

Probabilités conditionnelles

I. Généralités

- Une **expérience aléatoire** est une expérience dont le résultat dépend du hasard. Chacun des résultats possibles s'appelle **éventualité** ou **issue**. L'ensemble Ω (ou E) de tous les résultats possibles d'une expérience aléatoire s'appelle l'**univers** de l'expérience.
- Un **événement** A est une partie de l'univers Ω (un événement peut donc être constitué de zéro, une ou plusieurs issues de Ω).
- Un **événement élémentaire** est une partie de Ω qui ne contient qu'une seule issue.
- L'**événement contraire** de A , noté $\bar{A}$, est la partie constituée de toutes les issues de Ω qui ne sont pas dans A .

- $P(\emptyset)=0$ et $P(\Omega)=1$.
- La probabilité d'un événement A , notée $P(A)$, est la **somme des probabilités des événements élémentaires inclus dans A** .
- Pour tout événement A , on a $0 \leq P(A) \leq 1$.
- La somme des probabilités des événements élémentaires vaut 1.
- La probabilité de l'événement contraire d'un événement A est : $P(\bar{A})=1-P(A)$.

Lorsque tous les événements élémentaires d'une expérience aléatoire ont la même probabilité, on dit alors qu'ils sont **équiprobables**. On parle d'**expérience équiprobable** ou de **loi équirépartie**.

- L'événement « **A et B** », noté $A \cap B$, s'appelle l'**intersection** des événements A et B . Il est réalisé lorsque **les deux événements sont réalisés simultanément**.
- L'événement « **A ou B** », noté $A \cup B$, s'appelle la **réunion** des événements A et B . Il est réalisé lorsqu'**au moins l'un des deux événements est réalisé**.
- Deux événements A et B sont dits **incompatibles** si $A \cap B = \emptyset$.

Soit A un événement d'un univers Ω . Pour une loi équirépartie, on a :

$$P(A) = \frac{\text{nombre d'éléments de } A}{\text{nombre d'éléments de } \Omega} = \frac{\text{nombre de cas favorables à } A}{\text{nombre de cas possibles}}$$

Soit A et B deux événements d'un univers Ω .

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- Si A et B sont **incompatibles**, $P(A \cup B) = P(A) + P(B)$.

II. Probabilités conditionnelles

Soit A et B deux événements d'un univers Ω avec $P(A) \neq 0$.

La **probabilité conditionnelle** que l'événement B se réalise sachant que l'événement A est

réalisé est notée $P_A(B)$, et elle est définie par : $P_A(B) = \frac{P(A \cap B)}{P(A)}$.

La probabilité $P_A(B)$ vérifie bien $0 \leq P_A(B) \leq 1$ et $P_A(B) + P_A(\bar{B}) = 1$.

Soit A et B deux événements d'un univers Ω avec $P(A) \neq 0$ et $P(B) \neq 0$. Alors :

$$P(A) \times P_A(B) = P(B) \times P_B(A)$$

Utilisation de tableaux

Les tableaux à double entrées permettent une présentation claire de certaines expériences aléatoires et facilitent le calcul des probabilités conditionnelles.

	B	$\bar{B}$	Total
A	$P(A \cap B)$	$P(A \cap \bar{B})$	$P(A)$
$\bar{A}$	$P(\bar{A} \cap B)$	$P(\bar{A} \cap \bar{B})$	$P(\bar{A})$
Total	$P(B)$	$P(\bar{B})$	1

- $P(A \cap B)$ se lit à l'intersection de la ligne A et de la colonne B .
- $P(A)$ (respectivement $P(B)$) se lit sur la dernière colonne (respectivement la dernière ligne).
- $P_A(B)$ (ou $P_B(A)$) s'obtient en calculant le quotient des deux probabilités adéquates :

$$P_A(B) = \frac{P(A \cap B)}{P(A)} \text{ et } P_B(A) = \frac{P(A \cap B)}{P(B)}$$

Exemple

Un club sportif rassemble 180 membres répartis en juniors et seniors. On compte 135 seniors dont 81 hommes. Il y a 27 garçons parmi les seniors.

