Lycée Pierre-Gilles de Gennes
Mathématiques

BCPST
2022-2023

Notes de cours 01

Probabilités : des modèles finis aux modèles généraux

Table des matières

1 Modèles probabilistes : un exemple.

La tribu Néanderthal

On peut imaginer la situation suivante, un groupe de chasseurs-cueilleurs (une tribu ?) dont je fais partie, chasse le mammouth et la question vitale qui se pose est : va-t-on manger ce soir ?

Pour ce problème, les variables intéressantes sont certainement, pour chaque tir, la direction initiale du lancer, la puissance avec laquelle celui-ci est effectué et la distance parcourue par la sagaie.

Concentrons nous sur un seul tir.

Je laisse les spécialistes de mécanique établir que, en terrain plat, si la sagaie, de masse m (kg), est lancée avec un angle initial (rad) de α_0, une énergie cinétique initiale E_0 (J, $1\mathrm{J} = 1\mathrm{kg}^2\,\mathrm{m}^2\,\mathrm{s}^{-2}$) alors la distance parcourue [1] est, en m,

$$d = \frac{2.E_0 \sin(2\alpha_0)}{m.g} \tag{+}$$

La relation (+) est issue d'une modélisation très élaborée qui fait appel au calcul différentiel, aux concepts de force, d'accélération, etc...

Une fois la relation (+) connue, on peut vérifier [2] sa validité sur une expérience *ad-hoc* pour laquelle les 4 quantités E_0, α_0, m et d sont effectivement mesurables. Il s'agit de vérifier que le quotient $\frac{E_0 \sin(2\alpha_0)}{m.d}$ est essentiellement constant, valant g.

Imaginons que l'on ait un choix du lanceur de sagaie. Cette situation implique que l'angle de jet et l'énergie initiale varient tous les deux et sachant cela, on peut se demander les chances que soit atteinte une cible ayant une certaine largeur. Cela peut se mesurer expérimentalement en itérant suffisamment l'expérience pour évaluer la proportion de succès.

Dans ce problème, la/les variable/s aléatoire/s **primitive/s**, ce qui est « vraiment » tiré au hasard, c'est le type du tireur, ou, de façon qui s'en déduit immédiatement α_0, l'angle initial et E_0, l'énergie initiale.

Les autres paramètres, m, la masse de la sagaie, tout comme g, sont des constantes.

La variable intéressante pour répondre à la question est d la distance à laquelle se plante la sagaie.

La variable d est aussi aléatoire, elle dépend, via la formule (+) de α_0 et E_0. Sa distribution dépend de la distribution jointe de α_0 et E_0.

Les v.a. α_0, E_0 et d sont des variables aléatoires numériques pourvu que des unités physiques aient été choisies.

L'événement intéressant, c'est de savoir si la sagaie touche le mammouth ou pas, autrement dit, si la bête est entre les distances d_- et d_+, est-ce que $d \in [d_-, d_+]$?

D'un point de vue probabiliste, « $d \in [d_-, d_+]$ » est un événement et on cherche à connaitre sa probabilité.

Exercice 1.— Un groupe de trois chasseurs dispose d'une seule sagaie (de masse $m = 1$kg), la cible est un troupeau de mammouths à une distance entre 20 et 30 m.

Leurs forces sont variées, le plus malingre, resp. le moyen, resp. le plus fort développe au lancer une énergie de 100J, resp. 200J, resp. 400J.

Ils tirent au sort le lanceur et celui-ci tire sa sagaie avec (au hasard), l'un des angles (en rad) $\frac{\pi}{8}$, $\frac{\pi}{4}$ et $\frac{3\pi}{8}$.

Quelle est la probabilité de toucher un animal ? Quelle est la loi de la distance à laquelle la sagaie est lancée ? Et si le plus fort a deux fois plus chances d'être choisi que chacun des autres ?

Exercice 2.—Programmer un tir de sagaie suivant les règles exposées dans l'exercice 1, effectuer un grand nombre de tirages et tracer l'histogramme des résultats.

1. Si la vitesse initiale est v_0, la formule est $d = \frac{v_0^2 \sin(2\alpha_0)}{g}$, g étant l'accélération de la pesanteur, *as usual*
2. Vérifier la validité d'une formule sur des expériences réelles est un problème de statistiques.

2 Un résumé de la théorie des probabilités de première année

On rappelle rapidement un certain nombre de définitions de première année afin de pouvoir les replacer dans le contexte général.

La théorie des probabilités de première année permet de traiter les modèles probabilistes comportant essentiellement un nombre fini de variables aléatoires, chacune pouvant prendre une gamme finie de valeurs (numériques ou pas).

2.1 Espaces probabilisés finis

Définition 1 (Espace probabilisé fini)**.** *Un espace probabilisé fini* $(\Omega, \mathscr{T}, \mathbb{P})$ *est constitué*

(i) d'un ensemble fini, Ω*;*

(ii) de la « tribu » des événements, $\mathscr{T} = \mathscr{P}(\Omega)$*, qui n'est autre que l'ensemble de toutes les parties de* Ω*;*

(iii) d'une probabilité $\mathbb{P}$ *définie sur* $\mathscr{T}$*, à savoir une application* $\mathbb{P} : \mathscr{T} \to [0,1] \subset \mathbb{R}$ *qui à tout événement* $A \in \mathscr{T}$ *fait correspondre sa probabilité* $\mathbb{P}(A)$*.*

L'application $\mathbb{P}$ doit jouir d'un certain nombre de propriétés pour pouvoir être appelée « probabilité », la liste de ces propriétés est celle apparaissant dans la définition 11, p. 18. D'autres propriétés fondamentales sont listées dans la Proposition 12, p. 18

Définition 2 (Indépendance)**.** *Soit* $(\Omega, \mathscr{T}, \mathbb{P})$ *un espace probabilisé fini.*

Deux évenements $A, B \in \mathscr{T}$ *sont dits* indépendants *(sous la probabilité* $\mathbb{P}$*) si*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A).\mathbb{P}(B).$$

Remarques :

1. Si les événements A et B sont indépendants, les événements A et $\overline{B}$ le sont aussi car

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

et donc

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

2. Un événement A est indépendant de lui-même si et seulement si $\mathbb{P}(A) = 0$ ou $\mathbb{P}(A) = 1$, *i.e.* ssi A est négligeable ou presque-sûr [3].

3. L'indépendance d'une famille d'événements comportant (au moins) trois événements A, B et C *n'est pas* le fait [4] que

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A).\mathbb{P}(B).\mathbb{P}(C)$$

3. quasi-certain

4. On écrira la définition précise dans le chapitre suivant, elle est liée à l'indépendance de la famille des trois v.a. indicatrices des événements A, B et C.

2.2 Variables aléatoires

Définition 3 (Variables aléatoires)**.** *Soit* $(\Omega, \mathscr{T}, \mathbb{P})$ *un espace probabilisé fini. Une variable aléatoire* X *(sous-entendu : définie sur cet espace), à valeurs dans un ensemble* $\mathscr{X}$ *est une application* $X : \Omega \to \mathscr{X}$*.*

Soit X une variable aléatoire à valeurs dans un ensemble $\mathscr{X}$.

(i) Pour x, un élément de $\mathscr{X}$, l'évenement (probabiliste) décrit par la proposition "$X = x$" est noté[5] $\{X = x\}$ et est défini comme étant l'élément de $\mathscr{T}$:

$$\{X = x\} := \{\omega \in \Omega, X(\omega) = x\}$$

La probabilité de cet évenement est notée $\mathbb{P}(X = x)$, par simplification de la notation $\mathbb{P}(\{X = x\})$.

(ii) Dans ce cadre fini, l'ensemble $X(\Omega) \subset \mathscr{X}$ des valeurs effectivement prises par X est fini.

(iii) Si $X(\Omega) = \{x_1, \dots, x_N\} \subset \mathscr{X}$ est l'ensemble des valeurs prises, la *distribution* ou *loi* de X est caractérisée par la famille des probabilités des événements deux à deux incompatibles $\{X = x_1\}, \dots, \{X = x_N\}$, *i.e.* la famille $\mathbb{P}(X = x_1), \dots, \mathbb{P}(X = x_N)$.

Définition 4 (v.a. uniforme sur un ensemble fini)**.** *Soit* $x_1, \dots, x_K$*,* K *éléments distincts d'un ensemble* $\mathscr{X}$*,* X *une v.a. à valeurs dans* $\mathscr{X}$*. Si*

$$\forall k \in \{1, \dots, K\}, \mathbb{P}(X = x_k) = \frac{1}{K}$$

On dit que X *est distribuée suivant la loi uniforme sur* $\{x_k, k \in \{1, \dots, K\}\}$*. On note ce fait*

$$X \sim \mathscr{U}_{\{x_k, k \in \{1, \dots, K\}\}}$$

1. Une telle variable est (presque sûrement) à valeurs dans $\{x_k, k \in \{1, \dots, K\}\}$, on a (formule de la probabilité d'une disjonction de cas)

$$\mathbb{P}(X = x_1 \text{ ou } \dots \text{ ou } X = x_K) = \sum_{k=1}^{K} \mathbb{P}(X = x_k) = K.\frac{1}{K} = 1.$$

2. Etant donné un ensemble fini $\{x_k, k \in \{1, \dots, K\}\}$, il existe un ensemble fini Ω, une probabilité $\mathbb{P}$ et une v.a. X définie sur Ω telle que

$$X \sim \mathscr{U}_{\{x_k, k \in \{1, \dots, K\}\}}$$

Il suffit de prendre $\Omega = \{1, \dots, K\}$, $\mathbb{P}(A) = \frac{\#A}{K}$ et $X : \Omega \to \mathscr{X}$ définie par $\forall \omega \in \Omega, X(\omega) = x_\omega$.

