Université Montpellier II Année **2004-2005**
Mathématiques

DÉNOMBREMENT ET PROBABILITÉS FINIES

Jérôme DRONIOU, Daniel GUIN

---

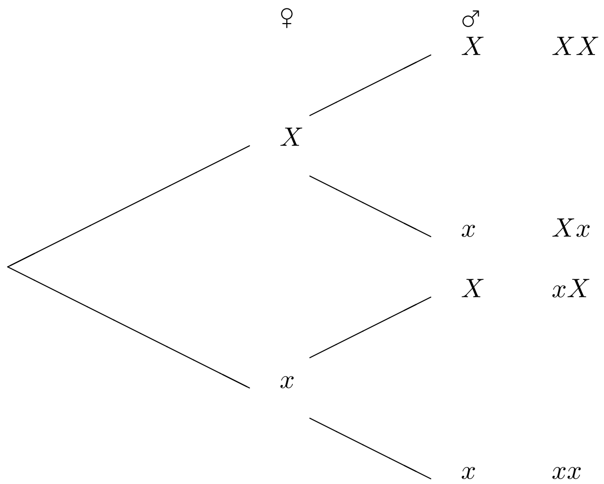
1. Introduction

Une expérience aléatoire $\mathcal{E}$, ou épreuve, est une expérience qui, répétée dans des conditions (apparemment) identiques, peut conduire à des résultats différents. L'ensemble de toutes les issues possibles de $\mathcal{E}$, noté Ω, s'appelle l'ensemble fondamental associé à $\mathcal{E}$. Les éléments de $\mathcal{E}$ s'appellent les évènements élémentaires. Un évènement lié à $\mathcal{E}$ (qui sera réalisé ou non suivant le résultat de l'expérience) correspond à un sous-ensemble de Ω. Le sous-ensemble vide correspond à un évènement impossible et Ω correspond à un évènement certain.

Exemples

1. Si $\mathcal{E}$ est "jeter un dé", alors $\Omega = \{1, 2, 3, 4, 5, 6\}$; "obtenir un six" est un évènement élémentaire ; "obtenir un nombre pair" est un évènement qui correspond au sous-ensemble $\{2, 4, 6\}$; "obtenir un nombre compris entre 1 et 6" est un évènement certain.
2. On considère une population dans laquelle on distingue un caractère déterminé par deux allèles X, x. L'épreuve $\mathcal{E}$ consiste en un croisement tel que chaque partenaire (mâle et femelle) donne un seul des éléments X ou x, qui se combine avec l'élément donné par l'autre partenaire. Cette combinaison est indépendante des éléments en jeu. On a alors
$$\Omega = \{XX, Xx, xx\}.$$

   On peut utiliser une représentation en arbre.



Le dictionnaire entre le langage des évènements et celui des ensembles est donné par le tableau ci-dessous

Langage des ensembles	Langage des évènements
Ω	issues de l'épreuve
$X \in \wp(\Omega)$	évènement
éléments de Ω	évènements élémentaires
$\emptyset$	évènement impossible
$\overline{X} = C_\Omega$ X	évènement contraire à X
$Z = X \cap Y$	X et Y : Z est réalisé si et seulement si X et Y sont réalisés
$Z = X \cup Y$	X ou Y : Z est réalisé si et seulement si l'un quelconque des évènements X ou Y est réalisé (ou les deux)
$X \cap Y = \emptyset$	évènements incompatibles
$X \subset Y$	si X réalisé, alors Y l'est.

2. Probabilités

Dans tout ce qui suit, on suppose que l'ensemble fondamental Ω est un ensemble fini. Le résultat d'une expérience aléatoire étant incertain, on souhaite prévoir quelle est la "chance" qu'un évènement se produise. Dans l'exemple 2 ci-dessus, les configurations possibles sont

♀\♂	X	x
X	XX	Xx
x	xX	xx

On note A le génotype XX, B le génotype Xx ou xX et C le génotype xx. Intuitivement, ce tableau suggère qu'il y a 1 chance sur 4 d'obtenir A, 2 chances sur 4 d'obtenir B, 1 chance sur 4 d'obtenir C.

Définition 2.1. — Soient $\mathcal{E}$ une expérience aléatoire et Ω l'ensemble fondamental associé. Une probabilité P sur Ω est une fonction $P : \wp(\Omega) \longrightarrow \mathbb{R}^+$ telle que

1. $P(\Omega) = 1$.
2. $\forall\ X \subset \Omega$, $\forall\, Y \subset \Omega$, $[X \cap Y = \emptyset] \Longrightarrow [P(X \cup Y) = P(X) + P(Y)]$.

