P}(\mathcal{N}(0,1) \le x).$$

Démonstration : il suffit de préciser dans l'énoncé du T.C.L. :
$(1) \quad \forall i,\ X_i \sim \mathcal{B}(1,p) \Rightarrow \mu = p$ et $\sigma^2 = p(1-p)$, et on sait que $S_n \sim \mathcal{B}(n,p)$.
$(2) \quad \forall i,\ X_i \sim \mathcal{P}(\lambda) \Rightarrow \mu = \lambda$ et $\sigma^2 = \lambda$, et on sait que $S_n \sim \mathcal{P}(n\lambda)$.
En pratique, on considère que l'estimation est raisonnable dès que $n \ge 30$. Pour la loi binomiale, l'approximation est d'autant meilleure que p est proche de $1/2$.

8 Convergence de variables aléatoires

8.1 Convergence en Probabilité

Définition. Soit $(X_n)_{n\ge1}$ une suite de variables aléatoires. On dit que la suite $(X_n)_{n\ge1}$ tend (ou converge) en probabilité vers la variable aléatoire X si

$$\forall \varepsilon > 0,\ \mathrm{P}(|X_n - X| \ge \varepsilon) \xrightarrow[n\to+\infty]{} 0.$$

On note

$$X_n \xrightarrow[n\to+\infty]{P} X$$

Théorème. Soit $(X_n)_{n\ge1}$ une suite de variables aléatoires indépendantes et identiquement distribuées, d'espérance $\mathrm{E}[X] = \mu$ et de variance $\mathrm{Var}[X] = \sigma_X^2 = \sigma^2$. Soit $\forall n \ge 1$,

$$S_n = \sum_{i=1}^{n} X_i \text{ et } \bar{X}_n = \frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Alors

$$\bar{X}_n \xrightarrow[n\to+\infty]{P} \mu$$

Démonstration. Comme les v.a. X_i sont i.i.d., on a

$$\mathrm{E}[S_n] = \sum_{i=1}^{n} \mathrm{E}[X_i] = n\mathrm{E}[X] = n\mu,$$

et

$$\mathrm{Var}[S_n] = \sum_{i=1}^{n} \mathrm{Var}[X_i] = n\mathrm{Var}[X] = n\sigma^2.$$

D'où

$$\mathrm{E}[\bar{X}_n] = \frac{1}{n}\mathrm{E}[S_n] = \mu,$$

et

$$\mathrm{Var}[\bar{X}_n] = \frac{1}{n^2}\mathrm{Var}[S_n] = \frac{\sigma^2}{n}.$$

Donc d'après l'inégalité de Bienaymé-Tchebytcheff,

$$\forall \varepsilon > 0,\ \mathrm{P}(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0.$$

Donc

$$\bar{X}_n \xrightarrow[n\to+\infty]{P} \mu.$$

8.2 Convergence en loi

Définition. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires de fonctions de répartition $F_{X_n}(x) = \mathrm{P}(X_n \leq x)$ et X une v.a. de fonction de répartition $F_X(x) = \mathrm{P}(X \leq x)$. On dit que la suite $(X_n)_{n\geq 1}$ tend (ou converge) en loi vers la variable aléatoire X si $\forall x$ tel que F_X est continue au point x,

$$F_{X_n}(x) \xrightarrow[n\to+\infty]{} F_X(x) \Leftrightarrow \mathrm{P}(X_n \leq x) \xrightarrow[n\to+\infty]{} \mathrm{P}(X \leq x).$$

On note

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X.$$

Théorème Central Limite (nouvelle formulation). Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées, d'espérance $\mathrm{E}[X] = \mu$ et de variance $\mathrm{Var}[X] = \sigma_X^2 = \sigma^2$. Soit $\forall n \geq 1$,

$$S_n = \sum_{i=1}^{n} X_i \text{ et } \bar{X}_n = \frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Alors
$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1),$$
ou de manière équivalente
$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

Théorème.
$$X_n \xrightarrow[n\to+\infty]{P} X \Rightarrow X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X.$$

9 Correction de continuité.

Théorème Central Limite. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées, d'espérance $\mathrm{E}[X] = \mu$ et de variance $\mathrm{Var}[X] = \sigma_X^2 = \sigma^2$. Soit $\forall n \geq 1$,

$$S_n = \sum_{i=1}^{n} X_i \text{ et } \bar{X}_n = \frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Alors

$$\mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\mathcal{N}(0,1) \leq x\right),$$

ou de manière équivalente

$$\mathbb{P}\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq x\right) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\mathcal{N}(0,1) \leq x\right).$$

On peut aussi dire que pour n suffisamment grand ($n \geq 30$ en général),

$$\mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) \simeq \mathbb{P}\left(\mathcal{N}(0,1) \leq x\right),$$

ou de manière équivalente

$$\mathbb{P}\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq x\right) \simeq \mathbb{P}\left(\mathcal{N}(0,1) \leq x\right).$$