3. Une v.a. uniforme c'est ce à quoi pense un quidam lorsque'on lui demande de tirer au hasard un nombre entre 1 et 10. Dans ce contexte « de la vie courante », la locution « au hasard » implique la plupart du temps uniformité (aucune valeur n'est privilégiée) et indépendance (si plusieurs tirages sont faits, ils ne dépendent pas les uns des autres). Pour nous, en théorie des probabilités, « au hasard » est insuffisant et on demande des précisions du type loi du tirage, dépendance ou indépendance, etc...

5. C'est une simplification typographique éludant les occurences des références à ω

A propos de la distribution d'une v.a. X dans ce contexte fini, on peut noter que si [6] $X(\Omega) \subset \{x_1, \ldots, x_M\} \subset \mathscr{X}$ alors

$$\sum_{k=1}^{M} \mathbb{P}(X = x_k) = 1$$

et la connaissance de $\mathbb{P}(X = x_1), \ldots, \mathbb{P}(X = x_M)$ implique la connaissance [7] de la distribution de X.

La notation d'événement (concernant la v.a. X à valeurs dans $\mathscr{X}$) s'étend naturellement à toutes les parties $\mathscr{X}'$ de $\mathscr{X}$. L'événement décrit par la proposition "$X \in \mathscr{X}'$" se note $\{X \in \mathscr{X}'\}$ et est défini comme

$$\left\{X \in \mathscr{X}'\right\} := \{\omega \in \Omega, X(\omega) \in \mathscr{X}'\}$$

Noter que si X et Y sont deux v.a. à valeurs respectivement dans $\mathscr{X}$ et $\mathscr{Y}$ alors l'accouplement $Z = (X,Y)$ est une v.a. à valeurs dans le produit $\mathscr{Z} = \mathscr{X} \times \mathscr{Y}$. Si $\mathscr{X}' \subset \mathscr{X}$, $\mathscr{Y}' \subset \mathscr{Y}$ et $\mathscr{Z}' = \mathscr{X}' \times \mathscr{Y}' \subset \mathscr{Z}$, on a l'égalité d'événements

$$\left\{X \in \mathscr{X}'\right\} \cap \left\{Y \in \mathscr{Y}'\right\} = \left\{X \in \mathscr{X}' \text{ et } Y \in \mathscr{Y}'\right\} = \left\{(X,Y) \in \mathscr{X}' \times \mathscr{Y}'\right\} = \left\{Z \in \mathscr{Z}'\right\}$$

La distribution de $Z = (X,Y)$ s'appelle la loi jointe de X et Y.

Proposition-Définition 5 (Indépendances)**.** *Soit $(\Omega, \mathscr{T}, \mathbb{P})$ un espace probabilisé fini.*

(i) *Deux v.a. X et Y à valeurs respectivement dans les ensembles $\mathscr{X}$ et $\mathscr{Y}$ sont dites* indépendantes *si A et B sont indépendants dès que A est un événement « ne concernant que » X et B est un événement « ne concernant que » Y, autrement dit si*

$$\forall \mathscr{X}' \subset \mathscr{X}, \forall \mathscr{Y}' \subset \mathscr{Y}, \mathbb{P}(X \in \mathscr{X}' \textit{ et } Y \in \mathscr{Y}') = \mathbb{P}(X \in \mathscr{X}').\mathbb{P}(Y \in \mathscr{Y}')$$

(ii) *Les v.a. X et Y sont indépendantes si (et seulement si)*

$$\forall x \in \mathscr{X}, \forall y \in \mathscr{Y}, \mathbb{P}(X = x \textit{ et } Y = y) = \mathbb{P}(X = x).\mathbb{P}(Y = y)$$

Remarques :

1. L'indépendance d'une famille finie de variables aléatoires est la variation naturelle sur (i) : pour trois variables aléatoires X,Y,Z à valeurs respectivement dans $\mathscr{X}$, $\mathscr{Y}$, $\mathscr{Z}$, on dit que celles-ci sont indépendantes si

$$\forall \mathscr{X}' \subset \mathscr{X}, \forall \mathscr{Y}' \subset \mathscr{Y}, \forall \mathscr{Z}' \subset \mathscr{Z}, \mathbb{P}(X \in \mathscr{X}' \text{ et } Y \in \mathscr{Y}' \text{ et } Z \in \mathscr{Z}') = \mathbb{P}(X \in \mathscr{X}').\mathbb{P}(Y \in \mathscr{Y}').\mathbb{P}(Z \in \mathscr{Z}')$$

2. Si X et Y sont indépendantes, la connaissance des distributions de X et Y entraine la connaissance de la loi jointe de X et Y, c'est à dire la distribution de l'accouplement $Z = (X,Y)$.
3. Réciproquement, la connaissance de la distribution de l'accouplement $Z = (X,Y)$, *i.e.* de la loi jointe du couple (X,Y) permet de retrouver (calcul des *lois marginales*) les distributions de X et de Y. On peut alors, en déduire l'indépendance (ou pas) des v.a. X et Y.
4. Une v.a. constante est indépendante de toute variable aléatoire, y compris d'elle-même. Une v.a. X est indépendante d'elle même si et seulement elle est presque sûrement constante : *i.e.* il existe x_0 tel que $\mathbb{P}(X = x_0) = 1$.

---

6. Noter la seule inclusion
7. La distribution complète d'une v.a. contient la quantité maximale d'information que l'on peut espérer d'une v.a. dans un modèle probabiliste. La connaissance d'un ensemble $\mathscr{X}' \subset \mathscr{X}$ suffisamment restreint dans lequel X prend ses valeurs est un premier pas fondamental dans cette recherche d'informations

2.3 Espérance

Définition 6 (Espérance)**.** *Si X est une v.a.* réelle*, i.e. à valeurs dans $\{x_1, \ldots, x_M\} \subset \mathbb{R}$, son espérance est par définition :*

$$\mathbb{E}(X) = \sum_{k=1}^{M} x_k.\mathbb{P}(X = x_k)$$

On remarque cette définition est bien posée car si X est (presque sûrement) à valeurs dans [8] $\{x_1, \ldots, x_N\} \subset \{x_1, \ldots, x_M\} \subset \mathbb{R}$ alors

$$\sum_{k=1}^{N} x_k.\mathbb{P}(X = x_k) = \sum_{k=1}^{M} x_k.\mathbb{P}(X = x_k)$$

du fait que $\mathbb{P}(X = x_{N+1}) = \cdots = \mathbb{P}(X = x_M) = 0$. Pour voir cela, on doit noter que par définition de la notation ensembliste les nombres réels (distincts) $x_{N+1} \ldots x_M$ ne font pas partie de l'ensemble $\{x_1, \ldots, x_N\}$ et donc la probabilité $\mathbb{P}(X \in \{x_{N+1}, \ldots, x_M\})$ est nulle.

On trouve parfois comme définition de l'espérance $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x.\mathbb{P}(X = x)$. C'est une définition correcte mais qui doit être complétée par la remarque ci-dessus. Pour calculer l'espérance, on peut faire la somme sur n'importe quel ensemble fini de valeurs contenant (p.s.) l'ensemble des valeurs prises.

Exemples :

1. L'espérance d'une v.a. constante est cette constante. En effet si X est constante valant x_1, X est à valeurs dans $\{x_1\}$ et $\mathbb{P}(X = x_1) = 1$. La formule donne
$$\mathbb{E}(X) = x_1.\mathbb{P}(X = x_1) = x_1$$
2. Une variable aléatoire X est dite « de BERNOULLI » (de paramètre de succès p, ce que l'on note $X \sim \mathscr{B}(p)$) si elle est à valeurs dans $\{0, 1\}$ avec $\mathbb{P}(X = 1) = p$. Son espérance est $\mathbb{E}(X) = p$. Une telle variable est indicatrice de son propre succès :
$$X = 1\!\!1_{\{X=1\}}$$
3. Une variable aléatoire X est dite « binomiale » (de paramètres de nombre d'essais n et de succès p, ce que l'on note $X \sim \mathscr{B}(n, p)$) si elle a la même loi qu'une somme de n variables de BERNOULLI indépendantes de même paramètre de succès p. Son espérance est $\mathbb{E}(X) = n.p$.

NB :On admet que l'espérance est une opération linéaire :

1. Si X, Y sont deux v.a. réelles prenant chacune un nombre fini de valeurs alors $Z = X + Y$ est aussi une v.a. réelle prenant un nombre fini de valeurs et on a (additivité de l'espérance)
$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$
2. Si X, est une v.a. réelle, $\lambda \in \mathbb{R}$ une constante alors $W = \lambda.X$ est aussi une v.a. réelle prenant un nombre fini de valeurs et on a
$$\mathbb{E}(\lambda.X) = \lambda.\mathbb{E}(X)$$
3. Plus généralement, Si $X_1, \ldots, X_n$ sont des v.a. réelles prenant un nombre fini de valeurs, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, des constantes alors $Z = \sum_{i=1}^{n} \lambda_i.X_i$ est aussi une v.a. réelle prenant un nombre fini de valeurs et
$$\mathbb{E}(Z) = \mathbb{E}(\sum_{i=1}^{n} \lambda_i.X_i) = \sum_{i=1}^{n} \lambda_i.\mathbb{E}(X_i)$$

8. Ce qui signifie $\mathbb{P}(X \in \{x_1, \ldots, x_N\}) = 1$

2.4 L'espérance et la formule de transfert

Si X est une v.a. à valeurs dans $\{x_k, k \in \{1,\dots,K\}\} \subset \mathscr{X}$ et $h : \mathscr{X} \to \mathbb{R}$ est une fonction quelconque. La variable aléatoire $Y = h(X)$ définie par

$$\forall \omega \in \Omega, Y(\omega) = h(X(\omega))$$

est une variable prenant les valeurs *réelles* $h(x_1),\dots,h(x_K)$. On a la *formule de transfert*

$$\mathbb{E}(h(X)) \quad = \quad \sum_{k=1}^{K} h(x_k)\mathbb{P}(X = x_k)$$

Elle permet de calculer l'espérance d'une fonction réelle quelconque de X. La donnée de la formule de transfert *générique*, avec une fonction *h quelconque*, est équivalente à la donnée de la distribution de X.