Proposition 2.2. — *La fonction P vérifie les propriétés suivantes :*

1. $\forall X \subset \Omega$, $\forall Y \subset \Omega$
$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$$
2. $\forall X \subset \Omega$, $\forall Y \subset \Omega$,
$$[X \subseteq Y] \Longrightarrow [P(X) \leq P(Y)]$$
3. $\forall X \subset \Omega$, $P(X) \leq 1$
4. $P(X) + P(\overline{X}) = 1$
5. $P(\emptyset) = 0$
6. *Si $(X_i)_{i=1,\dots,n}$, est une famille d'évènements incompatibles deux à deux, on a :*
$$P(\bigcup_{i=1}^{n} X_i) = \sum_{i=1}^{n} P(X_i).$$
7. $P(X) = P(X \cap Y) + P(X \cap \overline{Y}), \forall Y \subset \Omega$.

Démonstration :

1. On a $X \cap (Y \setminus (X \cap Y)) = \emptyset$ et $X \cup (Y \setminus (X \cap Y)) = X \cup Y$. D'où $P(X \cup Y) = P(X) + P(Y \setminus (X \cap Y))$. Mais $Y = (Y \setminus (X \cap Y)) \cup (Y \setminus X \cap Y)$, et, comme $(Y \setminus (X \cap Y)) \cap (X \cap Y) = \emptyset$, on a $P(Y) = P(Y \setminus (X \cap Y)) + P(X \cap Y)$. On en déduit que $P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$.
2. On a $Y = X \cup (Y \cap \overline{X})$: comme $X \cap (Y \cap \overline{X}) = \emptyset$, on a $P(Y) = P(X) + P(Y \cap \overline{X})$. Mais la fonction P étant à valeurs dans $\mathbb{R}^+$, $P(Y \cap \overline{X}) \geq 0$. D'où $P(X) \leq P(Y)$.
3. Toute partie X de Ω est telle que $X \subset \Omega$, donc, d'après 2., $P(X) \leq P(\Omega) = 1$.
4. On a $X \cap \overline{X} = \emptyset$ et $X \cup \overline{X} = \Omega$, d'où $P(X) + P(\overline{X}) = P(X \cup \overline{X}) = P(\Omega) = 1$.
5. On en déduit que $P(\emptyset) = 0$.
6. Se déduit de l'axiome 2, de "proche en proche".
7. Conséquence immédiate de l'axiome 2.

On déduit de la définition et de ces propriétés qu'une probabilité est une fonction définie sur $\wp(\Omega)$ et à valeurs dans l'intervalle $[0, 1]$.

Probabilités uniformes.

Définition 2.3. — La probabilité P sur Ω est dite <u>uniforme</u> si tous les évènements élémentaires sont équiprobables (*i.e.* ont la même probabilité).

Dans ce cas, si Card $\Omega = n$, d'après la propriété 6, chaque évènement élémentaire a une probabilité égale à $\frac{1}{n}$. On en déduit, encore à l'aide de la propriété 6, que pour tout évènement $X \subset \Omega$, on a

$$P(X) = \frac{\text{Card } X}{\text{Card } \Omega}.$$

On est donc amené à calculer des cardinaux d'ensembles.

3. Dénombrements

1. Arrangements sans répétitions.

   On forme les suites (ordonnées) constituées de p éléments distincts pris dans un ensemble à n éléments, $p \leq n$. Le nombre de ces suites est

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

   Convention : $0\,! = 1$

2. Arrangements avec répétitions. C'est le même processus que ci-dessus, mais on ne suppose plus que les éléments de la suite sont distincts - on peut répéter plusieurs fois le même élément. Le nombre de telles suites est alors n^p.

3. Permutations. Dans le cas des arrangements sans répétitions, si on suppose que $p = n$, on trouve $A_n^n = n\,!$. C'est le nombre de bijections d'un ensemble à n éléments dans lui-même. Une telle bijection s'appelle une permutation.

