Probabilités discrètes

1 Introduction

1.1 Vocabulaire et définitions

Univers : C'est l'ensemble des résultats possibles d'une expérience aléatoire. Ces résultats sont appelés événements élémentaires ou **éventualités** ou encore **issues**. On note en général l'univers U ou Ω.

Événement : C'est une partie de l'univers : elle correspond à une réunion d'événements élémentaires.

Si on répète un grand nombre de fois l'expérience, on admet que la fréquence d'apparition de l'événement élémentaire a se "stabilise" autour d'une certaine valeur, comprise entre 0 et 1 (par définition de la fréquence). Cette valeur est la **probabilité** de l'événement élémentaire a, notée en général $P(a)$.

La probabilité d'un événement A notée $P(A)$ est alors la somme des probabilités de tous les événements élémentaires constituant A.

Tout ceci constitue un énoncé vulgarisé de ce qu'on appelle la *loi des grands nombres* : Supposons qu'on reproduise n fois une expérience aléatoire. On admet que la fréquence de réalisation de l'événement A « tend » vers $P(A)$ lorsque n devient très grand.

1.2 Premières propriétés

1. La probabilité d'un événement est un nombre compris entre 0 et 1. La somme des probabilités de tous les événements élémentaires constituant U vaut 1 (car la somme des fréquences vaut toujours 1).
2. $P(\varnothing) = 0$: l'événement $\varnothing$ est impossible.
3. $P(U) = 1$: l'événement U est certain.
4. On dit qu'il y a équiprobabilité si tous les événements élémentaires ont même probabilité. Dans ce cas, on peut utiliser la formule :

$$\boxed{P(A) = \frac{\text{Nombre de cas favorables pour } A}{\text{Nombre de cas possibles}}}$$

2 Notions ensemblistes

2.1 Définitions

Soient A, B deux événements.

Réunion : $A \cup B$ est l'ensemble des éventualités qui appartiennent soit à A soit à B. Donc l'événement $A \cup B$ est réalisé lorsque A ou B est réalisé (ou les deux).

Intersection : $A \cap B$ est l'ensemble des éventualités qui appartiennent à la fois à A et à B. Donc l'événement $A \cap B$ est réalisé lorsque A et B sont réalisé tous les deux.

Incompatibilité : A et B sont incompatibles quand $A \cap B = \varnothing$ (pas d'éventualité appartenant à la fois à A et à B). Les événements A et B sont incompatibles lorsqu'ils ne peuvent pas être réalisés en même temps.

Contraire : L'événement contraire de A, noté $\overline{A}$ est l'ensemble de toutes les éventualités qui n'appartiennent pas à A. On a donc : $A \cap \overline{A} = \varnothing$ et $A \cup \overline{A} = U$.

Partition de l'univers : une partition de U est constituée par des ensembles incompatibles 2 à 2 et dont la réunion est U.

Autrement dit : A, B, C forment une partition de U si : $A \cup B \cup C = U$; $A \cap B = \varnothing$; $A \cap C = \varnothing$; $B \cap C = \varnothing$.

Exemple : A et $\overline{A}$ forment une partition de U. On définit de la même façon une partition d'un événement quelconque.

2.2 Propriétés

1. Si A et B sont *incompatibles*, alors $P(A \cup B) = P(A) + P(B)$.
   Conséquence : $P(A) + P(\overline{A}) = 1$
2. Cas général : quels que soient les événements A et B, $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$.
3. Principe des « probabilités totales » :
   Exprimons-le pour une partition constituée de trois événements ; la formule se transpose facilement au cas de deux, quatre, cinq, six événements ou plus.
   Si A, B, C forment une partition de U alors, pour tout événement E, $P(E) = P(E \cap A) + P(E \cap B) + P(E \cap C)$.
   Remarque : si E est l'univers, la formule devient : $P(A) + P(B) + P(C) = 1$.

