Analyse et Probabilité II
chapitre 2 : Combinatoire et Probabilité

Paragraphe 1 : Analyse combinatoire, introduction

On présente dans ce paragraphe quelques méthodes permettant de déterminer sans compter directement, c'est à dire sans un dénombrement direct, le nombre de résultats possibles lors d'une expérience particulière. Pour cela on part du principe suivant :
Supposons qu'une première expérience peut être représentée de $n_1 \in \mathbb{N}$ façon différentes, suivi d'une seconde expérience représentée par $n_2 \in \mathbb{N}$ façon différentes, suivi d'une troisième expérience représentée par $n_3 \in \mathbb{N}$ façon, ... , jusqu'à la p-ième expérience représentée par $n_p \in \mathbb{N}$ façon différentes. Alors le nombre total de façons différentes est le produit $n_1 \times n_2 \times \cdots \times n_p$.

Exemple : Une plaque d'immatriculation est définie par la donnée de deux lettres suivis de trois chiffres puis encore de deux lettres. Il n'y a aucune condition sur le premier groupe de deux lettres, les trois chiffres doivent être tous distincts, et le second groupe de deux lettres doivent être distinct. Le nombre total de possibilitées est de $26 \times 26 \times 10 \times 9 \times 8 \times 26 \times 25$.

Soit un entier naturel $n \in \mathbb{N}$, on définit l'entier noté $n!$ est appelé factorielle n comme le produit des entiers compris entre 1 et n. Si $n = 0$ par convention on pose $0! = 1$.

Paragraphe 2 : Permutations, arrangements, combinaisons

On se donne un ensemble Ω de n objets (deux à deux distincts).

— Soit une entier $r \leq n$, on appelle arrangements sans répétition de r objets prisent parmi n une suite ordonnée de r objets, c'est à dire on prend à la suite r objets de Ω sans remise. L'ordre dont les objets sont tirés est prise en compte. On dit aussi un arrangement d'indice r dans un ensemble de cardinal n. Le nombre total d'arrangement d'indice r est donnée par A_n^r où :

$$A_n^r = n(n-1)(n-2)\cdots(n-(r-1)) = \frac{n!}{(n-r)!}$$

Quelquefois on utilise la dénomination d'échantillon exhaustif de taille r.

— On appelle permutations de n objets toutes suites ordonnées de n objets c'est à dire un arrangement sans répétition de n objets prisent parmi n. Le nombre de permutations est donc $A_n^n = n!$.

— Soit une entier $r \leq n$, on appelle arrangement avec répétitions de r objets prisent parmi n une suite ordonnée de r objets, c'est à dire on prend à la suite r objets de Ω avec remise, il existe donc n manière de tirer le premier objet, puis n manière de tirer le second puisqu'il y a remise ... . L'ordre dont les objets sont tirés est prise en compte. Le nombre total de arrangements avec répétition de r objets est donnée par n^r. On dit aussi que c'est un échantillon non exhausttif de taille r.

— Soit une entier $r \leq n$, on appelle combinaisons ou combinaisons sans remise de r objet pris parmi n un sous ensemble de r objets de Ω, autrement dit on tire r objets de Ω sans remise mais l'ordre des objets n'est pas prise en compte. On dit aussi une combinaison d'indice r dans un ensemble de cardinal n. Le nombre total de combinaisons d'indice r est donnée par C_n^r où :

$$C_n^r = \frac{A_n^r}{r!} = \frac{n!}{r!(n-r)!}$$

On suppose que dans l'ensemble Ω les objets ne sont pas deux à deux distincts, mais qu'il existe un premier sous ensemble formé de n_1 objets identiques, un second sous ensemble formé de n_2 objets identiques, ... , jusqu'au dernier sous ensemble formé de n_k objets identiques. Donc en particulier $n_1 + n_2 + \cdots + n_k = n$, on tire n objets sans remise soit un arrangement d'indice n, le nombre total d'arrangements d'indice n est donnée par

$$\frac{n!}{n_1!n_2!...n_k!}$$

En effet si tous les éléments de tous les k sous ensembles sont deux à deux distincts, alors il y aura exactement $n!$ arrangements possibles, mais le premier sous ensemble contient exactement n_1 éléments qui formeront $n_1!$ arrangements possibles, étant tous identiques, ils seront indiscernables, il faut donc diviser par $n_1!$ le nombre $n!$. La raisonnement se poursuit avec les sous ensembles suivants. Attention la condition que Ω possè de n éléments est importante.

Exemple On veut écrire tous les mots de 5 caractères possibles à partir du mot $LILLE$. Si on écrit $LILLE$ sous la forme $L_1IL_2L_3E$ alors il y aura exactement 5! arrangements possibles. On remarque maintenant que les 6 arrangements :

$$L_1L_2L_3IE\ ;\quad L_2L_1L_3IE\ ;\quad L_3L_1L_2IE\ ;\quad L_1L_3L_2IE\ ;\quad L_2L_3L_1IE\ ;\quad L_3L_2L_1IE$$

forment le même mot si on supprime les indices. Les 6 arrangements ou permutations venant des 6 façons différentes de placer les trois L. Par conséquence il y a

$$\frac{5!}{3!} = 20$$

mots différents de 5 lettres.

Paragraphe 3 : Tribu

Soit un ensemble Ω donné, dans la théorie des probabilités il est appelé ensemble ou espace fondamentale, ou encore univers. Les éléments de Ω sont appelés les épreuves. On considère $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω. Un sous ensemble $\mathcal{A}$ de $\mathcal{P}(\Omega)$ est une tribu si

- $\Omega \in \mathcal{A}$,
- $\mathcal{A}$ est stable par passage au complémentaire, c'est à dire si $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$ où $\overline{A}$ désigne le complémentaire de A.
- $\mathcal{A}$ est stable par réunion dénombrable, c'est à dire pour une famille dénombrable $A_i \in \mathcal{A}$ on a $\cup A_i \in \mathcal{A}$

Remarque : Les éléments de $\mathcal{A}$ sont appelés des événements, un événement est donc un sous ensemble de Ω. Par passage au complémentaire, l'ensemble vide $\emptyset$ appartient forcément à toute tribu.

Soit une tribu $\mathcal{A}$ donnée, on dit que

- l'ensemble vide $\emptyset$ est l'événement impossible,
- l'ensemble Ω est l'événement certain,
- pour deux événements A et B, $A \cup B$ désigne l'événement A ou B, tandis que $A \cap B$ désigne l'événement A et B.
- le complémentaire d'un événement A noté $\overline{A}$ est l'événement contraire.

Exemple : On lance un dé, on note Ω l'ensemble de tous les résultats ou épreuves possible, donc $\Omega = \{1, 2, 3, 4, 5, 6\}$, l'ensemble $\mathcal{P}(\Omega)$ est une tribu. L' événement :

— le résultat est un nombre pair correspond au sous ensemble $A = \{2, 4, 6\}$ qui est un élément de $\mathcal{P}(\Omega)$.
— le résultat est un nombre impair correspond au sous ensemble $B = \{1, 3, 5\}$,
— le résultat est un nombre premier correspond au sous ensemble $C = \{2, 3, 5\}$,
— l'événement $A \cup C$ correspond à l'apparition d'un nombre pair ou premier,
— l'événement $B \cap C$ correspond à l'apparition d'un nombre impair premier,
— l'événement $\overline{C}$ correspond à l'apparition d'un nombre non premier.