Démonstration de la Formule de tranfert, (HP). Les K nombres $h(x_1),\dots,h(x_K)$ n'ont aucune raison d'être distincts. Il forment un ensemble fini que nous notons $\{y_1,\dots,y_L\}$. Par convention, dans cette écriture les y_ℓ sont distincts. A chacun de ces y_ℓ correspond l'ensemble des x_k tels que $h(x_k) = y_\ell$.

On a, par définition de l'espérance d'une v.a prenant un nombre fini de valeurs,

$$\mathbb{E}(Y) = \sum_{\ell=1}^{L} y_\ell.\mathbb{P}(Y = y_\ell)$$

Mais, pour chaque ℓ, on a

$$\mathbb{P}(Y = y_\ell) = \sum_{k\in\{1,\dots,K\}, h(x_k)=y_\ell} \mathbb{P}(X = x_k) = \sum_{k=1}^{K} 1\!\!1_{\{h(x_k)=y_\ell\}}\mathbb{P}(X = x_k)$$

On a donc

$$\begin{aligned}
\mathbb{E}(h(X)) \quad &= \quad \sum_{\ell=1}^{L}\sum_{k=1}^{K} y_\ell.1\!\!1_{\{h(x_k)=y_\ell\}}\mathbb{P}(X = x_k) \\
&= \quad \sum_{k=1}^{K}\sum_{\ell=1}^{L} 1\!\!1_{\{h(x_k)=y_\ell\}} h(x_k)\mathbb{P}(X = x_k) \\
&= \quad \sum_{k=1}^{K} \underbrace{\left(\sum_{\ell=1}^{L} 1\!\!1_{\{h(x_k)=y_\ell\}}\right)}_{=1} h(x_k)\mathbb{P}(X = x_k) \\
&= \quad \sum_{k=1}^{K} h(x_k)\mathbb{P}(X = x_k)
\end{aligned}$$

Ce qui est la formule de transfert annoncée. $\square$

Complément : Additivité de l'espérance (HP). La formule de transfert est à la base d'une démonstration de l'additivité de l'espérance de v.a.r réelles, via la formation d'un couple de v.a. Supposons que X soit à valeurs dans $\mathscr{X} = \{x_k, k \in \{1,\dots,K\}\} \subset \mathbb{R}$, que Y soit à valeurs dans $\mathscr{Y} = \{y_\ell, \ell \in \{1,\dots,L\}\} \subset \mathbb{R}$, et posons $Z = (X,Y)$ et $h : \mathbb{R}^2 \to \mathbb{R}$ la fonction « somme » définie par

$$\forall (x,y) \in \mathbb{R}^2, h(x,y) = x + y$$

La v.a. Z est à valeurs dans $\mathscr{Z} = \mathscr{X} \times \mathscr{Y} = \{(x_k,y_\ell), k \in \{1,\dots,K\}, \ell \in \{1,\dots,L\}\} \subset \mathbb{R}^2$, ensemble comportant $K.L$ éléments et on a, en démarrant par la formule de transfert pour Z et en utilisant ensuite les règles usuelles sur les sommes (doubles),

$$\begin{aligned}
\mathbb{E}(X+Y) &= \mathbb{E}(h(Z)) = \sum_{z\in\mathscr{Z}} h(z)\mathbb{P}(Z=z) \\
&= \sum_{k,l} h(x_k,y_\ell)\mathbb{P}((X,Y) = (x_k,y_\ell)) = \sum_{k,l}(x_k+y_\ell)\mathbb{P}(X = x_k \text{ et } Y = y_\ell) \\
&= \sum_{k,l} x_k\mathbb{P}(X = x_k \text{ et } Y = y_\ell) + \sum_{k,l} y_\ell\mathbb{P}(X = x_k \text{ et } Y = y_\ell) \\
&= \sum_k x_k \underbrace{\left(\sum_\ell \mathbb{P}(X = x_k \text{ et } Y = y_\ell)\right)}_{\mathbb{P}(X=x_k)} + \sum_\ell y_\ell \underbrace{\left(\sum_k \mathbb{P}(X = x_k \text{ et } Y = y_\ell)\right)}_{\mathbb{P}(Y=y_\ell)} \\
&= \mathbb{E}(X) + \mathbb{E}(Y)
\end{aligned}$$

□

2.5 Espérance et indépendance

Théorème 7 (Espérance et indépendance)**.** *Si X et Y sont indépendantes à valeurs réelles alors pour toutes fonctions $f,g : \mathbb{R} \to \mathbb{R}$,*

$$\mathbb{E}(f(X).g(Y)) = \mathbb{E}(f(X)).\mathbb{E}(g(Y))$$

La réciproque est trivialement vraie.

Démonstration. On utilise la formule de transfert associée à la variable aléatoire accouplée $Z = (X,Y)$ pour évaluer le membre de gauche, l'indépendance de X et Y donne la loi de Z : $\mathbb{P}(Z = (x_k,y_\ell)) = \mathbb{P}(X = x_k).\mathbb{P}(Y = y_\ell)$ et une manipulation de somme double mène alors au membre de droite via les formules de transfert pour X et Y.

Concernant la réciproque, sachant que pour A et B des parties de $\mathbb{R}$, on a

$$\mathbb{P}(X\in A) = \mathbb{E}(1\!\!1_{\{X\in A\}}), \mathbb{P}(Y\in B) = \mathbb{E}(1\!\!1_{\{Y\in B\}}), \mathbb{P}(X\in A \text{ et } Y\in B) = \mathbb{E}(1\!\!1_{\{X\in A \text{ et } Y\in B\}}) = \mathbb{E}(1\!\!1_{\{X\in A\}}.1\!\!1_{\{Y\in B\}})$$

la relation $\mathbb{P}(X \in A \text{ et } Y \in B) = \mathbb{P}(X \in A).\mathbb{P}(Y \in B)$ provient directement du choix $f = 1\!\!1_A$, $g = 1\!\!1_B$. □

Exercice 3.— Jet de deux dés à 4 faces, loi de la somme, loi de la différence.

Indication:Faire une décomposition des évenements suivant la valeur prise par le premier dé (*i.e.* formule des probabilités totales)

Exercice 4.—

1. Donner l'espérance d'une v.a. uniforme sur $\{0,\dots,n\}$, d'une v.a uniforme sur $\{1,\dots,n\}$.

2. Donner l'espérance d'une v.a. uniforme sur l'ensemble des entiers naturels pairs inférieurs à 500, (on remarquera que c'est le double d'une v.a. uniforme sur $\{0,\dots,250\}$, d'une v.a. uniforme sur l'ensemble des entiers naturels impairs inférieurs à 501 (se ramener au cas précédent en retranchant 1).

Exercice 5.— Une urne contient n jetons numérotés de 1 à n ; On effectue deux tirages avec remise ; Soit X le plus grand des deux numéros tirés, Y le plus petit.
1. Etablir une modélisation sensée de cette expérience. Quelles sont les v.a. naturelles ? Quelles sont leurs distributions ? Comment sont-elles l'une par rapport à l'autre ?
2. Donner la loi du couple (X,Y).
3. Donner les lois de X et Y. Donner leurs espérances.
4. Donner la loi de $Z = X - Y$. On donnera d'abord un ensemble raisonnable de valeurs pouvant être prises par Z. Espérance ? Variance ?

Exercice 6.— Fonction de répartition et formule de CAVALIERI.
1. Soit X une v.a. réelle ne prenant qu'un nombre fini de valeurs distinctes $\{x_1, \dots, x_K\}$. Montrer que pour tout $x \in \mathbb{R}$,

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{k=1}^{K} 1\!\!1_{\{x_k \leq x\}})\mathbb{P}(X = x_k) = \sum_{k: x_k \leq x} \mathbb{P}(X = x_k)$$

2. Montrer que si de plus, les valeurs prises par X sont *positives*, on a

$$\mathbb{E}(X) = \int_0^{+\infty} \mathbb{P}(X > x)\, dx = \int_0^{+\infty} (1 - F_X(x))\, dx$$

2.6 Variance et covariance

1. Si X est une v.a.r prenant un nombre fini de valeurs, sa variance $\mathbb{V}(X)$ est, par définition,
$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$$
2. Si X, Y sont deux v.a.r prenant un nombre fini de valeurs, leur covariance $\mathbb{C}ov(X,Y)$ est, par définition,
$$\mathbb{C}ov(X,Y) = \mathbb{E}((X - \mathbb{E}(X)).(Y - \mathbb{E}(Y)))$$

En développant et en utilisant la linéarité de l'espérance, on obtient, les formules de KOENIG–HUYGUENS

1. $$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$
2. $$\mathbb{C}ov(X,Y) = \mathbb{E}(X.Y) - \mathbb{E}(X).\mathbb{E}(Y)$$

La formule de transfert donne alors, en supposant X à valeurs dans $\{x_k, k \in \{1, \dots, K\}\} \subset \mathbb{R}$, Y à valeurs dans $\{y_\ell, \ell \in \{1, \dots, L\}\} \subset \mathbb{R}$, en posant

$$m_x = \mathbb{E}(X) = \sum_{k=1}^{K} x_k \mathbb{P}(X = x_k),\ m_y = \mathbb{E}(X) = \sum_{\ell=1}^{L} y_\ell \mathbb{P}(X = y_\ell),$$

$$\mathbb{V}(X) = \sum_{k=1}^{K} (x_k - m_x)^2 \mathbb{P}(X = x_k) = \sum_{k=1}^{K} x_k^2 \mathbb{P}(X = x_k) - m_x^2$$

$$\mathbb{C}ov(X,Y) = \sum_{k=1}^{K} \sum_{\ell=1}^{L} (x_k - m_x)(y_\ell - m_y)\mathbb{P}(X = x_k, Y = y_\ell) = \sum_{k=1}^{K} \sum_{\ell=1}^{L} x_k.y_\ell \mathbb{P}(X = x_k, Y = y_\ell) - m_x.m_y$$

Pour cela, on note que $Z = (X,Y)$ est une v.a. prenant un nombre fini de valeurs dans $\mathbb{R}^2$. On a appliqué la formule de transfert avec $h(Z) = h(X,Y) = (X - m_x).(Y - m_y)$. Et, pour la toute dernière formule, appliqué la formule de transfert avec $h(Z) = h(X,Y) = X.Y$.