3 Variable aléatoire discrète

3.1 Définition

On adopte la définition suivante (elle sera modifiée et précisée dans l'enseignement supérieur) :

Une **variable aléatoire discrète** est une variable qui ne prend qu'un nombre fini de valeurs x_1, x_2, ... , x_n, chacune avec une certaine probabilité. C'est l'équivalent d'un caractère quantitatif discret en statistiques. En particulier, on obtient une variable aléatoire en considérant les issues d'une expérience aléatoire dont les résultats sont numériques et en nombre fini (exemple : lancer d'un dé.)

On désigne habituellement une variable aléatoire avec une majuscule : X, Y, etc. La probabilité que X prenne la valeur x_i se note $P(X = x_i)$ (ou parfois $P(x_i)$ lorsqu'il n'y a pas d'ambiguïté.)

La **loi de probabilité** d'une v.a. X est l'application qui à toute valeur x_i prise par X associe la probabilité $P(X = x_i)$. On la représente habituellement à l'aide d'un tableau.

Valeur prise	x_1	x_2	...
Probabilité de cette valeur	$P(X = x_1)$	$P(X = x_2)$	...

3.2 Espérance d'une variable aléatoire

C'est l'équivalent de la moyenne en statistiques, les fréquences étant remplacées par les probabilités.

Soit X une variable aléatoire prenant les valeurs x_1, x_2, ... , x_n. **L'espérance** de X, notée $E(X)$ est le nombre :

$$E(X) = x_1 \times P(X = x_1) + x_2 \times P(X = x_2) + \cdots + x_n \times P(X = x_n)$$

Autrement dit, on multiplie chaque valeur par sa probabilité et on fait la somme.

Pour alléger les notations, posons : $p_1 = P(X = x_1)$, $p_2 = P(X = x_2)$, etc. On obtient :

$$\boxed{E(X) = p_1 \times x_1 + p_2 \times x_2 + \cdots + p_n \times x_n}$$

3.3 Variance et écart-type

Soit X une variable aléatoire prenant les valeurs x_1, x_2, ... , x_n. On pose comme précédemment : $p_1 = P(X = x_1)$, $p_2 = P(X = x_2)$, etc.

La variance de X, notée $V(X)$ est le nombre :

$$\boxed{V(X) = p_1 \times (x_1 - E(X))^2 + p_2 \times (x_2 - E(X))^2 + \cdots + p_n \times (x_n - E(X))^2}$$

Remarque : $V(X)$ est un nombre positif. En effet, les carrés sont positifs et les probabilités p_i sont des nombres positifs.

L'écart-type de X, noté $\sigma(X)$ est la racine carrée de la variance :

$$\boxed{\sigma(X) = \sqrt{V(X)}}$$

3.4 Propriétés

Soient a et b deux nombres réels. On note $aX + b$ la variable aléatoire qui prend les valeurs $ax_1 + b$, $ax_2 + b$, ... , $ax_n + b$ avec les probabilités respectives p_1, p_2, ... , p_n.

On a alors :

$$E(aX + b) = aE(X) + b$$
$$V(aX + b) = a^2V(X)$$

4 Répétition d'expériences identiques et indépendantes

4.1 Définition

On peut constituer une nouvelle expérience aléatoire en reproduisant plusieurs fois une expérience aléatoire donnée, en gardant la même procédure. On parle alors de **répétition d'expériences identiques**.

Par exemple : on répète plusieurs fois le lancer d'un dé.

Si le résultat de l'une quelconque de ces expérience ne dépend pas du résultat des autres expériences, on dit qu'elles sont indépendantes les unes des autres et on parle alors de répétitions d'expériences identiques **et indépendantes**.

Par exemple : Si on lance deux fois un dé, le résultat obtenu lors du premier lancer n'a pas d'influence sur le résultat du deuxième lancer.