4. Combinaisons. On forme les sous-ensembles à p éléments dans un ensemble à n éléments, $p \leq n$. La différence avec les arrangements sans répétitions est qu'ici l'ordre n'a pas d'importance. Autrement dit, deux suites à p éléments formées des mêmes éléments placés dans deux ordres différents constituent un même sous-ensemble à p éléments. Le nombre des sous-ensembles à p éléments dans un ensemble à n éléments, $p \leq n$, est

   $$\binom{n}{p} = \frac{n\,!}{p\,!\,(n-p)\,!}$$

   Rappel : formule du binôme. Pour tous nombres complexes a et b, et tout nombre entier positif n, on a

   $$(a+b)^n = \sum_{p=0}^{n} \binom{n}{p} a^p\, b^{n-p}$$

Proposition 3.1. — *Pour tous nombres entiers strictement positifs n et p, $p \leq n-1$, on a*

$$\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}.$$

Exemples :

1°) Pour un dé (non truqué) toutes les faces sont équiprobables. On peut donc appliquer la formule de la probabilité uniforme. Par exemple la probabilité d'obtenir, après un lancé, un nombre pair est égale à $\frac{1}{2}$, car

$$X = \{2, 4, 6\}, \qquad \Omega = \{1, 2, 3, 4, 5, 6\},$$

$$P(X) = \frac{\text{Card } X}{\text{Card } \Omega} = \frac{3}{6} = \frac{1}{2}.$$

2°) Dans l'exemple 2, les génotypes Xx et xX sont indiscernables. On a $\Omega = \{XX, Xx, xx\}$ et le diagramme page 1 montre que $P(XX) = \frac{1}{4}$, $P(Xx) = \frac{1}{2}$, et $P(xx) = \frac{1}{4}$. La probabilité P n'est pas uniforme.

Supposons que l'on soit dans le cas de "dominance stricte", *i.e.* X domine x, alors, dans la population correspondante, on distingue deux types d'individus.

- A : phénotype dominant,
- B : phénotype récessif.

Alors,

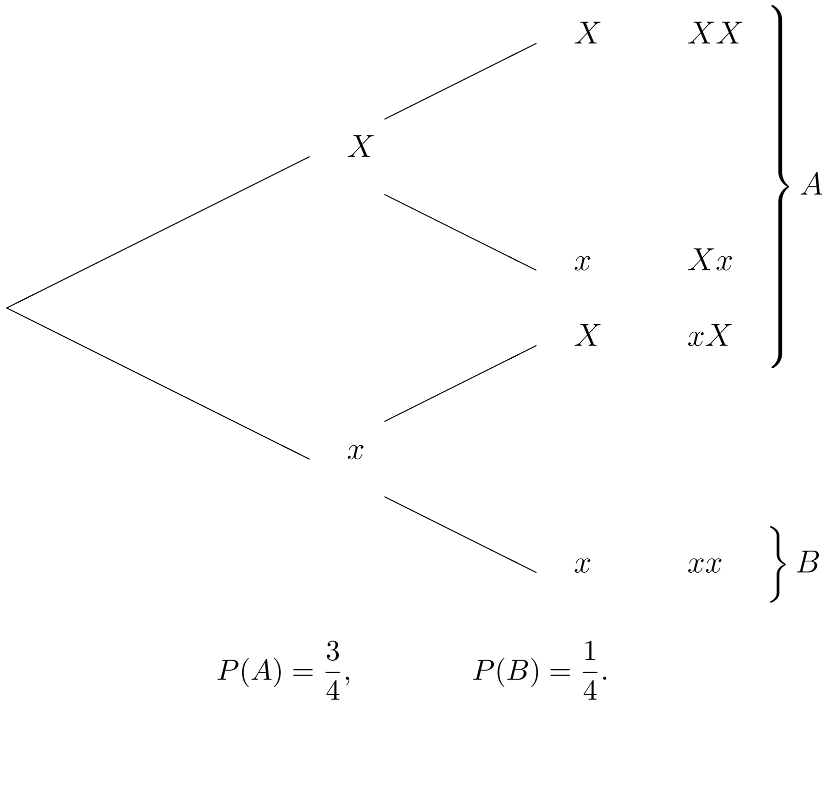
$$P(B) = P(xx) = \frac{1}{4}$$

$$P(A) = P(XX) + P(Xx) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

(Les évènements XX et Xx sont incompatibles). Dans ce cas, on peut se ramener à une probabilité uniforme. On considère

$$\Omega' = \{XX, Xx, xX, xx\},$$

alors $\qquad P(XX) = P(Xx) = P(xX) = P(xx) = \frac{1}{4}$,



$$P(A) = \frac{3}{4}, \qquad P(B) = \frac{1}{4}.$$

4. Probabilités conditionnelles

On considère une épreuve $\mathcal{E}$ et deux évènements X et Y associés à $\mathcal{E}$. On suppose qu'ils ne sont pas indépendants, *i.e.* la réalisation de l'un de ces évènements, par exemple Y, modifie la probabilité de l'autre, X.