En choisissant une femme au hasard, calculer la probabilité d'avoir une junior.

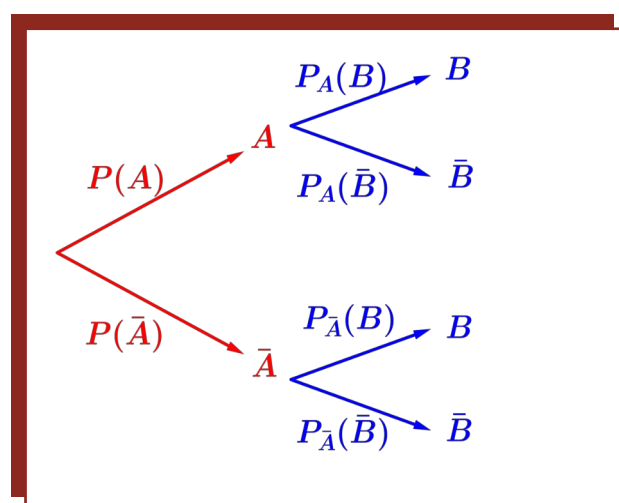
	J	S	Total
H	$\frac{27}{180} = 0,15$	$\frac{81}{180} = 0,45$	$0,15 + 0,45 = 0,6$
F	0,1	0,3	0,4
Total	0,25	$\frac{135}{180} = 0,75$	1

Ainsi $P_F(J) = \frac{P(F \cap J)}{P(F)} = \frac{0,1}{0,4} = 0,25$.

III. Arbre pondéré et calculs de probabilité

- La probabilité de l'événement correspondant à un chemin est obtenue en multipliant les probabilités portées par ses branches.
- La somme des probabilités sur les branches partant d'un même nœud est égale à 1.
- La probabilité de l'événement correspondant à plusieurs chemins est obtenue en additionnant les probabilités de chacun de ces chemins (**formule des probabilités totales**).

Le cas le plus fréquent correspond à la partition la plus simple (A et $\bar{A}$). Si on connaît les probabilités de B et $\bar{B}$ par l'intermédiaire de A et $\bar{A}$, on a l'arbre suivant :



- Le produit des probabilités inscrites sur chaque branche d'un chemin donne la probabilité de l'intersection des événements placés sur ce chemin.

$$P(A \cap B) = P(A) \times P_A(B)$$
- La somme des probabilités inscrites sur les branches issues d'un même nœud est égale à 1 (loi des nœuds).

$$P_A(B) + P_A(\bar{B}) = 1$$
- La probabilité d'un événement E est la somme des probabilités des chemins qui aboutissent à E .

$$P(B) = P(A \cap B) + P(\bar{A} \cap B)$$

$$= P(A) \times P_A(B) + P(\bar{A}) \times P_{\bar{A}}(B)$$

Exemple

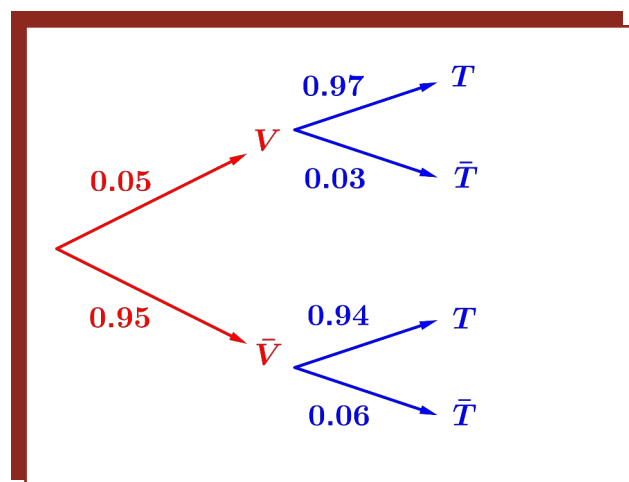
Dans un pays, il y a 5% de la population contaminée par le coronavirus. On dispose d'un test de dépistage de ce virus qui a les propriétés suivantes

- La probabilité qu'une personne contaminée ait un test positif est de 0,97 (sensibilité du test).
- La probabilité qu'une personne non contaminée ait un test négatif est de 0,94 (spécificité du test). On fait passer un test à une personne choisie au hasard dans cette population.