Théorème 8 (Covariance et indépendance)**.** *Si X et Y sont deux variables aléatoires réelles indépendantes alors*

$$\mathbb{C}ov(X,Y)=0$$

La réciproque est fausse.

Exercice 7.— Si X est une v.a.r prenant un nombre fini de valeurs entières, on définit, pour $t\in\mathbb{R}$,

$$f(t)=\mathbb{E}(t^X)$$

1. Déterminer f dans le cas où X suit

a. La loi uniforme sur $\{0,\dots,n\}$.

b. La loi de BERNOULLI $\mathscr{B}(p)$.

c. La loi binomiale $\mathscr{B}(n,p)$.

Vérifier que f est une fonction polynômiale. (Elle s'appelle la fonction génératrice de X).

Indication:On pourra commencer par écrire la formule de transfert calculant $\mathbb{E}(h(X))$ pour une fonction h adéquate.

2. Justifier que, en général, f est polynômiale, détermine entièrement la loi de X et que $\mathbb{E}(X)=f'(1)$, $\mathbb{E}(X(X-1))=f''(1)$.

3. Retrouver par ce biais espérance et variance d'une loi binomiale.

4. Retrouver par ce biais espérance et variance de la loi uniforme du a. Indication:On pourra écrire le DL2 en 1 de f

5. Retrouver par ce moyen le fait que si $X\sim\mathscr{B}(n_x,p)$, $Y\sim\mathscr{B}(n_y,p)$, X et Y sont indépendantes alors

$$X+Y\sim\mathscr{B}(n_x+n_y,p).$$

Exercice 8.—

1. Si X et Y sont deux variables de BERNOULLI et $\mathbb{C}ov(X,Y)=0$ alors X et Y sont indépendantes.

2. On suppose $X\sim\mathscr{U}_{\{-1,0,+1\}}$ et $Y=|X|$. Montrer : $\mathbb{C}ov(X,Y)=0$ et X et Y ne sont pas indépendantes.

Exercice 9.— Sur $n=100$ palourdes pêchées en Bretagne sud, on effectue la mesure de leur diamètre. On note d_k le diamètre de la palourde portant le numéro $k\in\{1,\dots,n\}$. On regroupe ces mesures en un tableau et on calcule (via Excel par exemple) leur moyenne $\overline{d}=\frac{1}{n}\sum_{k=1}^n d_k$ et l'écart-type $\sigma=\sqrt{\frac{1}{n}\sum_{k=1}^n(d_k-\overline{d})^2}$.

1. On prend une palourde au hasard et on note D son diamètre. Calculer $\mathbb{E}(D)$ et $\mathbb{V}(D)$ en fonction de $\overline{d}$ et σ.

2.a. Soit $\alpha>0$. Enoncer l'inégalité de BIENAYMÉ–TCHEBYCHEFF pour D au seuil $\alpha\sigma$.

2.b. Comment choisir α pour être sûr que moins d'un quart des palourdes ont un diamètre qui diffère du diamètre moyen de plus de $\alpha.\sigma$ en valeur absolue.

3. En fait, on met nos n palourdes dans un sac que l'on secoue, ce qui fait que les plus grandes palourdes vont plus facilement être tirées. On admet que la palourde numéro k est tirée avec probabilité

$$p_k=\frac{d_k}{n.\overline{d}}$$

3.a. Vérifier que $\sum_{k=1}^n p_k=1$.

3.b. On note D' le diamètre de la palourde tirée au sort suivant ce mode. Pourquoi, heuristiquement, $\mathbb{E}(D')\geq\mathbb{E}(D)$?

3.c. Donner une formule pour $\mathbb{E}(D')$ en fonction de $\mathbb{E}(D^2)$.

3.d. Démontrer mathématiquement l'inégalité heuristique.

2.7 Simulation, interprétation fréquentielle et représentation graphique.

En Python, sous `numpy`, on dispose [9] d'un certain nombre de fonctions (regroupées dans le module `numpy.random`) permettant d'obtenir la valeur d'une v.a. ayant une certaine distribution. Voir
`https://docs.scipy.org/doc/numpy/reference/routines.random.html`
Dans les simulations que nous affectuerons, trois fonctions Python sont particulièrement pratiques

1. `np.random.randint(n)` qui retourne un entier aléatoire uniforme entre 0 et `n-1`
2. `np.random.choice(liste,p=None)` qui retourne un des éléments de la `liste` par défaut tiré uniformément (et avec probabilité donnée par `p` si `p` est une liste de même taille que `liste` spécifiée par le mécanisme d'argument nommé avec valeur par défaut).
3. `np.random.rand()` qui retourne un nombre réel de l'ensemble $\left\{\frac{k}{N}, k \in \{0,\ldots,N-1\}\right\}$. Le nombre entier N est si grand que l'on considère que cette fonction retourne un nombre flottant tiré uniformément dans l'ensemble des flottants de $[0,1[$.

Nous verrons plus tard [10] que la troisième fonction suffit pour simuler tout type de variable aléatoire : c'est la plus fondamentale des trois.

Il est important de noter que chaque appel à l'une de ces fonctions Python produit une valeur indépendante des autres appels.

Dans les situations physiques que nous cherchons à modéliser, en un certain sens, le tirage au sort ne se produit qu'**une fois**.

Si je lance une pièce, le résultat de l'expérience est modélisé par la v.a. X dont les valeurs possibles sont Pile et Face. Le fait de dire

$$\mathbb{P}(X = \text{Pile}) = \frac{1}{2} = \mathbb{P}(X = \text{Face})$$

est une façon d'affirmer que l'on croit la pièce est équilibrée et que le tirage s'est effectué dans de bonnes conditions.

Il est clair cependant que si l'on ne fait qu'UN tirage, on ne peut être sûr de ceci. Une façon de consolider cette croyance est de procéder à un grand nombre de tirages semblables et de constater que *grosso modo*, Pile et Face sont sortis en quantités comparables. On calcule, sur un grand (disons $N = 1000$) nombre de tirages la *fréquence* d'apparition de Pile :

$$f_{\text{Pile}} = \frac{\text{nbre de Pile}}{N}$$

Le nombre f_{Pile} est compris entre 0 et 1 et, concernant, la fréquence d'apparition de Face, on a

$$f_{\text{Face}} = \frac{\text{nbre de Face}}{N} = \frac{1}{2} - f(\text{Pile})$$

Si le nombre f_{Pile} est suffisamment proche [11] de $\frac{1}{2}$, on pourra alors être convaincu que l'hypothèse

$$\mathbb{P}(X = \text{Pile}) = \frac{1}{2} = \mathbb{P}(X = \text{Face})$$

est raisonnable.

9. après importation par `import numpy as np`
10. c'est un thème récurrent du cours
11. De combien est *la* question fondamentale traitée dans le chapitre de statistiques

On peut noter que cette procédure revient à considérer N variables $X_1, \dots X_N$ ayant même distribution que X, **indépendantes**, à effectuer UN tirage de chacune de ces variables, et, une fois les résultats consignés, à calculer les fréquences d'apparition.

En Python, si on convient d'associer Face à 0 et Pile à 1, la fonction `np.random.randint(2)` simule un tirage de pièce « équilibrée » et la première boucle `for` du programme Python 1 remplit une liste de N tirages successifs et indépendants.

La technique d'accumulateur autour de la deuxième boucle `for` évalue la fréquence d'apparition de Pile. Cette fréquence est imprimée à l'écran à l'issue du calcul.

Listing 1 – python/PF-simple.py

```
import numpy as np
N=1000
liste=[]
for i in range(N):
    liste.append(np.random.randint(2))
fPile=0.0
for res in liste:
    fPile+=res
fPile=fPile/N
print("N=",N,";fréquence de Pile= ",fPile)
```

Supposons maintenant que l'on veuille simuler informatiquement une v.a.r X telle que

$$\mathbb{P}(X=0)=0.1, \mathbb{P}(X=0.5)=0.6 \text{ et } \mathbb{P}(X=1.1)=0.3$$

On utilise alors `np.random.choices([0,0.5,1.1],p=[0.1,0.6,0.3])`.

Comment vérifier, en tout cas, *grosso modo* que cette fonction Python fait bien ce qu'on attend d'elle ?

1. On peut comme précédemment, sur un grand nombre d'expériences, calculer les fréquences d'apparition de chaque résultat possible.
2. On peut aussi, après avoir consigné dans une liste les résultats de ces expériences, afficher un histogramme de cette série statistique avec « `bins` » bien choisis.

La commande `plt.hist(x,bins=[...], density = True)` affiche l'histogramme de la série statistique `x` lié au « `bins` ». Les « bins » déterminent les frontières d'intervalles contigus I de l'axe des abscisses et l'aire du rectangle au dessus de chaque intervalle est la fréquence d'apparition de l'événement $x \in I$, *i.e.* la proportion d'éléments de la liste `x` appartenant à l'intervalle I.