Exemple : On lance une pièce de monnaie trois fois de suite et on note la suite de piles (P) ou de face (F), l'univers de tous les résultats possibles est $\Omega = \{FFF, FFP, FPF, FPP, PFF, PPF, PFP, PPP\}$. Soit A l'événement correspondant à l'apparition de deux faces ou plus consécutives, alors $A = \{FFF, FFP, PFF\}$. Soit B l'événement trois faces identiques $B = \{FFF, PPP\}$. L'événement $A \cap B$ correspond à trois faces successives.

Exemple : On lance une paire de dés, l'ensemble de tous les résultats possibles est alors

$$\Omega = \{\ (i, j)\ /\ i \in \{1, 2, ..., 6\} \text{ et } j \in \{1, 2, ..., 6\}\}$$

L'événement la somme des deux résultats sortis est un nombre premier supérieur à 6 correspond à

$$A = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (5, 6), (6, 1), (6, 5)\}$$

Si $\mathcal{A}$ est une tribu, alors le couple $(\Omega, \mathcal{A})$ s'appelle un espace probabilisable. Dans le cas où Ω est un ensemble fini, on prendra pour $\mathcal{A}$ toujours $\mathcal{P}(\Omega)$.

Proposition 3.1
On se donne une application f de $\mathbb{E}$ dans $\mathbb{F}$ et une tribu $\mathcal{F}$ sur $\mathbb{F}$. Alors l'image réciproque $f^{-1}(\mathcal{F})$ est une tribu sur $\mathbb{E}$.

Preuve :
On rappelle que

$$f^{-1}(\mathcal{F}) = \{\ f^{-1}(B)\ \ /\ \ B \in \mathcal{F}\ \} = \{\ A \subset \mathbb{E}\ \ /\ \ f(A) \in \mathcal{F}\ \}$$

Montrons que $\mathcal{E} = f^{-1}(\mathcal{F})$ est une tribu. Il est clair que $\mathbb{F}$ appartient à $\mathcal{F}$, donc $\mathbb{E} = f^{-1}(\mathbb{F}) \in \mathcal{E}$. Vérifions la stabilité par passage au complémentaire, soit $A = f^{-1}(B) \in \mathcal{E}$. On a donc $A = \{\ x \in \mathbb{E}\ \ /\ \ f(x) \in B\ \}$, soit $\overline{A} = \{\ x \in \mathbb{E}\ \ /\ \ f(x) \notin B\ \} = \{\ x \in \mathbb{E}\ \ /\ \ f(x) \in \overline{B}\ \} = f^{-1}(\overline{B})$ qui est le résultat voulu. Reste à vérifier la stabilité par réunion dénombrable. Soit une famille dénombrable d'éléments A_n de $\mathcal{E}$, on a donc $A_n = f^{-1}(B_n)$ pour $B_n \in \mathcal{F}$. Or

$$\cup A_n = \cup f^{-1}(B_n) = f^{-1}(\cup B_n)$$

et $\cup B_n \in \mathcal{F}$, donc $f^{-1}(\cup B_n) \in \mathcal{E}$ c'est à dire $\cup A_n \in \mathcal{E}$.

Proposition 3.2
Soit une espace probabilisable $(\Omega, \mathcal{A})$ donné. On se donne un sous ensemble Ω' de Ω. Alors l'ensemble $\{\ A \cap \Omega'\ \ /\ \ A \in \mathcal{A}\ \}$ est une tribu sur Ω' appelé tribu trace de $\mathcal{A}$ sur Ω'.

Preuve :
Notons $\mathcal{A}' = \{\ A \cap \Omega'\ \ /\ \ A \in \mathcal{A}\ \}$. Comme $\Omega \in \mathcal{A}$ et $\Omega' = \Omega \cap \Omega'$, on en déduit $\Omega' \in \mathcal{A}'$. Soit $A' \in \mathcal{A}'$, mais $\Omega' - A' = (\Omega - A) \cap \Omega' = \overline{A} \cap \Omega'$, et donc $\Omega' - A' \in \mathcal{A}'$. Pour le dernier point, soit A'_n une famille dénonbrable d'événements de $\mathcal{A}'$, alors $A'_n = A_n \cap \Omega$ et donc

$$\cup A'_n = \cup(A_n \cap \Omega') = (\cup A_n) \cap \Omega'$$

mais $\cup A_n \in \mathcal{A}$ d'ou le résultat.

Proposition 3.3

Soit une famille dénombrable d'espaces probabilisables $(\Omega, \mathcal{A}_i)$ définie sur le même univers Ω. Alors $(\Omega, \cap\mathcal{A}_i)$ est un espace probabilisable.

Preuve :
Notons $\mathcal{A} = \cap\mathcal{A}_i$. Comme $\Omega \in \mathcal{A}_i$ pour tout i, on a $\Omega \in \cap\mathcal{A}_i$ c'est à dire $\Omega \in \mathcal{A}$. Soit $A \in \mathcal{A}$, donc $A \in \mathcal{A}_i$ pour tout i, d'ou $\overline{A} \in \mathcal{A}_i$ pour tout i soit $\overline{A} \in \mathcal{A}$. Enfin pour le dernier point soit B_j une famille dénombrable d'événements de $\mathcal{A}$, chaque B_j appartient à tout $\mathcal{A}_i$ pour tout i, donc $B_j \in \cap\mathcal{A}_i$, ceci étant vrai pour tout j, on a $\cup B_j \in \cap A_i$. D'ou le résultat.

Proposition 3.4

Soit un ensemble dénombrable Ω, on se donne un sous ensemble $\mathcal{F}$ de $\mathcal{P}(\Omega)$, alors il existe une plus petite tribu au sens de l'inclusion contenant $\mathcal{F}$. On l'appelle la tribu engendrée par $\mathcal{F}$.

Preuve :
Comme $\mathcal{P}(\Omega)$ est une tribu contenant $\mathcal{F}$, l'ensemble des tribus contenant $\mathcal{F}$ n'est pas vide, il suffit alors d'appliquer la proposition précédente.

Paragraphe 4 : Mesure et probabilité

On se donne un espace probabilisable $(\Omega, \mathcal{A})$. On appelle probabilité une fonction définie de $\mathcal{A}$ à valeurs dans $[0, 1]$ souvent noté μ

$$\mu \ : \ \mathcal{A} \longrightarrow \ [0,1] \ / \ A \ \longrightarrow \ \mu(A)$$

tel que
- $\mu(\Omega) = 1$.
- Pour tout suite d'éléments A_i de $\mathcal{A}$ deux à deux disjoints on a $\mu(\cup A_i) = \sum \mu(A_i)$.

Le triplet $(\Omega, \mathcal{A}, \mu)$ s'appelle un espace probabilisé.

Proposition 4.1

Soit un espace probabilisé $(\Omega, \mathcal{A}, \mu)$. On a alors les propriétés suivantes
- $\mu(\emptyset) = 0$,
- $\mu(\overline{A}) = 1 - \mu(A)$,
- $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$,
- Si $A \subset B$ alors $\mu(A) \leq \mu(B)$.

où A et B sont deux éléments de $\mathcal{A}$.