On note les événements suivants :

V : « la personne est contaminée par le virus »

T : « le test est positif »



- La probabilité que le test soit positif est :

$$P(T) = P(V \cap T) + P(\bar{V} \cap T)$$

$$= 0,05 \times 0,97 + 0,95 \times 0,94$$

$$= 0,9415$$

- La probabilité que la personne soit contaminée sachant que le test est positif est :

$$P_T(V) = \frac{P(V \cap T)}{P(T)}$$

$$= \frac{0,05 \times 0,97}{0,9415}$$

$$= 0,051514$$

IV. Indépendance

- Soit A et B deux événements d'un univers Ω .
On dit que A et B sont **indépendants** lorsque $P(A \cap B) = P(A) \times P(B)$.
- Soit A et B deux événements d'un univers Ω tels que $P(A) \neq 0$.
Les événements A et B sont **indépendants** si et seulement si $P_A(B) = P(B)$.

En réalisant successivement deux expériences aléatoires telles que les événements associés à la première soit indépendants des événements associés à la seconde, on dit que l'on réalise une **succession de deux épreuves indépendantes**.

- Deux expériences sont dites **indépendantes** si le résultat de l'une n'influe pas sur le résultat de l'autre.
- Deux expériences sont dites **identiques** si elles ont les mêmes issues et les mêmes probabilités pour chaque issue.

Dans une succession d'épreuves indépendantes, la probabilité d'une issue est le **produit des probabilités rencontrées sur le chemin** qui conduit à cette issue.

Répétition d'épreuves indépendantes et variables aléatoires

I. Modèle associé à une expérience aléatoire à deux épreuves indépendantes

En réalisant successivement deux expériences aléatoires telles que les événements associés à la première soit indépendants des événements associés à la seconde, on dit que l'on réalise une **succession de deux épreuves indépendantes**.

- Deux expériences sont dites **indépendantes** si le résultat de l'une n'influe pas sur le résultat de l'autre.
- Deux expériences sont dites **identiques** si elles ont les mêmes issues et les mêmes probabilités pour chaque issue.

II. Variables aléatoires

II.1 Variable aléatoire discrète

Soit une expérience aléatoire dont l'univers est l'ensemble Ω .

Une **variable aléatoire** est une fonction définie sur Ω et à valeurs dans $\mathbb{R}$. On la note X .

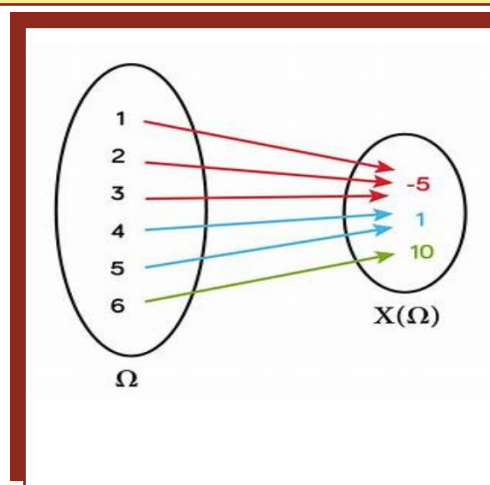
Exemple

On lance un dé équilibré à six faces et on observe le résultat affiché sur la face supérieure.

- « Si le résultat obtenu est 1, 2 ou 3, je perds 5 jetons »
- « Si le résultat obtenu est 4 ou 5, je gagne 1 jeton »
- « Sinon, je gagne 10 jetons »

On peut définir une variable aléatoire X qui décrit les gains de ce jeu. On a donc $X(1)=-5$, $X(2)=-5$, $X(3)=-5$, $X(4)=1$, $X(5)=1$, $X(6)=10$.

On peut représenter la situation par le schéma ci-contre.