Si on tire au hasard deux boules successivement dans une urne contenant une boule blanche et deux boules noires, il y a indépendance si on remet la première boule tirée dans l'urne. Par contre, si on ne remet pas dans l'urne la première boule tirée, il n'y a pas indépendance. En effet, la probabilité d'obtenir une boule blanche lors du premier tirage est $\frac{1}{3}$; mais si la boule blanche a été tirée en premier, la probabilité de tirer à nouveau une boule blanche est de 0 (car alors l'urne ne contient plus que les deux boules noires.)

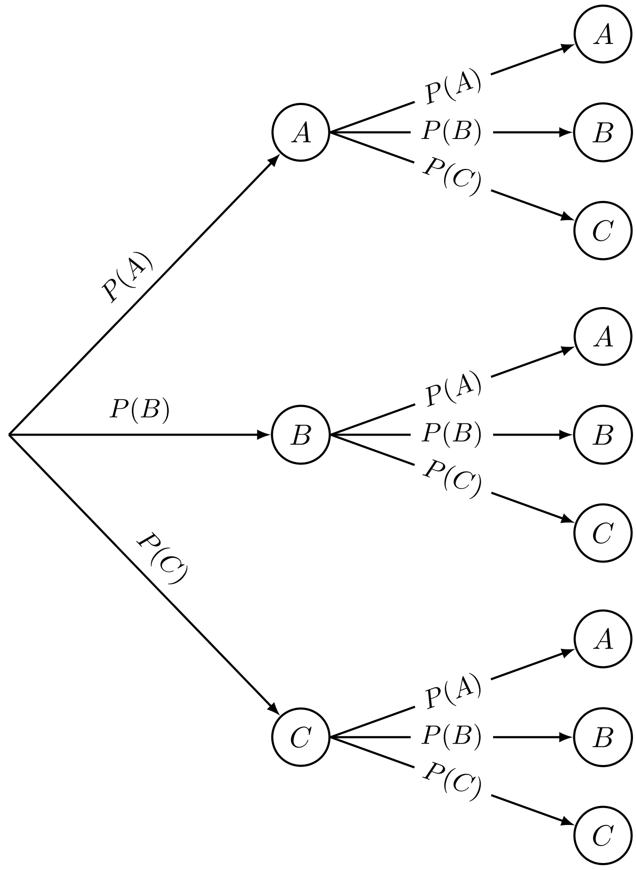
4.2 Propriété

Dans le cas de la répétition d'expériences identiques et indépendantes, on admet que la probabilité d'une liste (ordonnée) de résultats est le produit des probabilités de chaque résultat.

4.3 Arbres pondérés

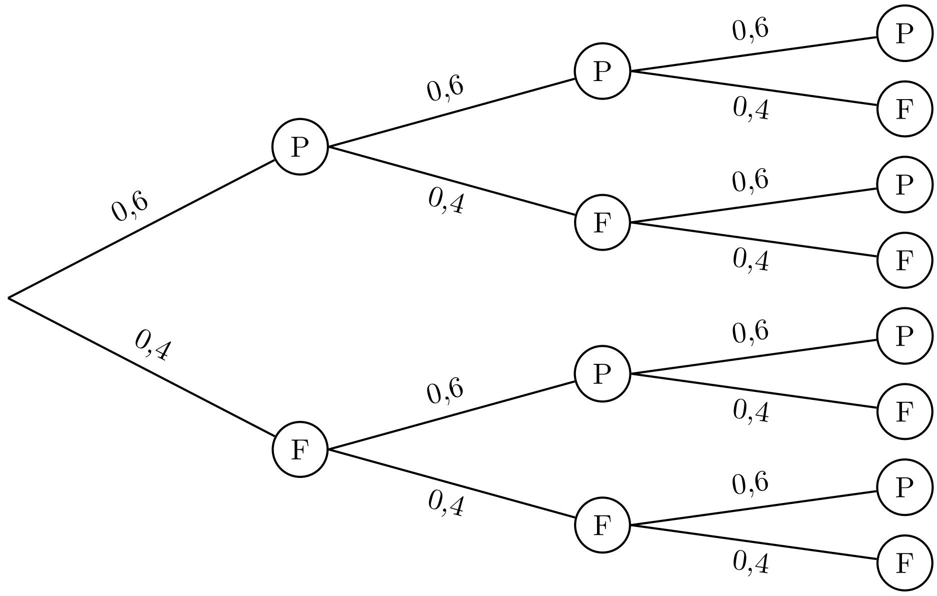
On peut représenter la situation précédente par un arbre pondéré. Chaque branche est affecté d'un coefficient égal à la probabilité de l'événement auquel elle aboutit. La somme des coefficients situés sur les branches qui partent d'un nœud vaut 1.