Listing 2 – python/tirage-3-val-simple.py

```
import numpy as np
import matplotlib.pyplot as plt
N=1000
liste=[]
x=[0,0.5,1.1] #Valeurs possibles
px=[0.1,0.6,0.3] #Proba d'apparition associée
for i in range(N):
    liste.append(np.random.choice(x,p=px))
plt.hist(liste,bins=[-0.25,0.25,0.75,1.25],normed=True)
plt.show()
plt.hist(liste,bins=[-0.01,0.01,0.49,0.51,1.09,1.11],normed=True)
plt.show()
```

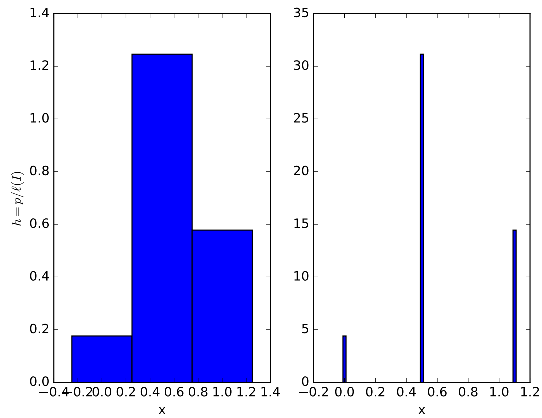


FIGURE 1 – Deux histogrammes d'une simulation de N=1000 valeurs de X, bins de largeur 0.5 autour des valeurs possibles, bins de largeur 0.02. La somme des aires des rectangles vaut 1.

Un apparté sur les bins et les histogrammes

Comme le montre la figure 2, le choix de bons « bins » est crucial pour la compréhension d'une série statistique d'une v.a discrète. Les rédacteurs du rapport G2E (figure 3) ont mieux compris ce problème.

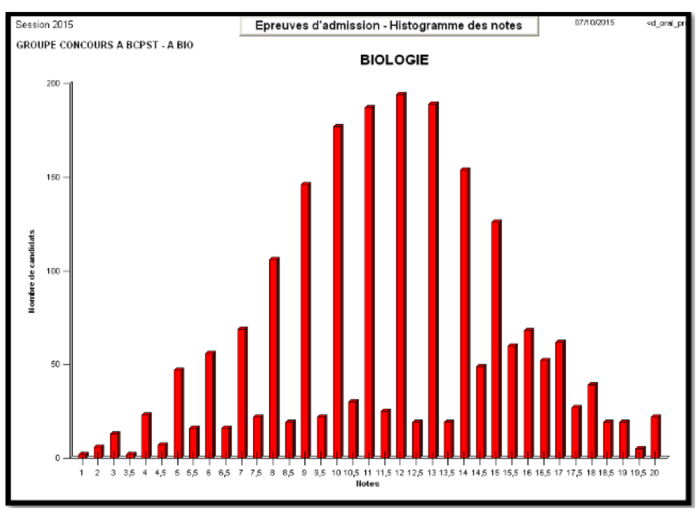


FIGURE 2 – Un histogramme des résultats de l'épreuve de Synthèse-Biologie AV 2015.

Représentation graphique d'une distribution et transformations.

Pour représenter graphiquement une distribution de v.a. réelle X, on place une barre (fléchée) de hauteur $\mathbb{P}(X = x_k)$ au dessus de chaque valeur x_k pouvant être prise par X. *cf.* figures 4 et 5

Une telle représentation est réminiscente de la présentation précédente en **histogramme** de données statistiques.

Une barre fléchée de hauteur p doit être lue comme représentant un rectangle de hauteur infinie, de largeur nulle mais d'aire p. (Une telle chose n'existe pas vraiment)

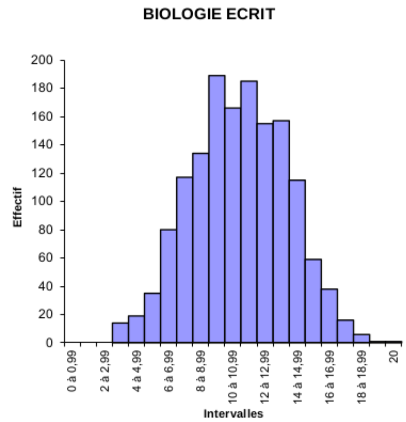


FIGURE 3 – Un histogramme des résultats de l'épreuve de Biologie G2E 2016.

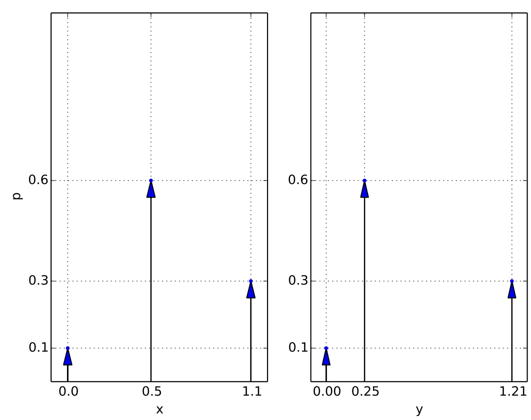


FIGURE 4 – La distribution d'une variable aléatoire X et la distribution de $Y = X^2$.

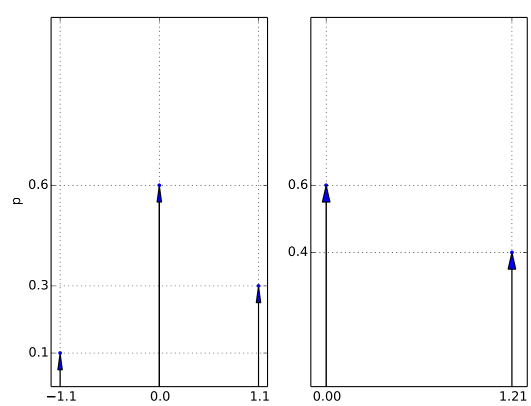


FIGURE 5 – Idem que pour la fig. 4. Noter que $\mathbb{P}(Y = 1.21) = \mathbb{P}(X = -1.1) + \mathbb{P}(X = 1.1)$

Exercice 10.— Représenter graphiquement la distribution d'une variable X suivant la loi binomiale $\mathscr{B}\left(4, \frac{1}{2}\right)$, puis la distribution de $Y = X - 2$ et enfin celle de $Z = Y^2 - 2$.

La fonction de répartition d'une v.a. à nombre fini de valeurs réelles : représentation graphique.

Soit X une variable aléatoire réelle prenant un nombre fini de valeurs. Sa fonction de répartition F_X est la fonction $F_X : \mathbb{R} \to [0,1] \subset \mathbb{R}$ définie par la formule

$$\forall x \in \mathbb{R}, F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(X \in]-\infty, x])$$

Les figures 6 et 7 montrent les graphes des fonctions F_X et F_Y. Noter le « saut » de ces fonctions aux valeurs prises par les variables avec probabilité > 0.

Dans le cas d'une v.a. X ne prenant qu'un nombre fini de valeurs réelles, la fonction F_X est constante par morceaux.

On peut vérifier sur ce cas précis les propriétés générales des fonctions de répartition présentées dans la Proposition 16, en page 20. Une telle fonction est

— croissante sur $\mathbb{R}$,

— continue à droite en tout point de $\mathbb{R}$,

— de limite 0 en $-\infty$, de limite 1 en $+\infty$.

Nous retrouverons cette représentation graphique dans la section **Simulations et construction de v.a prenant un nombre fini de valeurs**.

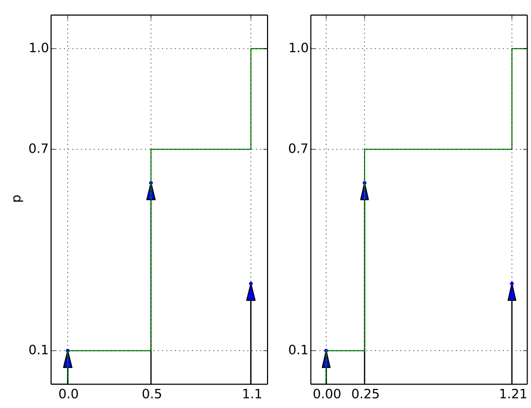


FIGURE 6 – Une variable aléatoire X, $Y = X^2$, distributions et fonctions de répartition.

Exemples

Exercice 11.— Si X suit la loi uniforme sur $\left\{\frac{k}{n}, k \in \{0, \dots, n\}\right\}$ où $n \in \mathbb{N}^*$. Tracer histogramme, fonction de répartition, idem pour $Y = X^2$. Formule de transfert générique pour une fonction pour X. Une remarque sur cette formule ?

Exercice 12.— Tracer, à la machine, histogramme et fonction de répartition pour la loi binomiale $\mathscr{B}(n, p)$.
1. Prendre pour p fixé, des valeurs de n augmentant. Observer l'évolution des graphiques.
2. Fixer $\lambda > 0$. Prendre des valeurs de n augmentant et $p = \frac{\lambda}{n}$

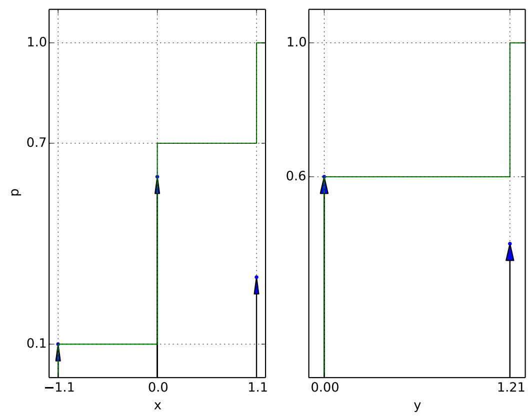


FIGURE 7 – Idem que pour la fig. 6.

2.8 Existence d'espace probabilisé supportant un modèle fini

Le théorème suivant permet de fonder mathématiquement les calculs dans un modèle probabiliste fini en affirmant l'existence d'un espace probabilisé supportant ce modèle :

Théorème 9 (Bonne fondation des modèles probabilistes finis)**.** *Soit $x_1,\dots,x_K$, K éléments distincts d'un ensemble $\mathscr{X}$, $p_1,\dots,p_K \in [0,1]$ tels que $\sum_{k=1}^K p_k = 1$ alors il existe un ensemble (fini) Ω, une probabilité $\mathbb{P}$ sur $\mathscr{T} = \mathscr{P}(\Omega)$ et une application $X : \Omega \to \mathscr{X}$ tels que*

$$\forall k \in \{1,\dots,K\}, \mathbb{P}(X = x_k) = p_k$$

Pour fonder tout modèle probabiliste fini, il suffit de comprendre que l'on peut accoupler les variables (primitives) du modèle en un seul n-uplet X. Celui-ci a une loi et c'est ce n-uplet que l'on tire au sort pour produire en une seule fois toutes les valeurs aléatoires nécessaires dans le modèle.