II.2 Loi de probabilité d'une variable aléatoire discrète

Soit X une variable aléatoire définie sur l'univers Ω . L'ensemble des valeurs prises par X est $\Omega = \{x_1; x_2; \dots; x_k\}$, où les valeurs sont rangées par ordre croissant. Le nombre x_i est associé à une ou plusieurs issues de Ω , avec $1 \leq i \leq k$.

- L'événement « $X = x_i$ » est l'ensemble des issues de Ω auxquelles on a associé la valeur x_i .
- L'événement « $X \geq x_i$ » est l'ensemble des issues de Ω auxquelles on a associé une valeur supérieure ou égale à x_i . On définit de même « $X > x_i$ », « $X < x_i$ », et « $X \leq x_i$ ».

La probabilité de l'événement « $X = x_i$ » est la probabilité de l'événement formé de toutes les issues associées au nombre x_i . On la note $P(X = x_i)$ ou p_i .

Une variable aléatoire X est définie sur l'univers Ω d'une expérience aléatoire.

Notons $\Omega = \{x_1, x_2, \dots, x_n\}$ l'ensemble des valeurs prises par X .

La loi de probabilité de X est la fonction qui à chaque x_i de Ω lui associe sa probabilité notée $P(X=x_i)$. On la représente sous forme d'un tableau de valeurs :

x_i	x_1	x_2	...	x_n
$P(X=x_i)$	$P(X=x_1)$	$P(X=x_2)$	...	$P(X=x_n)$

On a $p_1 + p_2 + \dots + p_n = \sum_{i=1}^n p_i = 1$ ou $P(X=x_1) + P(X=x_2) + \dots + P(X=x_n) = \sum_{i=1}^n P(X=x_i) = 1$.

II.3 Espérance

L'**espérance** mathématique de la variable aléatoire X est le réel noté $E(X)$ défini par :

$$E(X) = \sum_{i=1}^n p_i x_i = p_1 x_1 + p_2 x_2 + \dots + p_n x_n$$

Remarques

- La loi des grands nombres nous permet d'interpréter l'espérance de la loi de probabilité de X . Elle nous dit en effet qu'en répétant un grand nombre de fois l'expérience, les fréquences observées se rapprochent de la probabilité théorique. En conséquence, la moyenne des résultats obtenus se rapproche de l'espérance de la loi de probabilité de X . L'espérance est donc la moyenne que l'on peut espérer en répétant l'expérience un grand nombre de fois.

Pour le jeu proposé en exemple, l'espérance de $-\frac{1}{2}$ signifie que l'on peut « espérer » ou

« craindre » perdre en moyenne $-\frac{1}{2}$ de jeton par partie (ou perdre 1 jeton toutes les 2 parties).

- Si X représente le gain d'un jeu d'argent, si on a $E(X)=0$, alors on dit que le jeu est équitable.

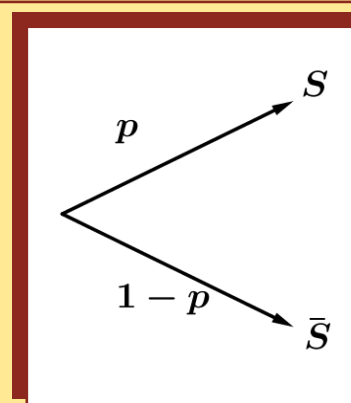
III. Épreuve, loi et schéma de Bernoulli

III.1 Épreuve de Bernoulli

Une **épreuve de Bernoulli** de paramètre p

(compris entre 0 et 1) est une expérience aléatoire comportant deux issues :

- l'une appelée **succès**, noté S , de probabilité p .
- l'autre appelée **échec**, noté $\bar{S}$ ou E , de probabilité $q=1-p$.



III.2 Loi de Bernoulli

Notons X la variable aléatoire prenant la valeur 1 en cas de succès et la valeur 0 en cas d'échec.

On dit que X suit la **loi de Bernoulli** de paramètre p .

Autrement dit, on a $P(X=1)=p$ et $P(X=0)=1-p$.