Preuve :
On a $\Omega \cap \emptyset = \emptyset$ et $\Omega \cup \emptyset = \Omega$, donc $\mu(\Omega) = \mu(\Omega \cup \emptyset) = \mu(\Omega) + \mu(\emptyset)$ d'ou $\mu(\emptyset) = 0$. Pour tout élément A de $\mathcal{A}$ on a par définition $\Omega = A \cup \overline{A}$ et $A \cap \overline{A} = \emptyset$, d'ou $\mu(\overline{A}) = 1 - \mu(A)$. Quand au troisième point on remarque $A \cup B = A \cup (B \setminus A \cap B)$, $A \cap (B \setminus A \cap B) = \emptyset$, $B = (B \setminus A \cap B) \cup (A \cap B)$ et que $(B \setminus A \cap B) \cap (A \cap B) = \emptyset$. On a alors le résultat. Maintenant si $A \subset B$ alors $B = A \cup B \setminus A$, $A \cap B - A = \emptyset$ d'ou $\mu(A) \leq \mu(B)$.

Exemple : On jette en l'air trois pièces de monnais et on note le nombre de faces obtenu. L'univers de tous les résultats possibles est $\Omega = \{0, 1, 2, 3\}$ et la tribu $\mathcal{A} = \mathcal{P}(\Omega)$. On définit une probabilité μ

sur $\mathcal{P}(\omega)$ en posant $\mu(\{0\}) = 1/8$, $\mu(\{1\}) = \mu(\{2\}) = 3/8$, et $\mu(\{3\}) = 1/8$. Si on considère les deux événements A obtenir au moins deux faces et B obtenir trois faces identiques, alors $A = \{2, 3\}$ et $B = \{0, 3\}$ et on a $\mu(A) = \mu(\{2\}) + \mu(\{3\}) = 4/8$ et $\mu(B) = \mu(\{0\}) + \mu(\{3\}) = 2/8$

Paragraphe 5 : Équiprobabilité

Soient un ensemble Ω de cardinal fini et la tribu $\mathcal{P}(\Omega)$, on dit que l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ est muni de la probabilité uniforme μ ou équiprobabilité si pour tout singleton $\{\omega\} \in \mathcal{P}(\Omega)$ on pose $\mu(\{\omega\}) = 1/\sharp(\Omega)$ où $\sharp(\Omega)$ est le cardinal de Ω. Tout événement $A \in \mathcal{P}(\Omega)$ est un sous ensemble de Ω, alors sa probabilité est de $\sharp(A)/\sharp(\Omega)$.
L'expression choisir au hasard un point dans un ensemble Ω signifie que l'on considère l'espace probablilisé $(\Omega, \mathcal{P}(\omega), \mu)$ où μ est la probabilité uniforme.

Exemple : Dans un jeu de 52 cartes, l'ensemble Ω est l'ensemble de toutes les cartes. On tire une carte au hasard et on veut connaitre la probabilité que ce soit un pique. On considère l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mu)$ muni de la probabilité uniforme. L'événement A est un pique contient 13 cartes, donc la probabilité de tirer un pique est de 13/52. L'événement B est une tête (un valet, une reine, un roi) est constitué de 12 cartes, donc la probabilité de tirer une tête est de 12/52. Enfin l'événement $A \cap B$ correspond tirer une pique représentant une tête, il n'y a que 3 cartes vérifiant ces conditions, donc la probabilité est de 3/52.

Exemple : On dispose d'un lot de 12 articles dont 4 sont défecteux, et on tire au hasard deux articles. L'univers Ω est donc l'ensemble formé par toutes les combinaisons de 2 articles prisent parmi 12, la tribu $\mathcal{P}(\Omega)$. Soient les deux événements $A = \{$ les deux articles sont défectueux $\}$ et $B = \{$ aucun articles n'est défectueux$\}$. Pour calculer $\mu(A)$ et $\mu(B)$, on remarque que

— il y a $C_{12}^2 = 66$ manières différentes de choisir 2 articles parmi 12, qui est donc le cardinal de Ω.
— il y a $C_4^2 = 6$ manières différentes de choisir 2 articles défectueux parmi 4, soit le nombre de combinaisons de 2 articles défectueux parmi 4.
— il y a $C_8^2 = 28$ manières différentes de choisir 2 articles non défectueux parmi 8.

donc $\mu(A) = 6/66 = 1/11$ et $\mu(B) = 28/66 = 14/33$. Si on cherche à connaitre la probabilité qu'au moins un article est défectueux, c'est le complémentaire de B, donc on obtient la probabilité de 1-14/33 = 19/33.

Exemple : Soit un ensemble de n personnes, on cherche à calculer la probabilité que ces n personnes aient des dates de naissances tous distincts. Pour cela si on designe par un n-uplet $(d_1, d_2, ...d_n)$ où d_i est la date de naissance de la i personne, les dates de naissances des n personnes, l'univers Ω est alors l'ensemble de tous les n-uplet possibles. Et on cherche à dénombrer le nombre de n-uplet dont il n'y a pas deux d_i identiques. La première personne a le choix de 365 jours possibles, la seconde de 364 possible, ... , et ainsi de suite, la n personne $365 - (n - 1)$ choix possible. Comme le cardinal de Ω est 365^n la probabilité cherchée est donc de

$$\frac{365 \cdot 364 \cdot 363 \cdots (365 - (n - 1))}{365^n}$$

Paragraphe 6 : Probabilité conditionnelle

Soient un espace probabilisé $(\Omega, \mathcal{A}, \mu)$ et B un événement tel que $\mu(B) > 0$. Comme B est un événement, c'est aussi un sous ensemble de Ω, on peut donc considérer la tribu trace de $\mathcal{A}$ sur B : $\{ A \cap B \ / \ A \in \mathcal{A} \}$. Sur cette tribu on considère la probabilité μ_B suivante :

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

Vérifions que c'est bien une probabilité. Il est clair que $\mu_B(A) \in [0, 1]$, que $\mu_B(\Omega) = 1$, reste juste à vérifier la troisième condition. Soit une suite d'éléments A_i de $\mathcal{A}$ deux à deux disjoints on a

$$\mu_B(\cup A_i) = \mu(\cup A_i \mid B) = \frac{\mu((\cup A_i) \cap B)}{\mu(B)} = \frac{\mu(\cup(A_i \cap B))}{\mu(B)} = \frac{\sum \mu(A_i \cap B)}{\mu(B)} = \sum \mu(A_i \mid B) = \sum \mu_B(A_i)$$

ce qui démontre que μ_B est bien une probabilité sur la tribu trace de $\mathcal{A}$ sur B. On dit aussi que $\mu_B(A) = \mu(A \mid B)$ est la probabilité de l'événement A sachant que l'événement B est réalisé ou plus simplement la probabilité de A sachant B.

Exemple : On jette une paire de dés non pipés. On note A l'évènnement : un 2 est sorti, et B l'événement : la somme est égal à 6, calculer la probabilité de A sachant B. L'univers Ω correspond à l'ensemble des couples (i, j) où i et j varient entre 1 et 6, représentant le numéro sorti. On veut donc calculer la probabilité qu'un 2 est tiré sachant que la somme est 6. Les couples dont la somme est 6 sont $(1, 5)$, $(2, 4)$, $(3, 3)$, $(4, 2)$, $(5, 1)$, et les couples contenant au moins un 2 sont $(2, 4)$, $(4, 2)$. Donc la probabilité cherché est $2/5$. Noter que la probabilité de A est $11/36$, ce qui est inférieur à $2/5$.