Exemple : Répétition de deux expériences à trois issues A, B, C, de façon identique et indépendante :



L'intérêt de l'arbre est le suivant : une liste de résultats correspond à un parcours sur l'arbre, partant du nœud principal à gauche pour aboutir à un des nœuds terminaux. La probabilité de cette liste est alors le produit des probabilités rencontrées sur chacune des branches du parcours.

Exemple numérique : On lance une pièce truquée. La probabilité de tomber sur pile est de 0,6 et la probabilité de tomber sur face est 0,4.



On peut calculer, par exemple : $P(\mathrm{P,F,P}) = 0{,}6 \times 0{,}4 \times 0{,}6 = 0{,}144$.

5 Loi binomiale

5.1 Définition

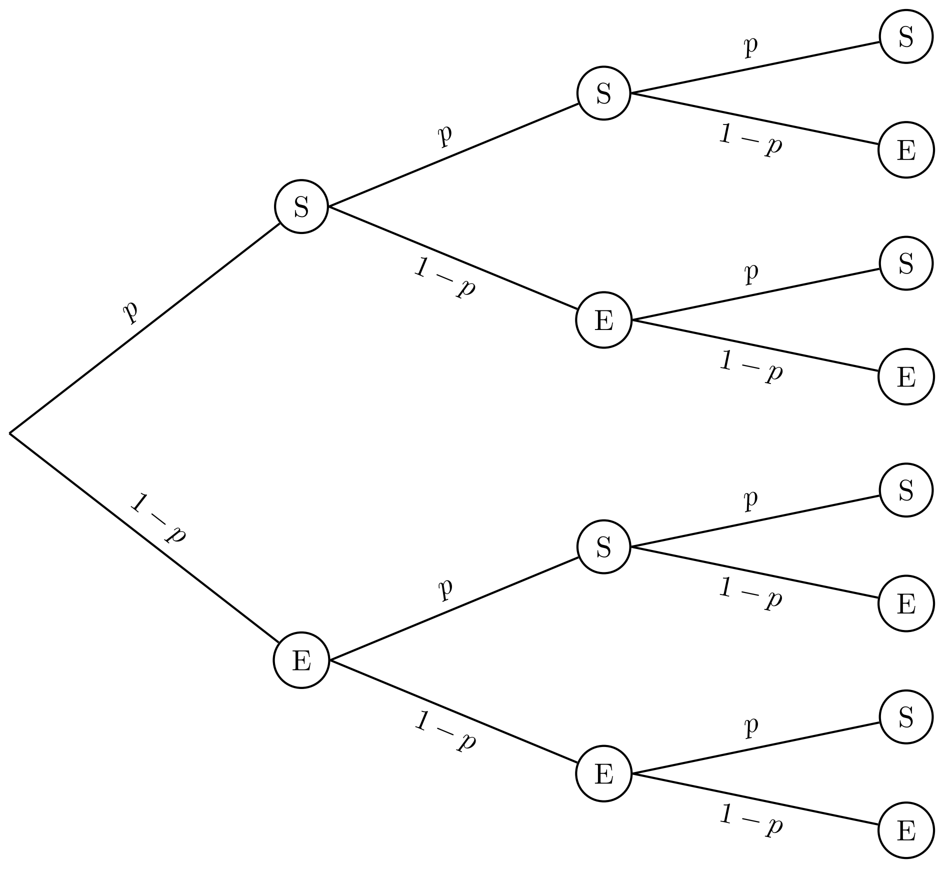
5.1.1 Schéma de Bernoulli

Une **épreuve de Bernoulli** est une expérience aléatoire qui n'a que deux issues (l'univers n'a donc que deux éventualités). On note traditionnellement ces deux issues S (succès) et E (échec). On note $p = P(S)$ la probabilité d'un succès.