D'un point de vue informatique, cela reviendrait à tirer d'un coup toutes les valeurs aléatoires dont on a besoin à l'aide d'un seul appel à la fonction `np.random.choice` en utilisant comme argument la liste de tous les n-uplets possibles. Ce n'est pas l'usage. Celui-ci est plutôt de tirer au sort les valeurs aléatoires au fûr et à mesure de leur utilisation par le programme.

Les méthodes de simulation (informatique) de v.a. que nous rencontrerons suggèrent l'existence d'un univers Ω (muni d'une probabilité $\mathbb{P}$) universel, *i.e.* valable pour tous les modèles. Le tirage au sort sur une expérience est en fait le tirage d'une suite de valeurs issues d'appels successifs à la fonction `np.random.rand`.

Exercice 13.—On pose, pour $n \in \mathbb{N}^*$ et $q \in]0,1[$,

$$p_1 = (1-q), \dots p_k = (1-q)q^{k-1}, p_n = (1-q)q^{n-1} \text{ et } p_0 = q^n$$

1. Montrer qu'il existe une variable aléatoire X à valeurs dans $\{0,\dots,n\}$ vérifiant :

$$\forall k \in \{0,\dots,n\}, \mathbb{P}(X = k) = p_k$$

2. Donner $\mathbb{E}(X)$.
3. Donner une fonction Python `X(q,n)` retournant une valeur simulée de X avec les paramètres `q` et `n`. (Pour les 3/2, on pourra utiliser `np.random.choice`. Pourles 5/2, le faire uniquement avec `np.random.rand`)

3 L'axiomatique de KOLMOGOROV

Disclaimer

On donne maintenant les définitions générales d'espace probabilisé, de tribu (d'événements), de probabilité qui vont permettre de traiter le cas de variables prenant possiblement une infinité de valeurs et le cas d'un nombre infini de variables.

C'est une théorie très abstraite, ces définitions sont au programme mais, comme on le verra assez rapidement, on va peu s'en préoccuper.

C'est une généralisation de ce qui se passe pour les modèles finis, il faut « simplement » arrêter de penser que les v.a. ne prennent qu'une gamme finie de valeurs. Ceci implique que l'on puisse traiter une plus large variété de problèmes mais impose des limitations plus subtiles que la seule finitude.

Tribu d'événements

Définition 10. *Soit Ω un ensemble non vide. $\mathscr{T}$ un sous-ensemble*[12] *de $\mathscr{P}(\Omega)$. On dit que $\mathscr{T}$ est une* tribu *(d'événements) si*

1. *$\emptyset \in \mathscr{T}$, $\Omega \in \mathscr{T}$, respectivement l'*événement impossible *et l'*événement certain
2. *$\mathscr{T}$ est stable par passage au complémentaire.*
$$\text{Si } A \in \mathscr{T}, \text{ alors } \overline{A} \in \mathscr{T}$$
3. *$\mathscr{T}$ est stable par $\cup$ dénombrable*[13] *: Si $(A_n)_{n\in\mathbb{N}}$ est une famille d'éléments de $\mathscr{T}$ alors $\cup_{n\in\mathbb{N}} A_n \in \mathscr{T}$ où*
$$\cup_{n\in\mathbb{N}} A_n = \{\omega \in \Omega, \exists n \in \mathbb{N}, \omega \in A_n\}$$
$\mathscr{T}$ est stable par $\cap$ dénombrable[14] *: Si $(A_n)_{n\in\mathbb{N}}$ est une famille d'éléments de $\mathscr{T}$ alors $\cap_{n\in\mathbb{N}} A_n \in \mathscr{T}$ où*
$$\cap_{n\in\mathbb{N}} A_n = \{\omega \in \Omega, \forall n \in \mathbb{N}, \omega \in A_n\}$$

Cette définition permet un glissement du langage, la possibilité d'écrire les « événements » de la réalité comme des objets mathématiques.

1. Un élément $\omega \in \Omega$ est appelé un élément de hasard (en anglais *random element*). On trouve aussi la locution 'configuration élémentaire'. J.P KAHANE a proposé de l'appeler un *randon*.
2. Un événement, c'est ce qui peut avoir lieu ou pas quand un ω est choisi. C'est un ensemble de randons qui ont une propriété logique commune. L'événement a lieu si le ω choisi fait partie de l'ensemble, il n'a pas lieu sinon.

Avec le vocabulaire des événements, les axiomes de tribu se lisent comme suit

1. L'événement impossible, c'est ce qui est toujours faux, n'a jamais lieu. L'événement certain c'est ce qui est toujours vrai, a toujours lieu.
2. Si A est un événement, l'événement contraire $\overline{A}$ est aussi un événement. Ces deux événements ne peuvent avoir lieu simultanément, ils sont incompatibles.

12. Les éléments de $\mathscr{T}$ sont donc des parties de Ω.
13. union dénombrable
14. intersection dénombrable

3. Si A_1, A_2 sont deux événements, $A_1 \cup A_2 = (A_1 \text{ ou } A_2)$ est un événement. C'est l'ensemble des randons ω faisant partie de l'événement A_1 ou de l'événement A_2. C'est l'événement « alternative de A_1 et A_2 ».
4. Si A_1, A_2 sont deux événements, $A_1 \cap A_2 = (A_1 \text{ et } A_2)$ est un événement. C'est l'ensemble des randons ω faisant partie de l'événement A_1 et de l'événement A_2. C'est l'événement « conjonction de A_1 et A_2 ».
5. Les deux derniers axiomes montrent que l'on peut faire une conjonction ou une alternative d'une infinité *dénombrable* d'événements et obtenir encore un événement.

Finalement, les relations logiques de causalité et d'incompatibilité se traduisent aussi dans ce langage :

1. Un événement A implique un événement B si $A \subset B$. Cela signifie que si A a lieu lorsque ω est choisi, alors B a lieu.
2. Deux événements A et B sont dits « équivalents » si $A = B$. Cela signifie que lorsque ω est choisi A a lieu si et seulement si B a lieu.
3. Deux événements sont dits « incompatibles » si leur conjonction est impossible, *i.e.* si $A \cap B = \emptyset$. Cela signifie que si A a lieu lorsque ω est choisi, alors B n'a pas lieu et réciproquement.

Probabilité sur une tribu

Définition 11. *Soit $\mathscr{T}$ une tribu sur un ensemble Ω. Une application $\mathbb{P} : \mathscr{T} \to [0,1] \subset \mathbb{R}$ est appelée une* (mesure de) probabilité. *Si*

1. $\mathbb{P}(\emptyset) = 0$ *et* $\mathbb{P}(\Omega) = 1$.
2. *Si $A_1, \dots, A_n, \dots$ est une famille (au plus dénombrable) d'événements (mutuellement) incompatibles alors*

$$\begin{aligned}\mathbb{P}(\cup_{n\in\mathbb{N}} A_n) &= \mathbb{P}(A_1 \cup \dots \cup A_n \cup \dots) \\ &= \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n) + \dots \\ &= \lim_{N\to+\infty} \sum_{n=1}^{N} \mathbb{P}(A_n)\end{aligned}$$

Si Ω est un ensemble quelconque, $\mathscr{P}(\Omega)$ est une tribu.

1. C'est celle que l'on considère le plus souvent lorsque Ω est fini. On peut alors définir une probabilité sur cette tribu à l'aide d'un système de poids (*cf.* Théorème 9).
2. Lorsque Ω est infini, il y a potentiellement une difficulté relevant de la logique pour définir une mesure de probabilité intéressante sur $\mathscr{P}(\Omega)$. Cela explique pourquoi on se restreint à une tribu d'événements plus petite que $\mathscr{P}(\Omega)$.

En conséquences directes des axiomes

Proposition 12. *1. Si $A \subset B$ alors $\mathbb{P}(A) \leq \mathbb{P}(B)$*

2. (Formule de Poincaré*)*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

En langage presque courant, on peut lire

1. La probabilité d'un événement impossible est nulle. La probabilité d'un événement certain est 1.
2. Si A_1 et A_2 sont incompatibles alors la probabilité de l'alternative de A_1 et A_2 est la somme des probabilités.
3. Si un événement A implique un événement B, la probabilité de A est inférieure à celle de B.

Variables aléatoires réelles

Définition 13. *Soit un espace probabilisé* $(\Omega,\mathscr{T},\mathbb{P})$.

Une variable aléatoire réelle, X, est une application $X:\Omega\to\mathbb{R}$ *telle que pour tout intervalle* $I\in\mathbb{R}$

$$\underbrace{\{X\in I\}}_{\text{notation probabiliste}} := \{\omega\in\Omega, X(\omega)\in I\} = \underbrace{X^{-1}(I)}_{\text{notation ensembliste, HP}}$$

est un événement, i.e. est dans $\mathscr{T}$.

Variables aléatoires réelles, opérations

Une fois que l'on a une ou plusieurs v.a.r, on peut en fabriquer de nouvelles par toutes les opérations algébriques et compositions avec les fonctions numériques de variable réelle usuelles.