Exemple : Une personne rend visite à une famille ayant deux enfants. L'un des enfants est un garçon et entre dans la pièce. Calculer la probabilité pour que l'autre enfant soit un garçon sachant qu'il est plus jeune. Et si on ne sait rien. L'univers $\Omega = \{ (g, g), \ (g, f), \ (f, g), \ (f, f) \}$ où on suppose que la première composante correspond à l'ainé. Si on note A l'événément : l'un des enfants est un garçon, alors les couples qui satisfont à cette conditions sont (g, g), (g, f), (f, g) et B l'événement : le plus jeune est un garçon, ce sont les couples (g, g), (f, g). On veut calculer $\mu(\{(g, g)\} \mid A)$ ou $\mu(\{(g, g)\} \mid B)$, or on sait que $\mu(A) = 3/4$ et $\mu(B) = 2/4$. Si l'enfant qui entre est un garçon on veut calculer $\mu(\{(g, g)\} \mid A) = \mu(\{(g, g)\} \cap A)/\mu(A) = (1/4)/(3/4) = 1/3$. Si le plus jeune est un garçon alors on veut calculer $\mu(\{(g, g)\} \mid B) = \mu(\{(g, g)\} \cap B)/\mu(B) = (1/4)/(2/4) = 2/4$.

La relation $\mu(A \mid B) = \mu(A \cap B)/\mu(B)$ peut s'écrire $\mu(A \cap B) = \mu(B)\mu(A \mid B)$. Supposons que l'on dispose de quatre événements A_1, A_2, A_3, et A_4, alors on peut écrire :

$$\mu(A_1 \cap A_2 \cap A_3) = \mu(A_3 \cap (A_1 \cap A_2)) = \mu(A_1 \cap A_2)\mu(A_3 \mid A_1 \cap A_2) = \mu(A_1)\mu(A_2 \mid A_1)\mu(A_3 \mid A_1 \cap A_2)$$

ce qui donne avec quatre événements

$$\begin{aligned} \mu(A_1 \cap A_2 \cap A_3 \cap A_4) &= \mu(A_1 \cap A_2 \cap A_3)\mu(A_4 \mid A_1 \cap A_2 \cap A_3) \\ &= \mu(A_1)\mu(A_2 \mid A_1)\mu(A_3 \mid A_1 \cap A_2)\mu(A_4 \mid A_1 \cap A_2 \cap A_3) \end{aligned}$$

d'ou plus généralement avec n événements on obtient la formule

$$\mu(A_1 \cap A_2 \cdots \cap A_n) = \mu(A_1)\mu(A_2 \mid A_1)\mu(A_3 \mid A_1 \cap A_2) \cdots \mu(A_n \mid A_1 \cap A_2 \cdots A_{n-1})$$

Exemple : Soit un lot de 12 articles dont quatre sont défecteux, on tire à la suite sans remise 3 articles, calculer la probabilité de n'obtenir aucun article défectueux. La probabilité de tirer un article non défectueux lors du premier tirage est de 8/12, lors du second tirage 7/11, puis de 6/10 pour le dernier tirage. Donc la probabilité cherchée est de (8/12)(7/11)(6/10)=14/55.

Paragraphe 7 : Processus stochastiques finis

On appelle processus stochastiques fini une suite fini d'expériences dont chacune a un nombre fini de résultats possibles avec des probablilités données. Un moyen de le représenter est d'utiliser un diagramme en arbre. Le résultat du paragraphe précédent permet de calculer la probabilité pour que un résultat représenté par un chemin donné soit obtenu.

Exemple : On dispose de trois boites telles que

1. La boite numéro 1 contient 10 ampoules dont 4 sont défectueux,
2. La boite numéro 2 contient 8 ampoules dont 2 sont défectueux,
3. La boite numéro 3 contient 13 ampoules dont 3 sont défectueux,

On choisit une boite au hasard et on tire une ampoule, calculer la probabilité que l'ampoule soit défectueuse. Il y a 1/3 de chance de tirer la première boite et 4/10 de tirer une ampoule défectueuse, ainsi la probabilité de tirer une ampoule défectueuse dans la boite numéro 1 est de $1/3 \times 4/10 = 2/15$. Le même raisonnement sur la seconde boite donne $1/3 \times 2/8 = 1/12$, et pour la troisième $1/3 \times 3/13 = 1/13$. Au final on obtient la probabilité de 2/15+1/12+1/13.

Exemple : On jette une pièce de monnaie biaisée de telle sorte que $\mu(F) = 2/3$ et $\mu(P) = 1/3$. Si c'est F qui apparait, on choisit au hasard un nombre entre 1 et 9, si c'est P qui apparaît, on choisit un nombre entre 1 et 5. Calculer la probabilité que ce soit un nombre pair. Il y a 4 nombres pairs entre 1 et 9, donc la probabilité de tirer un nombre pair entre 1 et 9 est de 4/9. De même la probabilité de tirer un nombre pair entre 1 et 5 est de 2/5. Donc la probabilité cherchée est de $2/3 \times 4/9 + 1/3 \times 2/5$.

Paragraphe 8 Théorème de Bayes

Soient l'univers Ω et $(A_i)_{i \in I}$ une famille fini ou dénombrable d'événements de $\mathcal{P}(\Omega)$. On dit que cette famille forme une partition de Ω si $A_i \cap A_j = \emptyset$ et $\Omega = \cup_{i \in I} A_i$. Supposons donné une partition de Ω telle que $(\forall i \in I)(\mu(A_i) > 0)$, et soit B un élément de $\mathcal{P}(\Omega)$, alors

$$\begin{aligned} \mu(B) &= \mu(\Omega \cap B) \\ &= \mu((\cup_{i \in I} A_i) \cap B) \\ &= \sum_{i \in I} \mu(A_i \cap B) = \sum_{i \in I} \mu(A_i)\mu(B \mid A_i) \end{aligned}$$

Soit un $j \in I$ arbitraire on a alors :

$$\mu(A_j \mid B) = \frac{\mu(A_j \cap B)}{\mu(B)} = \frac{\mu(A_j)\mu(B \mid A_j)}{\sum_{i \in I} \mu(A_i)\mu(B \mid A_i)}$$

Cette formule est appelé la formule de Bayes.

Exemple : On considère trois machines M_1, M_2 et M_3 produisant respectivement 50, 30 et 20 pièces, dont 3 défectueuses pour M_1, 5 pour M_2 et 1 pour M_3. Si on prend au hasard une pièce, quelle est la probabilité qu'elle soit défectueuse ? Supposons que l'on ait tiré une pièce défectueuse, quelle est la probabilité qu'elle ait été fabriqué par la machine M_1. L'univers Ω correspond à l'ensemble des 100 pièces, M_1, M_2 et M_3 forment une partition de Ω (en identifiant M_i avec l'ensemble des pièces qu'elle a fabriqué). Notons A l'événement tirer une pièce défectueuse, alors on a

$$\begin{aligned}
\mu(A) = \mu((M_1 \cup M_2 \cup M_3) \cap A) &= \mu(M_1 \cap A) + \mu(M_2 \cap A) + \mu(M_3 \cap A) \\
&= \mu(M_1)\mu(A \mid M_1) + \mu(M_2)\mu(A \mid M_2) + \mu(M_3)\mu(A \mid M_3) \\
&= 50/100 \times 3/50 + 30/100 \times 5/30 + 20/100 \times 1/20 = 9/100
\end{aligned}$$

Maintenant on veut calcule $\mu(M_1 \mid A)$, la formule de Bayes donne

$$\mu(M_1 \mid A) = \frac{\mu(M_1)\mu(A \mid M_1)}{\mu(M_1)\mu(A \mid M_1) + \mu(M_2)\mu(A \mid M_2) + \mu(M_3)\mu(A \mid M_3)} = \frac{3/100}{9/100} = 1/3$$

soit une chance sur trois que la pièce défectueuse soit sortie de M_1.