Exemples : on lance une pièce ; on tire une boule dans une urne contenant des boules de deux couleurs, etc.

Un **schéma de Bernoulli** est une expérience qui est la répétition de plusieurs épreuves de Bernoulli *identiques* et *indépendantes*. On note n le nombre d'épreuves répétées.

Illustration avec $n = 3$:



Exemples :

1. On lance 20 fois de suite la même pièce.
2. On tire 10 fois de suite une boule dans une urne contenant des boules de deux couleurs et on la remet à chaque fois dans l'urne (tirage avec remise).

Remarque : on peut schématiser une issue par une liste ordonnée de n lettres ≪ S ≫ ou ≪ E ≫.

5.1.2 Loi binomiale

On réalise un schéma de Bernoulli. Soit X la variable aléatoire qui compte le nombre de succès obtenus. Les valeurs prises par X sont donc tous les entiers de 0 à n. On appelle **loi binomiale** de paramètres n et p, et on note $\mathcal{B}(n; p)$ la loi de probabilité de X.

Elle associe à tout entier $k \in [0 ; 1 ; ... ; n]$ la probabilité d'obtenir k succès. Cette probabilité dépend du nombre n de répétitions et de la probabilité p d'un succès.

Exemple : on jette une pièce trois fois de suite. Quelle est la probabilité d'obtenir 2 fois pile ?

Il y a trois manière d'obtenir 2 piles : PPF, PFP et FPP. La probabilité de chacun de ces événements est : $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$ puisque les épreuves sont indépendantes. La probabilité d'obtenir 2 piles est donc de $\frac{3}{8}$.

5.2 Coefficients binomiaux

5.2.1 Définition

On construit l'arbre pondéré associé à un schéma de Bernoulli de paramètre n et p (rappel : n est le nombre de répétitions et p la probabilité d'un succès au cours d'une épreuve de Bernoulli.) Le nombre de chemins correspondant à k succès est appelé **coefficient binomial** « k parmi n », et noté $\begin{pmatrix} n \\ k \end{pmatrix}$.

Remarque : De façon équivalente, on peut dire que parmi les mots de n lettres qu'on peut fabriquer en utilisant seulement les lettres « S » et « E », $\binom{n}{k}$ est le nombre de mots comportant exactement k fois la lettre « S ».

5.2.2 Propriétés

pour tout $n \in \mathbb{N}$, pour tout k entier entre 0 et n :

1. $\binom{n}{0} = \binom{n}{n} = 1$
2. $\binom{n}{1} = n$
3. $\binom{n}{k} = \binom{n}{n-k}$
4. $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$

5.2.3 Application : Le triangle de Pascal

On peut calculer les $\binom{n}{k}$ de proche en proche à l'aide du tableau ci-contre :

- on place des 1 sur la première colonne et la diagonale ;
- on obtient un autre nombre du tableau en ajoutant le nombre juste au-dessus et celui situé à gauche sur la ligne précédente.

n \ k	0	1	2	3	4	5
0	1					
1	1	1				
2	1	2	1			
3	1	3	3	1		
4	1	4	6	4	1	
5	1	5	10	10	5	1

5.3 Formule générale de la loi binomiale

Soit X une v.a. suivant la loi binomiale $\mathcal{B}(n;p)$ de paramètres n et p. Pour tout entier $k \in [0;n]$, on a :

$$\boxed{P(X = k) = \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k}}$$

5.4 Espérance, variance et écart-type

Soit X une v.a. suivant la loi binomiale $\mathcal{B}(n;p)$ de paramètres n et p.

L'espérance de X est : $\boxed{E(X) = np}$.