Proposition 14. *Soit X, Y, Z,... des v.a.r*

1. *(Stabilité par CL) Si* λ, μ *sont des nombres réels alors* $\lambda.X+\mu.Y$ *est une v.a.r.*
2. *(Stabilité par composition I) Si* $h:I\to\mathbb{R}$ *est une fonction réelle de variable réelle, X à valeurs dans I, un intervalle de* $\mathbb{R}$*, alors* [15] $Y=h(X)$ *définie par*
$$\forall\omega\in\Omega, Y(\omega)=[h(X)](\omega)=h(X(\omega))=(h\circ X)(\omega)$$
est une variable aléatoire réelle.
3. *(Stabilité par composition II) Si* $h:\mathbb{R}^2\to\mathbb{R}$ *est une fonction réelle de deux variables réelles,* $Z=h(X,Y)$ *définie par*
$$\forall\omega\in\Omega, Z(\omega)=[h(X,Y)](\omega)=h(X(\omega),Y(\omega))$$
est une variable aléatoire réelle.

Commentaires sur les hypothèses : on peut prendre h continue, continue par morceaux, monotone, et d'une façon générale, toute fonction que l'on rencontre dans la pratique.

Exemples : Prendre une v.a, mettre au carré, prendre le log, l'exp, en prendre deux, prendre leur max, leur min, leur produit, leur somme, en prendre trois...

Variables aléatoires réelles, loi, fonction de répartition

Définition 15. *Soit un espace probabilisé* $(\Omega,\mathscr{T},\mathbb{P})$. *Soit X une variable aléatoire réelle. La loi de X est la donnée de la famille de nombres* $\{\mathbb{P}(X\in I), I \text{ intervalle de } \mathbb{R}\}$

Plus modestement, on peut se contenter de la sous-famille des intervalles $]\infty,x]$, $x\in\mathbb{R}$ pour reconstruire la famille complète via les propriétés axiomatiques de $\mathbb{P}$. La loi de X est donc donnée par la connaissance de la fonction de répartition de X, définie par

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

Par exemple, pour $a<b$, on a

$$\mathbb{P}(X\in\,]a,b])=F_X(b)-F_X(a),\ \mathbb{P}(X\in\,]a,+\infty[)=1-F_X(a)$$

15. modulo des hypothèses très faibles sur h, toutes les fonctions h que nous utilisons vérifient ces hypothèses relevant de la logique mathématique la plus fondamentale.

Proposition 16. *La fonction de répartition F_X d'une v.a.r X est une fonction $\mathbb{R} \to [0,1]$, vérifiant*

1. *F_X est croissante*
2. $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to+\infty} F_X(x) = 1$
3. *(Hors programme) F_X est continue à droite en tout point*

3.1 V.a. uniforme sur $[0,1]$

Définition et loi

On présente maintenant l'exemple fondamental de v.a.r pouvant prendre une gamme infinie de valeurs.

Définition 17. *Soit U une variable aléatoire réelle. On dit que U suit la* loi uniforme *sur* $[0,1]$ *si, pour tout intervalle I de $\mathbb{R}$,*

$$\mathbb{P}(U \in I) = longueur(I \cap [0,1])$$

Dans ce cas, on note $U \sim \mathscr{U}_{[0,1]}$.

U est p.s. à valeurs dans $[0,1]$. En effet, on a

$$\mathbb{P}(U \notin [0,1]) = \mathbb{P}(U \in]-\infty,0]) + \mathbb{P}(U \in [1,-\infty[) = 0+0 = 0$$

On peut comparer le résultat suivant (admis, difficile) avec le théorème 9, plus faible.

Théorème 18 (Existence d'un modèle probabiliste fondamental)**.** *Il existe un ensemble Ω, une tribu $\mathscr{T}$ sur Ω, une probabilité $\mathbb{P}$ sur $\mathscr{T}$ et une v.a. $U : \Omega \to \mathbb{R}$ tels que*

$$U \sim \mathscr{U}_{[0,1]}$$

Ayant à disposition un tel modèle $(\Omega, \mathscr{T}, \mathbb{P})$, on peut aller plus loin et y construire (miracle de l'infini avec lequel on peut toujours trouver un peu de place libre)une suite $(U_n)_{n\in\mathbb{N}}$ de v.a. sur $(\Omega, \mathscr{T}, \mathbb{P})$, indépendantes, distribuées uniformément sur $[0,1]$. Sur cet espace, à l'aide de cette suite de v.a. uniformes et indépendantes, on peut, en utilisant les techniques de simulation/construction de v.a. décrites dans les sections **Simulations et construction de v.a prenant un nombre fini de valeurs** et **Lois conditionnelles, Simulation informatique,** construire par exemple une chaine de MARKOV de matrice de transition donnée, les modèles de tirage au sort avec ou sans remise présentés dans la suite et tout autre modèle pouvant se programmer. Il suffit d'assimiler un flottant (informatique) de l'intervalle $[0,1]$ avec un nombre réel de ce même intervalle.

Fonction de répartition d'une loi uniforme

— Calcul et graphe. On a, si $U \sim \mathscr{U}_{[0,1]}$,

$$\forall u \in \mathbb{R}, F_U(u) = \begin{cases} 0 & \text{si } u \leq 0 \\ u & \text{si } 0 \leq u \leq 1 \\ 1 & \text{si } u \geq 1 \end{cases}.$$

— Fonction de répartition du carré d'une telle variable ? Soit $V = U^2$. V est p.s. à valeurs dans $[0,1]$, on a, pour $0 \leq v \leq 1$,

$$\mathbb{P}(V \leq v) = \mathbb{P}(U^2 \leq v \text{ et } U \in [0,1]) = \mathbb{P}(0 \leq U \leq \sqrt{v}) = \sqrt{v}$$

et donc

$$\forall v \in \mathbb{R}, F_V(v) = \begin{cases} 0 & \text{si } v \leq 0 \\ \sqrt{v} & \text{si } 0 \leq v \leq 1 \\ 1 & \text{si } v \geq 1 \end{cases}.$$

— *cf.* fig. 8.

Remarque sur la probabilité de prendre une valeur « exactement ».

$$\forall x \in \mathbb{R}, \mathbb{P}(U = x) = \mathbb{P}(U \in [x,x]) = 0$$

En conséquence, U est aussi p.s. à valeurs dans $]0,1[$.

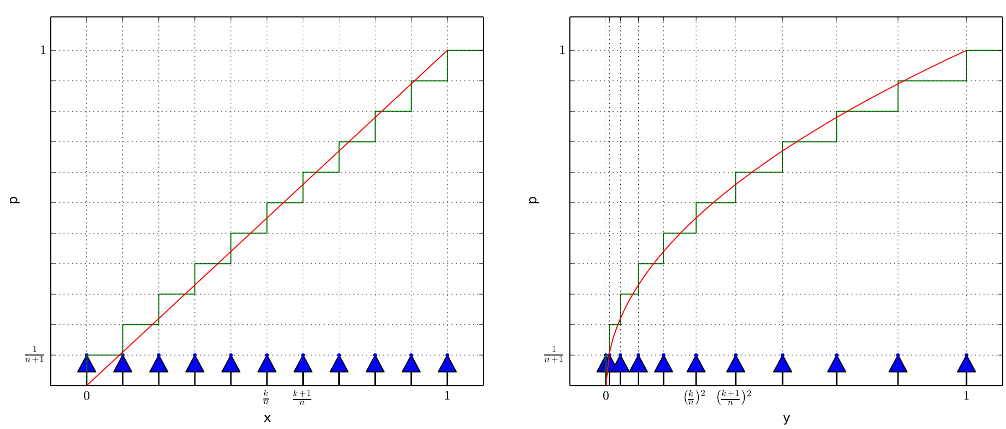


FIGURE 8 – $X \sim \mathscr{U}_{[0,1]}$, $Y = X^2$, fonctions de répartition, comparaison avec le cas discret.

V.a. uniforme sur un intervalle borné

Définition 19. *Soit X une variable aléatoire réelle, a, $b \in \mathbb{R}$, $a < b$. On dit que X suit la* loi uniforme *sur $[a,b]$ si, pour tout intervalle I de $\mathbb{R}$,*

$$\mathbb{P}(X \in I) = \frac{longueur(I \cap [a,b])}{longueur([a,b])} = \frac{1}{b-a} longueur(I \cap [a,b])$$

Dans ce cas, on note $X \sim \mathscr{U}_{[a,b]}$.

1. Si $U \sim \mathscr{U}_{[0,1]}$ alors en posant $X = a + (b-a)U$, $X \sim \mathscr{U}_{[a,b]}$.
2. Si $X \sim \mathscr{U}_{[a,b]}$ alors en posant $U = \frac{X-a}{b-a}$, $U \sim \mathscr{U}_{[0,1]}$.

Simulations et construction de v.a prenant un nombre fini de valeurs.

Pour simuler (informatiquement) des variables aléatoires, les langages informatiques disposent en général d'un *générateur* fondamental de nombres aléatoires. L'instruction s'appelle toujours `rand()` ou `random()`.

Pour nous ce sera `np.random.rand()` où `np` est l'alias de `numpy` Chaque appel à ce générateur retourne un nombre aléatoire de distribution uniforme sur $[0,1[$, indépendant des précédents. Cela signifie que si l'on fait plusieurs appels successifs, la suite des nombres obtenus a les mêmes propriétés statistiques (distribution, moyenne, ...) qu'une suite (abstraite) de valeurs prises par $U_1,\dots,U_n$ où $U_1,\dots,U_n$ sont des v.a indépendantes et

$$\forall k \in \{1,\dots,n\}\,, U_k \sim \mathscr{U}_{[0,1]}$$

La question qui se pose est la suivante : à partir d'une v.a. uniforme sur $[0,1]$, U, comment construire une v.a. (prenant un nombre fini de valeurs) ayant une loi discrète finie donnée. Reformulé pour Python, la question est : comment réécrire une fonction du type `np.random.choice` en utilisant `np.random.rand` ?

Supposons que l'on doive simuler une v.a. X prenant trois valeurs distinctes (non nécessairement numériques), x_1, x_2 et x_3. L'information dont on dispose est la distribution de X

$$p_1 := \mathbb{P}(X = x_1),\, p_2 := \mathbb{P}(X = x_2),\, p_3 := \mathbb{P}(X = x_3),$$

La méthode est simple (graphique) On découpe l'intervalle $[0,1]$ en trois intervalles disjoints I_1, I_2 et I_3 de longueurs respectives p_1, p_2 et p_3. On tire au sort U grâce au générateur uniforme. Si $U \in I_1$, on prend $X = x_1$, si $U \in I_2$, on prend $X = x_2$,...