Exemple : On applique un test médical sur des patients. Si le patient est malade, le test est positif dans 99%, mais il y a aussi 2% de cas où le test se révèle positif mais le patient n'est pas malade. Sachant qu'il 0,1% de la population qui est atteint, calculer la probabilité qu'un patient soit malade sachant que le test est positif. Notons M l'événement le patient est malade et T l'événement le test est positif. L'univers Ω est l'ensemble de la population, la partition ceux qui sont malades M et ceux qui ne le sont pas noté $\overline{M}$. Alors on a

$$\mu(T \mid M) = 0,99, \qquad \mu(T \mid \overline{M}) = 0,02, \qquad \mu(M) = 1/1000$$

et la formule de Bayes donne :

$$\mu(M \mid T) = \frac{\mu(T \mid M) \times \mu(M)}{\mu(M) \times \mu(T \mid M) + \mu(\overline{M}) \times \mu(T \mid \overline{M})} = 1/21$$

Paragraphe 9 : Événements indépendants

Soient un espace probabilisé $(\Omega, \mathcal{A}, \mu)$, A et B deux événements de Ω. On dit qu'ils sont indépendants selon la probabilité μ ou simplement indépendants quand aucune confusion n'est possible si

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

Notons que si $\mu(A) > 0$ et $\mu(B) > 0$ alors on a la suite d'équivalence

$$\mu(A \cap B) = \mu(A) \cdot \mu(B) \Longleftrightarrow \mu(A \mid B) = \mu(A) \Longleftrightarrow \mu(B \mid A) = \mu(B)$$

On dit que les événements A et B sont disjoints si $A \cap B = \emptyset$, c'est donc une propriété ensembliste, et ils sont dis incompatibles si $\mu(A \cap B) = 0$. Deux événements disjoints sont donc incompatibles, mais la réciproque est fausse. Enfin deux événements disjoints sont indépendant si la probabilité d'un des deux événements est nulle.

Exemple : On jette trois fois de suite une pièce de monnaie non pippé, l'ensemble Ω est donc

$$\Omega = \{FFF, FFP, FPF, FPP, PFF, PFP, PPF, PPP\}$$

On choisit $\mathcal{A} = \mathcal{P}(\Omega)$ et μ la probabilité uniforme. Soient les trois événements

1. $A = \{FFF, FFP, FPF, FPP\}$ le premier jet donne face,

2. $B = \{FFF, FFP, PFF, PFP\}$ le second jet donne face,
3. $C = \{FFF, FFP, PFF\}$ deux jets consécutifs de face.

On a alors

$$A \cap B = \{FFF, FFP\}, \quad A \cap C = \{FFF, FFP\}, \quad B \cap C = \{FFF, FFP, PFF\}.$$

et donc

$$\mu(A) = \mu(B) = 1/2, \quad \mu(C) = 3/8, \quad \mu(A \cap B) = 1/4, \quad \mu(A \cap C) = 1/4, \quad \mu(B \cap C) = 3/8.$$

Donc les événements A et B sont indépendants, mais A et C ainsi que B et C ne sont pas indépendants. On notera qu'il est naturel que A et B sont indépendants, il n'y a aucune raison que le second jet donne face soit influencé par le résultat du premier jet.

Exemple : On jette deux pièces de monnaies non pippées, l'ensemble Ω est $\{FF, FP, PF, PP\}$, on choisit $\mathcal{A} = \mathcal{P}(\Omega)$ et μ la probabilité uniforme. Soient les trois événements suivants

1. $A = \{FF, FP\}$ face apparaît sur la première pièce,
2. $B = \{FF, PF\}$ face apparaît sur la seconde pièce,
3. $C = \{FP, PF\}$ face apparaît sur une seule des deux pièces.

On a alors

$$\mu(A) = \mu(B) = \mu(C) = 1/2, \quad \mu(A \cap B) = \mu(A \cap C) = \mu(B \cap C) = 1/4, \quad \mu(A \cap B \cap C) = 0$$

Autrement dit les événements A, B, et C sont deux à deux indépendants mais ne sont pas mutuellement indépendants. On dit qu'une famille fini d'événements $(A_i)_{1 \leq i \leq n}$ sont mutuellement indépendants si

$$(\ \forall J \subset \{1, 2, ..., n\}\ )(\ \mu(\cap_{i \in J} A_i) = \prod_{i \in J} \mu(A_i)\ )$$

Proposition 9.1

Soient un espace probabilisé $(\Omega, \mathcal{A}, \mu)$, A et B deux événements de Ω. Si A et B sont indépendants alors A et $\overline{B}$, $\overline{A}$ et B, $\overline{A}$ et $\overline{B}$, sont indépendants. De plus si $\mu(A) = 0$ ou 1 alors tout événement $C \in \mathcal{A}$ est indépendant à A.

Preuve :

On a

$$\mu(A) = \mu(A \cap B) + \mu(A \cap \overline{B}) = \mu(A)\mu(B) + \mu(A \cap \overline{B}),$$

donc

$$\mu(A \cap \overline{B}) = \mu(A) - \mu(A)\mu(B) = \mu(A)(1 - \mu(B)) = \mu(A)\mu(\overline{B})$$

Le reste s'obtient pas symétrie. Maintenant si $\mu(A) = 0$, comme $\mu(A \cap B) \leq \mu(A)$, obligatoirement $\mu(A \cap B) = 0$, et donc $\mu(A \cap B) = 0 = \mu(A)\mu(B)$. Enfin si $\mu(A) = 1$ alors $\mu(\overline{A}) = 0$ c'est à dire que $\overline{A}$ est indépendant à tout événement, et par symétrie A l'est.

Paragraphe 10 : Variable aléatoire, espérance, variance

Soient deux espaces probabilisables $(\Omega_1, \mathcal{A}_1)$ et $(\Omega_2, \mathcal{A}_2)$. Une application X de Ω_1 dans Ω_2 est appelée une variable aléatoire si l'image réciproque de tout élément $B \in \mathcal{A}_2$ par X, $X^{-1}(B)$ est un élément de $\mathcal{A}_1$. Ainsi X est une variable aléatoire si

$$(\ \forall B \in \mathcal{A}_2\ )(\ X^{-1}(B) = \{\ \omega \in \Omega_1\ /\ X(\omega) \in B\ \} \in \mathcal{A}_1\ )$$

Si $\Omega_2 = \mathbb{R}$ et $\mathcal{A}_2$ est la tribu engendrée par les intervalles ouverts de $\mathbb{R}$, une variable aléatoire de Ω_1 dans $\mathbb{R}$ est appelé une variable aléatoire réelle, la somme et le produit de deux variables aléatoires réelles est une variable aléatoire. Noter que si Ω_1 est un ensemble fini et $\mathcal{A}_1 = \mathcal{P}(\Omega_1)$, le résultat est clair. Soit X une variable aléatoire réelle, si a est un réel, au lieu de noté $X^{-1}(\{a\})$ l'image réciproque par X du singleton $\{a\}$ il est noté $(X = a)$, de même $X^{-1}([\alpha, \beta]) = (\alpha \leq X \leq \beta)$. La notation s'étend au cas général d'une variable aléatoire X entre Ω_1 et Ω_2.
On suppose que l'espace probabilisable $(\Omega_1, \mathcal{A}_1)$ est muni d'une probabilité μ_1, soit l'espace probabilisé $(\Omega_1, \mathcal{A}_1, \mu_1)$, la donné d'une variable aléatoire X entre Ω_1 et Ω_2 permet de transformer l'espace probabilisable $(\Omega_2, \mathcal{A}_2)$ en un espace probabilisé en posant sur $\mathcal{A}_2$ la probabilité définie comme suit

$$(\forall A_2 \in \mathcal{A}_2)(\mu_2(A_2) = \mu_1(X = A_2) = \mu_1(X^{-1}(A_2)) = \mu(\{\omega \in \Omega_1 \ / \ X(\omega) \in A_2 \ \} \ )$$