La variance de X est : $\boxed{V(X) = np(1-p)}$ et donc l'écart-type de X est : $\boxed{\sigma(X) = \sqrt{np(1-p)}}$.

6 Échantillonnage et prise de décision en lien avec la loi binomiale

6.1 Introduction

On étudie, dans une population, une caractéristique A. On fait l'hypothèse que la proportion d'individus ayant cette caractéristique est p. Pour tester cette hypothèse, on réalise un échantillon de taille n en choisissant au hasard, de façon indépendante, n individus.

On calcule f, la fréquence de la caractéristique A dans l'échantillon (c'est à dire la proportion d'individus ayant cette caractéristique.) Si f est suffisamment proche de p, on va accepter l'hypothèse, et dans le cas contraire, la rejeter. Le but est de donner un critère permettant de préciser cette idée de "suffisamment proche".

Si on choisit un individu au hasard, la probabilité qu'il ait la caractéristique A est p. Le choix de n individus au hasard et indépendamment est un schéma de Bernoulli, donc la variable aléatoire X comptant le nombre d'individus ayant la caractéristique A dans l'échantillon suit la loi binomiale $\mathcal{B}(n; p)$.

6.2 Intervalle de fluctuation au seuil 95%

L'objectif est de déterminer l'intervalle $[a\,;b]$ le plus petit possible tel que $P(a \leq X \leq b) \geq 0{,}95$. On choisit a et b de la façon suivante :

- a est le plus petit entier tel que $P(X \leq a) > 0{,}025$
- b est le plus petit entier tel que $P(X \leq b) \geq 0{,}975$.

L'intervalle de fluctuation à 95% est $I = \left[\frac{a}{n}\,;\frac{b}{n}\right]$. On remarque que I est à peu près centré en p.

Si on constitue un grand nombre d'échantillons, on va observer en général que 95 fois sur 100 environ, le nombre d'individu ayant la caractéristique A dans l'échantillon appartient à $[a\,;b]$, autrement dit que f appartient à I.

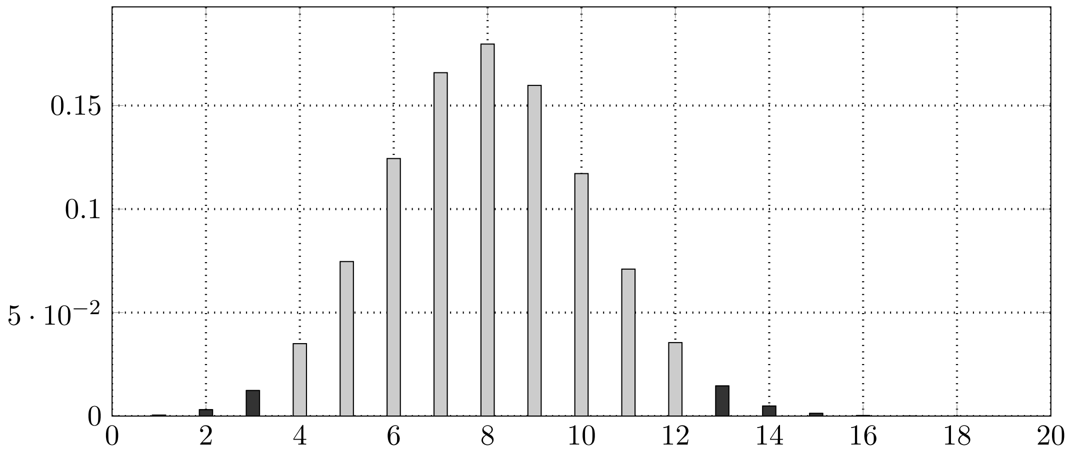
6.3 Règle de décision

- Si $f \in I$, alors on accepte l'hypothèse.
- Si $f \notin I$ alors on rejette l'hypothèse, au risque de 5% de la rejeter à tort.

6.4 Exemples

Les diagrammes ci-dessous représentent les probabilités de chaque valeur.