Définition 4.1. — On pose

$$P_Y(X) = \frac{P(X \cap Y)}{P(Y)},$$

qu'on appelle probabilité conditionnelle de X sachant Y, ou probabilité conditionnelle de X sachant que Y est réalisé.

Au lieu de la notation $P_Y(X)$, on peut aussi utiliser la notation $P(X|Y)$.

Exemple :
On considère la situation précédente : A : phénotype dominant, B : phénotype récessif.

Génotypes :	XX	Xx	xx
Probabilités :	1/4	1/2	1/4

On cherche la probabilité d'obtenir, lors d'un croisement, un génotype homozygote, évènement H, sachant que l'on observe un phénotype A.

$$P(H \cap A) = P(XX) = \frac{1}{4}, \;\; P(A) = \frac{3}{4}$$

$$P_A(H) = \frac{P(H \cap A)}{P(A)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Proposition 4.2. —

1. *Si* $Y \subset X$, *alors* $P(X \cap Y) = P(Y)$, *d'où* $P_Y(X) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(Y)}{P(Y)} = 1$,
2. $P(Y)P_Y(X) = P(X)P_X(Y) = P(X \cap Y)$,
3. $P_Y(X) + P_Y(\overline{X}) = 1$,
4. $P(X) = P(Y)P_Y(X) + P(\overline{Y})P_{\overline{Y}}(X)$.

Cette dernière formule est appelée "formule des probabilités totales".

Evènements indépendants.

On suppose maintenant que la réalisation d'un des évènements X ou Y ne modifie pas la probabilité de l'autre.

Définition 4.3. — Les évènements X et Y sont indépendants si $P(X \cap Y) = P(X)P(Y)$.

On déduit de ce qui précède que deux évènements X et Y sont indépendants si et seulement si $P_Y(X) = P(X), P_X(Y) = P(Y)$.

Attention : deux évènements incompatibles de probabilités non nulles ne sont pas indépendants : $X \cap Y = \emptyset \implies 0 = P(X \cap Y) = P(X)P_X(Y)$, alors que $P(X)P(Y) \neq 0$.

Théorème de Bayes. Ce théorème permet de déterminer la probabilité pour qu'un évènement qui est supposé déjà réalisé soit dû à une certaine cause plutôt qu'une autre.

Théorème 4.4. — *Soient X et Y deux évènements associés à une épreuve. Alors*

$$P_Y(X) = \frac{P(X)P_X(Y)}{P(X)P_X(Y) + P(\overline{X})P_{\overline{X}}(Y)}.$$

Généralisation : Soient $X_1, \cdots, X_n, Y$ des évènements associés à une épreuve, tels que :

i) les X_i, $i = 1, \cdots, n$ sont incompatibles deux à deux,

ii) $\forall i$, $i = 1, \cdots, n$, $\quad P(X_i) \neq 0$,

iii) $Y = \bigcup_{i=1}^{n} X_i$.

alors pour tout k, $k = 1, \cdots, n$, on a

$$P_Y(X_k) = \frac{P(X_k)P_{X_k}(Y)}{\Sigma_{i=1}^{n}(P(X_i)P_{X_i}(Y))}.$$

Exemple :

On considère une population dont 1 individu sur 100 est atteint d'une maladie. On a mis au point un test qui permet de détecter cette maladie avec 8 chances sur 10 de reconnaître un malade (test positif) et 9 chances sur 10 de reconnaître un sujet sain (test négatif). On cherche la probabilité qu'un sujet ayant un test positif soit malade. On considère les évènements suivants : T = test positif, M = sujet malade, S = sujet sain (avec M = $\overline{S}$). On a

$$P(M) = \frac{1}{100}, \qquad P_M(T) = \frac{8}{10},$$
$$P(S) = \frac{99}{100}, \qquad P_S(T) = \frac{1}{10}.$$

D'où

$$\begin{aligned} P_T(M) &= \frac{(P(M)P_M(T)}{P(M)P_M(T) + P(S)P_S(T)} \\ &= \frac{0,01 \times 0,8}{0,01 \times 0,8 + 0,99 \times 0,1} \\ &\simeq 0,075. \end{aligned}$$

5. Compléments

1. **Loi binomiale.**

Une alternative est une épreuve à deux issues, que l'on désignera sous le nom de succès, noté S, ou échec, noté E. On suppose que l'évènement S a une probabilité $P(S) = p$. D'où $P(E) = q = 1 - p$. On suppose que l'on répète cette épreuve n fois. Par exemple tirage de boules avec remise ou jeu de "pile ou face". Alors, si on note E_k l'évènement "obtenir k succès" ($0 \leq k \leq n$), au cours de n répétitions de l'épreuve, on a

$$P(E_k) = \binom{n}{k} p^k \, q^{n-k}.$$

D'où le nom de loi binomiale, notée $\mathcal{B}(n, p)$.