On note que μ_2 est parfaitement bien définie car X est une varaible aléatoire et donc l'ensemble $(X = A_2)$ est par construction un élément de $\mathcal{A}_1$ et donc on peut prendre sa probabilité. Cette probabilité μ_2 s'appelle la loi de la variable aléatoire X ou plus simplement la loi de X, elle est notée P_X. Donc $P_X(A_2) = \mu_1(X = A_2)$.
On dit qu'une variable aléatoire X est discrète si l'ensemble d'arrivé Ω_2 est un ensemble fini ou dénombrable. Soit donc une variable aléatoire discrète X, notons $\Omega_2 = \{\omega_1, \omega_2, ...\omega_n\} = X(\Omega_1)$ où l'on suppose que c'est un ensemble fini, on définit alors l'espérance (*resp* la variance) de X noté $E(X)$ (*resp* $Var(X)$) comme étant le nombre

$$E(X) = \sum_{i=1}^{n} \omega_i \mu(X = \omega_i) = \omega_1 \mu(X = \omega_1) + \omega_2 \mu(X = \omega_2) + \cdots + \omega_n \mu(X = \omega_n)$$

$$(\ resp \ \ Var(X) = \sum_{i=1}^{n} (\omega_i - E(X))^2 \mu(X = \omega_i) = E(X^2) - E^2(X) \ )$$

Si Ω_2 est un ensemble discret infini, on définit l'espérance(*resp* la variance) de X comme étant la limite, si elle existe, de la suite

$$u_n = \sum_{i=1}^{n} \omega_i \mu(X = \omega_i)$$

$$(\ resp \ \ v_n = \sum_{i=1}^{n} (\omega_i - E(X))^2 \mu(X = \omega_i) \ )$$

Noter que cette espérance et ou variance peut ne pas exister.

Proposition 10.1

On suppose que Ω_2 est contenu dans $\mathbb{R}$, soient deux variables aléatoires X et Y définies sur le même espace probabilisable $(\Omega_1, \mathcal{A}_1)$, alors pour tout $\alpha \in \mathbb{R}$, $X + \alpha$, αX, $X + Y$, et XY sont des variables aléatoires.

Preuve :
Si $\mathcal{A}_1 = \mathcal{P}(\Omega_1)$ le résultat est clair, dans le cas contraire on admet le résultat d'autant plus que la définition d'une variable aléatoire réelle n'est pas clair. L'ensemble $\mathcal{A}_2$ n'est pas précisé.

Exemple : On lance une paire de dés non pippés, l'espace Ω_1 est l'ensemble des 36 couples possibles, la première (*resp* seconde) composante indiquant le nombre sorti pour le premier(*resp* second) dé

$$\Omega_1 = \{ \ (i, j) \ /1 \leq i \leq 6 \quad 1 \leq j \leq 6 \ \}$$

On choisit $\mathcal{A}_1 = \mathcal{P}(\Omega)$ et μ_1 la probabilité uniforme et la variable aléatoire X définie par $X(i,j) = \text{Max}(i,j)$. L'ensemble Ω_2 est donc $\{1,2,3,4,5,6\}$ avec $\mathcal{A}_2 = \mathcal{P}(\Omega_2)$. Comme on a

— $(X = 1) = \{(1,1)\}$,

— $(X = 2) = \{(1,2),(2,1)(2,2)\}$,

— $(X = 3) = \{(1,3),(2,3),(3,3),(3,2),(3,1)\}$,

— $(X = 4) = \{(1,4),(2,4),(3,4),(4,4),(4,3),(4,2),(4,1)\}$,

— $(X = 5) = \{(1,5),(2,5),(3,5),(4,5),(5,5),(5,4),(5,3),(5,2),(5,1)\}$,

— $(X = 6) = \{(1,6),(2,6),(3,6),(4,6),(5,6),(6,6),(6,5),(6,4),(6,3),(6,2),(6,1)\}$.

La loi de X est alors décrite comme suit

$$P_X(1) = 1/36,\ P_X(2) = 3/36,\ P_X(3) = 5/36,\ P_X(4) = 7/36,\ P_X(5) = 9/36,\ P_X(6) = 11/36.$$

Son espérance est donnée par

$$E(X) = 1\frac{1}{36} + 2\frac{3}{36} + 3\frac{5}{36} + 4\frac{7}{36} + 5\frac{9}{36} + 6\frac{11}{36} = \frac{161}{36}$$

Pour calculer la variance $Var(X)$, on a :

$$E(X^2) = 1^2\frac{1}{36} + 2^2\frac{3}{36} + 3^2\frac{5}{36} + 4^2\frac{7}{36} + 5^2\frac{9}{36} + 6^2\frac{11}{36} = \frac{791}{36}$$

et donc

$$Var(X) = E(X^2) - E^2(X) = \frac{791}{36} - (\frac{161}{36})^2 = \frac{2555}{36^2}$$

Si on veut par exemple calculer la probabilité que le maximum tiré soit un nombre impair, il s'agit de l'événement $\{1,3,5\} \in \mathcal{A}_2$, donc sa probabilité est

$$P_X(\{1,3,5\}) = P_X(1) + P_X(3) + P_X(5) = 15/36$$

Exemple : On lance trois fois de suite une pièce de monnaie truquée telle que la probabilité d'obtenir face F est de $2/3$ et pile P de $1/3$. L'ensemble Ω_1 est $\{FFF, FFP, FPF, FPP, PFF, PFP, PPF, PPP\}$, on considère la variable aléatoire qui à un lancé associe le plus grand nombre de faces successive, l'ensemble Ω_2 est donc $\{0,1,2,3\}$. Comme on a

— $(X = 0) = \{PPP\}$,

— $(X = 1) = \{FPF, FPP, PFP, PPF\}$,

— $(X = 2) = \{FFP, PFF\}$,

— $(X = 3) = \{FFF\}$.