Application à la loi binômiale. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi de Bernoulli de paramètre p. Alors $\mathrm{E}[X] = p$ et $\mathrm{Var}[X] = \sigma_X^2 = p(1-p)$. De plus S_n suit une loi binômiale $\mathcal{B}(n,p)$. Le T.C.L. s'écrit donc dans ce cas $\forall x \in \mathbb{R}$, pour n suffisamment grand,

$$\mathbb{P}\left(\frac{\mathcal{B}(n,p) - np}{\sqrt{np(1-p)}} \leq x\right) \simeq \mathbb{P}\left(\mathcal{N}(0,1) \leq x\right).$$

Supposons qu'on cherche plutôt une approximation de

$$\mathbb{P}\left(\mathcal{B}(n,p) \leq k\right),\ k \in \mathbb{N}.$$

On peut donc écrire

$$\begin{aligned}
\mathbb{P}\left(\mathcal{B}(n,p) \leq k\right) &= \mathbb{P}\left(\frac{\mathcal{B}(n,p) - np}{\sqrt{np(1-p)}} \leq \frac{k-np}{\sqrt{np(1-p)}}\right) \\
&\simeq \mathbb{P}\left(\mathcal{N}(0,1) \leq \frac{k-np}{\sqrt{np(1-p)}}\right) \\
&\simeq \mathbb{P}\left(\mathcal{N}\left(np, np(1-p)\right) \leq k\right).
\end{aligned}$$

On peut donc dire qu'**on approche la loi** $\mathcal{B}(n, p)$ **par la loi** $\mathcal{N}(np, np(1-p))$**.**

Correction de continuité. Dans ce qui précède, on obtient une meilleure approximation de $\mathbb{P}(\mathcal{B}(n,p) \leq k)$ par la formule suivante

$$\mathbb{P}(\mathcal{B}(n,p) \leq k) \simeq \mathbb{P}(\mathcal{N}(np, np(1-p)) \leq k + 0,5).$$

En effet la loi $\mathcal{B}(n, p)$ est la loi d'une variable à valeurs entières que nous approchons par la loi d'une variable à valeurs réelles. Il faut donc voir en réalité la variable $\mathcal{B}(n, p)$ que nous approchons comme **l'arrondie à l'entier près** d'une variable de loi $\mathcal{N}(np, np(1-p))$. On a donc l'équivalence

$$\mathcal{B}(n,p) \leq k \Leftrightarrow \mathcal{N}(np, np(1-p)) \leq k + 0,5.$$

De la même façon, on utilisera de préférence les équivalences suivantes :

$$\begin{aligned}
\mathcal{B}(n,p) \geq k &\quad\Leftrightarrow\quad \mathcal{N}(np, np(1-p)) \geq k - 0,5,\\
\mathcal{B}(n,p) < k &\quad\Leftrightarrow\quad \mathcal{N}(np, np(1-p)) \leq k - 0,5,\\
\mathcal{B}(n,p) > k &\quad\Leftrightarrow\quad \mathcal{N}(np, np(1-p)) \geq k + 0,5.
\end{aligned}$$

Remarque 5. L'utilisation de la correction de continuité :

- permet de garder une cohérence par rapport au passage au complémentaire. En effet

$$\begin{array}{rcl}
\mathbb{P}(\mathcal{B}(n,p) \leq k) & \simeq & \mathbb{P}(\mathcal{N}(np, np(1-p)) \leq k + 0,5),\\
& \Longleftrightarrow & \\
1 - \mathbb{P}(\mathcal{B}(n,p) \leq k) & \simeq & 1 - \mathbb{P}(\mathcal{N}(np, np(1-p)) \leq k + 0,5),\\
& \Longleftrightarrow & \\
\mathbb{P}(\mathcal{B}(n,p) > k) & \simeq & \mathbb{P}(\mathcal{N}(np, np(1-p)) \geq k + 0,5).
\end{array}$$

- permet de définir une approximation assez précise de la probabilité ponctuelle $\mathbb{P}(\mathcal{B}(n, p) = k)$. On peut en effet écrire que

$$\begin{aligned}
\mathbb{P}(\mathcal{B}(n,p) = k) &= \mathbb{P}(\mathcal{B}(n,p) \leq k) - \mathbb{P}(\mathcal{B}(n,p) < k)\\
&\simeq \mathbb{P}(\mathcal{N}(np, np(1-p)) \leq k + 0,5) - \mathbb{P}(\mathcal{N}(np, np(1-p)) \leq k - 0,5).
\end{aligned}$$

Remarque 6. La correction de continuité s'applique d'une manière plus générale dans tous les cas où on approche une loi à valeurs entières par une loi continue (en particulier par une loi normale). Si X est discrète à valeurs dans $\mathbb{N}$ et si sa loi est approchée par une loi $\mathcal{N}(\mu, \sigma^2)$, alors

$$\mathbb{P}(X \leq k) \simeq \mathbb{P}(\mathcal{N}(\mu, \sigma^2) \leq k + 0,5).$$