Exemple :

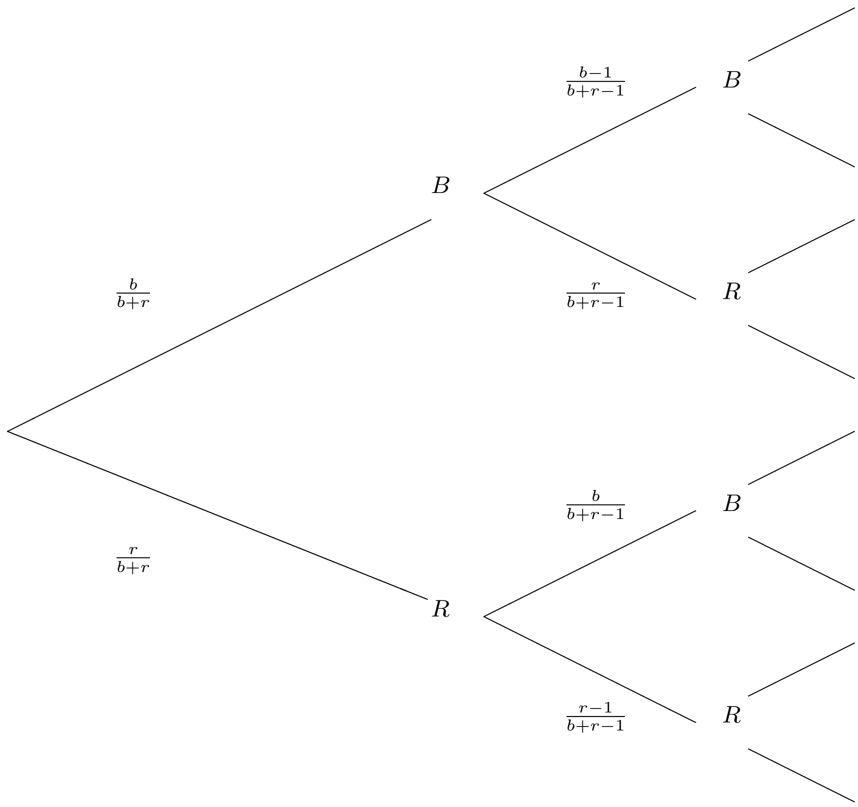
On suppose qu'au jeu de pile ou face, chaque issue est équiprobable, de probabilité 1/2. La probabilité d'obtenir 3 "face" en 5 lancers est

$$P(3) = \binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{5}{16}.$$

2. **Loi hypergéométrique**.

Une urne contient b boules blanches et r boules rouges. L'épreuve consiste à tirer une boule sans la remettre dans l'urne. Quelle est la probabilité de l'évènement E_k : "obtenir k boules blanches dans une série de n tirages" ?

On peut représenter l'épreuve par le diagramme



L'évènement E_k est réalisé lorsqu'on a tiré, en n tirages, k boules blanches et $n-k$ boules rouges. Toutes les configurations de tirages de k boules blanches et $n-k$ boules rouges ont la même probabilité, quel que soit l'ordre. D'où la probabilité de l'une d'elles est, en posant $N = b + r$:

$$\begin{aligned} &= \frac{b}{N}\cdot\frac{b-1}{N-1}\cdot\frac{b-2}{N-2}\cdots\frac{b-k+1}{N-k+1}\times \\ &\quad \frac{r}{N-k}\cdot\frac{r-1}{N-k-1}\cdots\frac{r-n+k+1}{N-n+1} \\ &= \frac{b!}{(b-k)!}\frac{r!}{(r-n+k)!}\frac{(N-n)!}{N!} \end{aligned}$$

Il y a $\binom{n}{k}$ configurations possibles d'où

$$P(E_k) \; = \; \frac{\binom{b}{k}\binom{r}{n-k}}{\binom{N}{n}}$$

C'est la loi hypergéométrique. $\mathcal{H}(n, b, r)$.

Exemple :

Dans une population de N individus, on prélève un individu que l'on marque et que l'on relâche dans la population. On répète l'opération. La population est alors formée de deux types d'individus : ceux qui sont marqués et ceux qui ne le sont pas. On capture (successivement) n individus, parmi lesquels X sont marqués. Cette expérience correspond à une loi hypergéométrique.