La loi de X est alors donnée par

$$P_X(0) = \mu(PPP) = 1/27,\ P_X(1) = \mu(\{FPF, FPP, PFP, PPF\}) = 4/27{+}2/27{+}2/27{+}2/27 = 10/27,$$

$$P_X(2) = \mu(\{FFP, PFF\}) = 4/27 + 4/27 = 8/27,\ P_X(3) = \mu(FFF) = 8/27$$

Son espérance est donnée par

$$E(X) = 0\frac{1}{27} + 1\frac{10}{27} + 2\frac{8}{27} + 3\frac{8}{27} = \frac{53}{27}$$

Pour calculer la variance $Var(X)$, on a :

$$E(X^2) = 0^2\frac{1}{27} + 1^2\frac{10}{27} + 2^2\frac{8}{27} + 3^2\frac{8}{27} = \frac{114}{27}$$

et donc

$$Var(X) = E(X^2) - E^2(X) = \frac{114}{27} - (\frac{53}{27})^2 = \frac{269}{27^2}$$

Exemple : Une boite contient douze articles dont trois sont défectueux. On tire trois articles, l'espace Ω_1 est l'ensemble constitué de tous les triplets possibles formés par les douze articles, $\mathcal{A}_1 = \mathcal{P}(\Omega_1)$, on considère la variable aléatoire X qui à un triplet indique le nombre d'articles défectueux. Donc $\Omega_2 = \{0, 1, 2, 3\}$, et la loi de X n'est autre que la probabilité de tirer un lots sans aucune pièce défectueuse, avec une pièce défectueuse, ... On a donc

— $P_X(0) = \mu(X = 0)$ la probabilité de tirer trois articles dont aucun n'est défectueux,
— $P_X(1) = \mu(X = 1)$ la probabilité de tirer trois articles dont un et un seul est défectueux,
— $P_X(2) = \mu(X = 2)$ la probabilité de tirer trois articles dont deux et deux seulement sont défectueux,
— $P_X(3) = \mu(X = 3)$ la probabilité de tirer trois articles tous défectueux.

Or

— le nombre de triplets possible, soit le cardinal de Ω_1 est $C_{12}^3 = 220$,
— le nombre de triplets sans aucune pièce défectueuse est $C_9^3 = 84$,
— le nombre de triplets avec une pièce défectueuse est $3C_9^2 = 108$,
— le nombre de triplets avec deux pièces défectueuse est $9C_3^2 = 27$,
— et il n'y a qu'un seule triplet avec toute les pièces défectueuses.

On a donc

$$P_X(0) = 84/220, \quad P_X(1) = 108/220, \quad P_X(2) = 27/220, \quad P_X(3) = 1/220$$

Son espérance est donnée par

$$E(X) = 0\frac{84}{220} + 1\frac{108}{220} + 2\frac{27}{220} + 3\frac{1}{220} = \frac{165}{220}$$

Pour calculer la variance $Var(X)$, on a :

$$E(X^2) = 0^2\frac{84}{220} + 1^2\frac{108}{220} + 2^2\frac{27}{220} + 3^2\frac{1}{220} = \frac{225}{220}$$

et donc

$$Var(X) = E(X^2) - E^2(X) = \frac{225}{220} - (\frac{165}{220})^2 = \frac{22275}{220^2}$$

Paragraphe 11 : Fonction de répartition.

On suppose que Ω_2 est contenu dans $\mathbb{R}$, soit une variable aléatoire X de l'univers Ω_1 dans Ω_2. On définit la fonction de répartition de la variable aléatoire X comme la fonction notée F_X

$$F_X \ : \mathbb{R} \longrightarrow \mathbb{R} \ / \ x \longrightarrow F_X(x) = P_X(]-\infty, x]) = \mu(X \leq x)$$

Proposition 11.1

La fonction de répartition F_X d'une variable aléatoire X vérifie les propriétés suivantes :

— $(\ \forall x \in \mathbb{R}\ )(\ 0 \leq F_X(x) \leq 1\ )$
— est une fonction croissante,
— est continue à droite
— $\lim_{x\to-\infty} F_X(x) = 0$ et $\lim_{x\to+\infty} F_X(x) = 1$

Preuve :

La première propriété découle de la définition de probabilité, il n'y a rien à dire. Pour deux réels α et β tels que $\alpha \leq \beta$ on a $]-\infty, \alpha] \subset]-\infty, \beta]$, donc $F_X(\alpha) \leq F_X(\beta)$. Pour démontrer le troisième point, considérons une suite de réelles (u_n) décroissante de limite 0. Pour tout $\alpha \in \mathbb{R}$ on a alors $]\alpha, \alpha + u_n] =]-\infty, \alpha + u_n] \backslash]-\infty, \alpha]$ et donc $P_X(]\alpha, \alpha + u_n]) = F_X(\alpha + u_n) - F_X(\alpha)$. Mais la suite d'intervalles $]\alpha, \alpha + u_n]$ est décroissante au sens de l'inclusion dont la limite est l'ensemble vide, on a donc

$$\lim_{n \to +\infty} P_X(]\alpha, \alpha + u_n]) = P_X(\cap_{n\in\mathbb{N}}]\alpha, \alpha + u_n]) = P_X(\emptyset) = 0$$

On en déduit donc

$$\lim_{n \to +\infty} F_X(\alpha + u_n) = F_X(\alpha)$$

c'est à dire que F_X est continue à droite. Pour démontrer le dernier point, on considère la suite d'intervalles $]-\infty, -n]$, qui est décroissante de limite l'ensemble vide $\emptyset$, on a donc

$$\lim_{n \to +\infty} F_X(-n) = \lim_{n \to +\infty} P_X(]-\infty, -n]) = P_X(\cap_{n=1}^{+\infty}]-\infty, -n]) = P_X(\emptyset) = 0$$

Comme $F_X(-x)$ est une fonction décroissante, on en déduit le résultat. Pour démontrer l'autre limite, il suffit de considérer la suite croissante d'intervalles $]-\infty, n]$.

Pour aller un peu plus loin.

Soient deux espaces probabilisables $(\Omega_1, \mathcal{A}_1)$ et $(\Omega_2, \mathcal{A}_2)$ tel que $\Omega \subset \mathbb{R}$. Une variable aléatoire X de Ω_1 dans Ω_2 est dite continue si sa fonction de répartition est continue. Donc une variable aléatoire non discrète n'est pas forcément continue. On dit qu'une variable aléatoire continue X admet une densité si il existe une fonction f telle que

$$F_X(x) = \int_{-\infty}^{x} f(t)dt$$

La fonction f est appelée la densité de la variable aléatoir X. Pour une telle variable aléatoire, on définit son espérance comme étant

$$E(X) = \int_{\mathbb{R}} tf(t)dt$$

et sa variance

$$Var(X) = E((X - E(X)^2) = \int_{\mathbb{R}} (t - E(X))^2 f(t)dt$$

Noter que ces quantités ne sont pas forcément bien définies.

Proposition 11.2

Soient une variable aléatoire X (resp Y) continue de densité f (resp g) et $\alpha \in \mathbb{R}$ un nombre réel. On suppose que l'espérance et la variance de X de Y sont bien définies, alors on a

— $E(X + \alpha) = E(x) + \alpha$,
— $E(\alpha X) = \alpha E(X)$,
— $E(X + Y) = E(X) + E(Y)$,
— $Var(X + \alpha) = Var(X)$,
— $Var(\alpha X) = \alpha^2 Var(X)$,

Preuve :
Découle simplement des propriétés de l'intégrale.

$$E(X + \alpha) = \int_{\mathbb{R}} (t + \alpha)f(t)dt = \int_{\mathbb{R}} tf(t)dt + \alpha \int_{\mathbb{R}} f(t)dt = E(X) + \alpha$$

$$Var(X+\alpha) = E((X+\alpha)^2) - E^2(X+\alpha) = \int_{\mathbb{R}} (t+\alpha)^2 f(t)dt - (E(X)+\alpha)^2$$
$$= \int_{\mathbb{R}} t^2 f(t)dt + 2\alpha \int_{\mathbb{R}} tf(t)dt + \alpha^2 \int_{\mathbb{R}} f(t)dt - E^2(X) - 2\alpha E(X) - \alpha^2$$
$$= \int_{\mathbb{R}} t^2 f(t)dt - E^2(X) = E(X^2) - E^2(X) = Var(X)$$