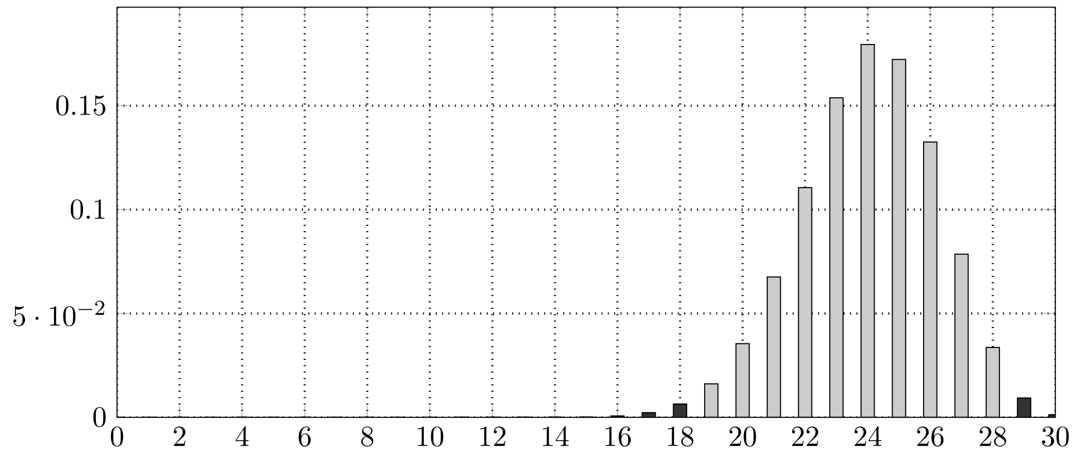
1. n=20 ; p=0,4



On trouve $a = 4$ et $b = 12$.

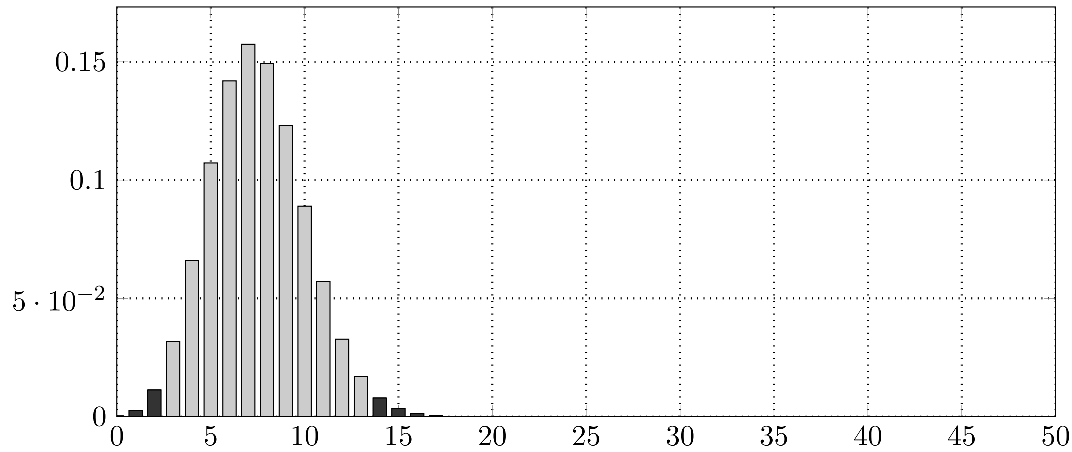
Ainsi, l'intervalle de fluctuation au seuil 95% est : $I = [0{,}2\,; 0{,}6]$.

2. n=30 ; p=0,8



On trouve $a = 19$ et $b = 28$. Ainsi, $I \simeq [0{,}63\ ;0{,}93]$.

3. n=50 ; p=0,15



On trouve $a = 3$ et $b = 13$. Ainsi, $I = [0{,}06\ ;0{,}26]$.