On peut résumer la loi de Bernoulli par le tableau suivant :

x_i	1	0
$P(X=x_i)$	p	$1-p$

Soit X une variable aléatoire qui suit la loi de Bernoulli de paramètre p .

L'**espérance** de X est $E(X)=p$.

III.3 Schéma de Bernoulli

Soit n un entier naturel et p un réel compris entre 0 et 1.

Un **schéma de Bernoulli** de paramètres n et p est une expérience aléatoire consistant à répéter n fois, de façon **identique et indépendante**, une épreuve de Bernoulli de paramètre p .

IV. Échantillonnage

IV.1 Simulation d'échantillonnage

Lorsqu'on travaille sur une population de grande taille, il est rarement possible d'avoir accès aux données relatives à l'ensemble de la population. On utilise alors un **échantillon** de cette population.

Un échantillon de taille n est une sélection de n individus choisis au hasard dans une population.

Lorsque l'on répète n fois, de façon indépendante, une expérience aléatoire à deux issues, on obtient une série de n résultats que l'on appelle échantillon de taille n associé à une épreuve de Bernoulli.

Remarque

Simuler un échantillon avec un tableur ou un programme en langage Python permet d'étudier des séries statistiques comportant un très grand nombre de données.

IV.2 Étude des échantillons

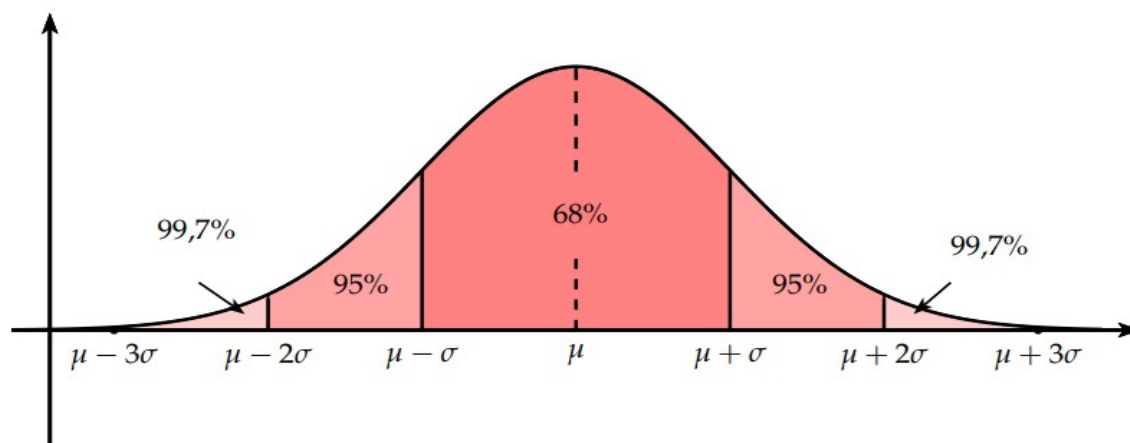
Sur plusieurs échantillons de même taille, la fréquence d'un caractère observée varie d'un échantillon à l'autre. C'est ce qu'on appelle la **fluctuation d'échantillonnage**.

Remarque

Plus la taille des échantillons est grande, plus le phénomène de fluctuation diminue. Les fréquences se rapprochent alors de la proportion théorique, c'est-à-dire de la probabilité théorique.

On étudie une simulation de N échantillons de taille n et on note σ l'écart type de la série des fréquences obtenues et μ la proportion théorique.

- En moyenne, environ 68% des fréquences sont dans l'intervalle $[\mu - \sigma; \mu + \sigma]$.
- En moyenne, environ 95% des fréquences sont dans l'intervalle $[\mu - 2\sigma; \mu + 2\sigma]$.
- En moyenne, plus de 99% des fréquences sont dans l'intervalle $[\mu - 3\sigma; \mu + 3\sigma]$.



L'écart-type de la série des fréquences est de l'ordre de $\frac{1}{2\sqrt{n}}$.

Autrement dit, la valeur de σ est proche de $\frac{1}{2\sqrt{n}}$.