Soient deux variables aléatoires X et Y définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mu)$ à valeurs respectives dans $(\Omega_X, \mathcal{A}_X)$ et $(\Omega_Y, \mathcal{A}_Y)$. On dit que ces deux variables aléatoires sont indépendantes si

$$(\ \forall A_X \in \mathcal{A}_X\ )(\ \forall A_Y \in \mathcal{A}_Y\ )(\ \mu((X = A_X) \cap (Y = A_Y)) = \mu(X = A_X) \times \mu(Y = A_Y)\ )$$

La définition se généralise à une famille finie de variables aléatoires. Revenons à la situation des deux variables aléatoires X et Y sans hypothèse d'indépendance mais on suppose qu'elles sont distrètes. On peut alors définir sur l'ensemble produit $\Omega_X \times \Omega_Y$ une probabilité en posant $\nu(x_i, y_j) = \mu((X = x_i) \cap (Y = y_j))$. Il faut montrer que ν est bien une probabilité. Il est clair que ν est à valeurs comprises entre 0 et 1. Il reste à montrer que $\nu(\Omega_1 \times \Omega_2) = 1$. Pour cela on note $\Omega_X = \{x_1, x_2, ...x_m, ...\}$ et $\Omega_Y = \{y_1, y_2, ...y_n, ...\}$, on a alors

$$\Omega_X \times \Omega_y = \cup_{(i,j)} \{(x_i, y_j)\}$$

Soient deux indices j_1 et j_2 distincts, comme $(X = x_i, Y = y_j) = \{\ \omega \in \Omega\ / X(\omega) = x_i \text{ et } Y(\omega) = y_j\ \}$ alors $(X = x_i, Y = y_{j_1}) \cap (X = x_i, Y = y_{j_2}) = \emptyset$ car les conditions $Y = y_{j_1}$ et $Y = y_{j_2}$ sont incompatibles, de plus $\cup_j (X = x_i, Y = y_j) = (X = x_i)$, donc

$$\mu(X = x_i) = \mu\left(\cup_j (X = x_i, Y = y_j)\right) = \sum_j \mu(X = x_i, Y = y_j) = \sum_j \nu(x_i, y_j)$$

on en déduit alors

$$1 = \mu(\Omega) = \mu\left(\cup_i (X = x_i)\right) = \sum_i \mu(X = x_i) = \sum_i \sum_j \nu(x_i, y_j)$$

qui est le résultat cherché.

Paragraphe 12 : Exemples de lois discrètes.

On se donne une variable aléatoire réelle X de Ω_1 dans $\Omega_2 \subset \mathbb{R}$.

— Loi uniforme discrète.
  La variable aléatoire X est dite de loi uniforme discrète finie si Ω_2 est un ensemble fini et $P_X(\{\omega\}) = \mu(X = \omega) = 1/\mathrm{Card}\Omega_2$.

— Loi de Bernouilli.
  La variable aléatoire X est dite de Bernouilli de paramètre $p \in [0, 1]$ si $\Omega_2 = \{0, 1\}$ et $P_X(\{0\}) = \mu(X = 0) = 1 - p$, $P_X(\{1\}) = \mu(X = 1) = p$.
  Par exemple si on veut modéliser le tirage dans une urne contenant des boules rouges et des boules bleues, l'événement $(X = 1)$ désigne le tirage donne une boule rouge, et $(X = 0)$ l'événement le tirage donne une boule bleue.

— Loi binômiale.
  La variable aléatoire X est dite de loi binômiale (n, p) où $n \in \mathbb{N}^*$ et $p \in [0, 1]$ noté $B(n, p)$ si $\Omega_2 = \{0, 1, 2, ..., n\}$ et $P_X(\{k\}) = \mu(X = k) = C_n^k pk(1 - p)^{n-k}$

Cette loi apparait par exemple si on effectue n épreuves successives indépendantes ou on note à chaque fois la réalisation ou non d'un certain événement A, on obtient une suite de A et de $\overline{A}$. À cette suite d'épreuves on associé l'événement ω et la variable aléatoire $X(\omega) =$ nombre de réalistaion de A. Cette variable suit une loi binomiale. La loi binômiale est une généralisation de la loi de Bernouilli.

— Loi hypergéométrique.
La variable aléatoire X est dite de loi hypergéométrique de paramètre (n, N, M) où n, N, et M sont des entiers tels que $M < N$ et $n \leq N$ si $\Omega_2 = \mathbb{N} \cap [\max(0, n - (N - M)), \min(n, M)]$ et

$$P_X(\{k\}) = \mu(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

Par exemple si on effectue n tirage successives sans remise dans une urne contenant N objets dont M possédant une certaine caractéristique, on note ω l'événement correspondant à ces n tirages successives et $X(\omega)$ le nombre d'objets tirés possédant cette caractéristique. Alors X suit une loi hypergéométrique. Effectuer n tirages successives sans remise est identique à tirer n objets parmi N, il y a donc C_N^n échantillons possibles. Pour calculer la probabilité d'obtenir k objets ayant cette caractéristique, il faut dénombrer tous les échantillons contenant exactement k objets ayant la caractéristique parmi M soit C_M^k. Chacun d'eux contient aussi exactement $n - k$ objets n'ayant pas la caractéristique, il y en a C_{N-M}^{n-k}, on trouve donc bien l'expression

$$P_X(\{k\}) = \mu(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

— Loi géométrique ou de Pascal.
La variable aléatoire X est dite de loi géométrique ou de Pascal de paramètre $p \in [0,]$ si $\Omega_2 = \mathbb{N}^*$ et $P_X(\{k\}) = \mu(X = k) = (1 - p)^{k-1} p$.
Par exemple on effectue des épreuves successives indépendantes jusqu'à la réalisation d'un événement A, et on note X le nombre d'épreuve effectuées. Calculer $\mu(X = k)$ correspond à calculer la probabilité qu'à la k - ième épreuve l'événement A est sorti, mais alors cet événement n'est pas sorti dans les $k - 1$ épreuves précédentes. Si on pose $\mu(A) = p$ alors on a $\mu(X = k) = (1 - p)^{k-1} p$.

— Loi binômiale négative.
La variable aléatoire X est dite de loi binômiale négative de paramètre (n, p) où $n \in \mathbb{N}^*$ et $p \in [0, 1]$ si $\Omega_2 = \mathbb{N}$ et $P_X(\{k\}) = \mu(X = k) = C_{k-1}^{n-1} p^n (1 - p)^{k-n}$.
Contrairement au cas précédent, on effectue une suite d'épreuves jusqu'à ce qu'il y ait exactement n éveénements A qui soient sortis. On note X le nombre d'épreuves effectuées. L'événement $(X = k)$ avec $k \geq n$ signifie qu'à la k épreuve l'événement A s'est réalisé exactement n fois et que durant les $k - 1$ épreuves précédentes, l'événement A s'est réalisé $n - 1$ fois ni plus ni moins. Si on suppose que $\mu(A) = p$ alors on a

$$P_X(\{k\}) = \mu(X = k) = C_{k-1}^{n-1} p^n (1 - p)^{k-n}$$

Le coefficient C_{k-1}^{n-1} correspond au nombre de manière de choisir $n - 1$ places parmi $k - 1$. Par rapport à la loi binômiale où c'est le nombre d'épreuves qui est fixé et où on observe le nombre de réalisation d'un certain événement, dans la binômiale négative c'est le nombre de réalistion d'un certain événement qui est fixé.