6.5 Intervalle de fluctuation à un autre seuil

Même principe. On cherche à déterminer l'intervalle $[a\ ;b]$ le plus petit possible tel que $P(a \leq X \leq b) \geq s$, où s est donné. Par exemple, $s = 99\%$ ou $s = 90\%$.

7 Probabilités conditionelles

7.1 Définition

Soit P une probabilité sur l'univers U. Soit A un événement possible : $P(A) \neq 0$. On définit une nouvelle probabilité P_A sur U en posant, pour tout événement B : $\boxed{P_A(B) = \dfrac{P(B \cap A)}{P(A)}}$.

$P_A(B)$ se lit « probabilité de B sachant A ». On l'interprète ainsi : c'est la probabilité que B soit réalisé sachant que A est réalisé.

Notons que ceci implique : $P(B \cap A) = P_A(B) \times P(A)$.

Exemple 1 : On lance un dé. Quelle est la probabilité d'obtenir le 2 sachant qu'on obtient un nombre pair ?

Posons : A = « on obtient un nombre pair » et B = « on obtient le 2 ».

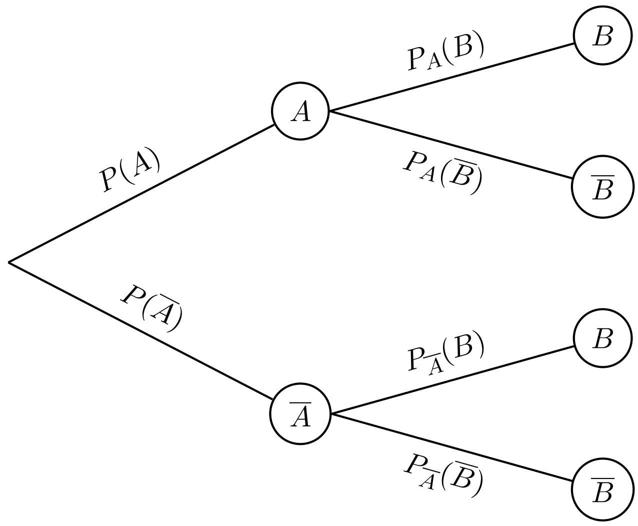
On calcule facilement : $P(A) = 3/6 = 1/2$; $P(A \cap B) = 1/6$ donc $P_A(B) = (1/6) \div (1/2) = 1/3$. C'est conforme à l'intuition puisqu'il y a 3 façons d'obtenir un nombre pair.

Exemple 2 : Dans les mêmes conditions, quelle est la probabilité d'obtenir le 3 sachant qu'on obtient un nombre pair ?

Posons : C = « obtenir le 3 ». Alors $P(A \cap C) = 0$, donc $P_A(C) = 0$. C'est logique, puisque l'événement est impossible.

7.2 Arbre de probabilités

On représente les probabilités conditionnelles sur les branches du deuxième niveau, ainsi :



On trouve par exemple : $P(A \cap B) = P(A) \times P_A(B)$ en suivant les branches du haut.

7.3 Application du principe des probabilités totales

On l'écrit pour une partition de trois ensembles ; ça se généralise à une partition formée d'un nombre quelconque d'ensembles.

Soient A, B, C trois événements différents de $\varnothing$. Si A, B, C forment une partition de U alors pour tout événement E :

$$P(E) = P(E \cap A) + P(E \cap B) + P(E \cap C)$$

Par conséquent :

$$P(E) = P_A(E) \times P(A) + P_B(E) \times P(B) + P_C(E) \times P(C)$$

7.4 Indépendance

Définition :

Soient A et B deux événements. On dit que A et B sont indépendants si $\boxed{P(A \cap B) = P(A) \times P(B)}$.

Si A et B sont différents de $\varnothing$, cette définition est équivalente à :

- $P_A(B) = P(B)$
- $P_B(A) = P(A)$

Intuitivement, le fait qu'un événement soit réalisé n'a aucune influence sur la probabilité de réalisation de l'autre.

Table des matières