En d'autres termes, on pose

$$I_1 = [0,p_1[\,,\, I_2 = [p_1, p_1+p_2[\,,\, I_3 = [p_1+p_2, 1[\text{ et}$$

$$X = \begin{cases} x_1 & \text{si } U \in I_1 \\ x_2 & \text{si } U \in I_2 \\ x_3 & \text{si } U \in I_3 \end{cases}$$

En termes de fonction sur Ω,

$$\forall \omega \in \Omega, X(\omega) = \begin{cases} x_1 & \text{si } U(\omega) \in I_1 \\ x_2 & \text{si } U(\omega) \in I_2 \\ x_3 & \text{si } U(\omega) \in I_3 \end{cases}$$

On a $\{X = x_1\} = \{U \in I_1\}$ et donc $\mathbb{P}(X = x_1) = \mathbb{P}(U \in I_1) = p_1$, etc...

En Python cela donne la fonction suivante

```
def VA3finie(p):
    """
    VA3finie(p): liste de 3 nombres positifs sommant à 1
    retourne un tirage aléatoire d'un nombre entre 0 et 2 avec
    la distribution p
    P(VAfinie(p)=k)=p[k]
    """
    u=np.random.rand() #tire un nombre uniformement entre 0 et 1
    if(u<=p[0]):
        return 0
```

```
    if(u<=p[0]+p[1]):
        return 1
    if(u<=p[0]+p[1]+p[2]):#RQ: ne sert à rien, automatiquement vrai si p correct
        return 2
    return -1 #On n'arrive jamais à ce point si p correct
```

Pour finir, généralisons au cas de K valeurs. Si X prend un nombre fini de valeurs numériques ordonnées $x_1 < x_2 < \cdots < x_K$, avec probabilités p_1, p_2,..p_K, ce mécanisme peut s'exprimer à l'aide de la fonction de répartition de X. Si F (ne dépendant que de la distribution connue de X) est définie par

$$\forall x \in \mathbb{R}, F(x) := \mathbb{P}(X \leq x)$$

(On pose $x_0 := -\infty$, $F(x_0) := 0$)

$$X = x_k \text{ si } U \in I_k := [F(x_{k-1}), F(x_k)[$$

Ce qui donne en Python :

```
def VAfinie(p):
    """
    VAfinie(p): p liste de nombres positifs sommant à 1
    retourne un tirage aléatoire d'un nombre entre 0 et len(p)-1 avec
    la distribution p
    P(VAfinie(p)=k)=p[k]
    """
    u = np.random.rand() #tire un nombre uniformement entre 0 et 1
    q = 0.0
    for y in range(len(p)) :
        q += p[y]
        if (u<=q) :
            return y
    return -1 #si erreur!!!
```

3.2 Probabilités et combinatoire : variables aléatoires non numériques

La combinatoire permet de dénombrer des objets mathématiques satisfaisant certaines propriétés, qu'il s'agisse de nombres, d'applications, de parties d'un ensemble ou de p-uplets d'éléments pris dans un ensemble.

On rappelle rapidement le vocabulaire et les quelques résultats élémentaires devant être connus. Soit $\mathscr{X}$ un ensemble fini de cardinal $n > 0$, $p \in \mathbb{N}$

- Une permutation de $\mathscr{X}$ est une bijection $\mathscr{X} \to \mathscr{X}$. On note $\mathfrak{S}(X)$ l'ensemble de toutes les permutations de $\mathscr{X}$. Il y a $n!$ permutations de $\mathscr{X}$, $n! = \#\mathfrak{S}(X)$. C'est aussi le nombre de bijections $\{1,\ldots,n\} \to \mathscr{X}$, *i.e.* de numérotations des éléments de $\mathscr{X}$.
- Si $p > 0$, une p-liste [16] à valeurs dans $\mathscr{X}$ est un p-uplet d'éléments de $\mathscr{X}$. Il y a n^p p-listes à valeurs dans $\mathscr{X}$. Il s'agit aussi du nombre d'applications $\{1,\ldots,p\} \to \mathscr{X}$.

16. Si $p = 0$, la notation n'a pas vraiment de sens (0-uplet, kezako ?), il y a par contre–c'est de la pure logique–$n^0 = 1$ application $\emptyset \to \mathscr{X}$, celle dont le graphe est l'ensemble vide.

— Si $p > 0$, une p-liste sans répétition à valeurs dans $\mathscr{X}$ est un p-uplet d'éléments de $\mathscr{X}$ dont les composantes sont deux à deux distinctes. Si $p > n$, il n'y a aucune p-liste sans répétition à valeurs dans $\mathscr{X}$. Si $p < n$, il y en a–nombre d'arrangements– $A_n^p = n.(n-1).\dots.(n-p+1) = \frac{n!}{(n-p)!}$. Il s'agit aussi du nombre d'applications *injectives* $\{1,\dots,p\} \to \mathscr{X}$.

— Une p-combinaison à valeurs dans $\mathscr{X}$ est un p-uplet *non ordonné* d'éléments distincts de $\mathscr{X}$. Il y a $\binom{n}{p}$ p-combinaisons à valeurs dans $\mathscr{X}$. Il s'agit aussi du nombre de parties de $\mathscr{X}$ comportant p éléments ou du nombre d'applications $\mathscr{X} \to \{0,1\}$ dont la somme des valeurs prises vaut p. On a

$$\forall p \in \{0,\dots,n\}, \binom{n}{p} = \frac{A_n^p}{p!} = \frac{n!}{(n-p)!p!}$$

Lorsque l'on calcule une probabilité d'événement, on a plus que souvent recours à des décompositions de l'événement en union d'événements élémentaires incompatibles et à la propriété d'additivité. Les dénombrements servent à évaluer le nombre de tels événements élémentaires ; connaitre le nombre de termes d'une somme est toujours utile pour évaluer cette somme.

D'un point de vue légèrement décalé, la modélisation d'un problème peut faire intervenir de façon naturelle des variables aléatoires non numériques, typiquement *uniformément distribuées sur un ensemble fini* composé d'ensembles ou d'autres objets combinatoires. La connaissance du nombre d'objets de cet ensemble satisfaisant telle ou telle propriété est un moyen d'évaluer une probabilité par une formule du type

$$\frac{\text{Nbre de configurations favorables}}{\text{Nbre de configurations total}}$$

Les exercices suivants illustrent ce type de calculs.

Exercice 14.— Une urne contient N boules numérotées de 1 à N. On tire, sans remise (ou simultanément, ce qui revient au même) n boules et on note X le plus grand des numéros tirés.
1. Donner un ensemble raisonnable dans lequel X prend ses valeurs.
2. En calculant par dénombrements, donner la loi de X.
3. Déterminer l'espérance de X.

⋆⋆⋆

Dans l'exercice suivant, ne traiter que la partie C.

Exercice 15.— Extrait du problème 1, G2E 2016.

Partie C

Quand un laboratoire choisit les échantillons

On revient au cas général et on suppose dorénavant que n est un entier naturel supérieur ou égal à 2.

Les n échantillons de l'usine de traitement sont relevés par un laboratoire qui choisit d'analyser une partie de ces échantillons (partie choisie de façon équiprobable, éventuellement réduite à l'ensemble vide ou égale à l'ensemble des n échantillons).

Pour une semaine donnée, on considère X la variable aléatoire correspondant au nombre d'échantillons analysés.

C.1.a. Rappeler $\#\mathscr{P}(E)$ puis démontrer que :

$$\forall k \in \{0,\dots,n\}, \mathbb{P}(X=k) = \frac{\binom{n}{k}}{2^n}.$$

C.1.b. Quelle est la loi suivie par X ? Quels sont les paramètres de cette loi ? **C.1.c.** Quel est le nombre moyen d'échantillons analysés par ce laboratoire ?

C.2. Calculer le plus petit entier naturel n non nul tel qu'on est certain à strictement plus de 99% qu'au moins un échantillon a été analysé.

Partie D

Quand deux laboratoires choisissent les échantillons

On suppose dorénavant que deux laboratoires L_1 et L_2 choisissent chacun de manière indépendante et équiprobable une partie quelconque des n échantillons. La partie des échantillons analysés par L_1 est notée $\mathscr{P}_1$ et la partie des échantillons analysés par L_2 est notée $\mathscr{P}_2$.

D.1. On note X_1 la variable aléatoire correspondant au nombre d'échantillons analysés par L_1 et B désigne l'événement $\mathscr{P}_1 \subset \mathscr{P}_2$. **D.1.a.** Démontrer que :

$$\forall k \in \{0,\dots,n\}, \mathbb{P}(B|X_1 = k) = \frac{1}{2^k}$$

D.1.b. En déduire $\mathbb{P}(B)$.

D.2. On note Y la variable aléatoire correspondant au cardinal de $\mathscr{P}_1 \cap \mathscr{P}_2$. **D.2.a.** Les variables aléatoires X_1 et Y sont-elles indépendantes ? **D.2.b.** Calculer $\mathbb{P}(Y = i|X_1 = k)$ pour tout $(i,k) \in \{0,\dots,n\}^2$. **D.2.c.** En déduire la loi de Y.

D.3.a. Démontrer que :

$$\forall k \in E, k\binom{n}{k} = n\binom{n-1}{k-1}$$

D.3.b. Déduire enfin des questions précédentes que :

$$\sum_{(\mathscr{P}_1,\mathscr{P}_2)\in\mathscr{P}(E)^2} \#(\mathscr{P}_1 \cap \mathscr{P}_2) = n4^{n-1}.$$

3.3 Variables aléatoires réelles, Espérances

cf. Le polycopié spin-off.

4 Probabilités conditionnelles

Nous n'avons pas encore abordé la question des probabilités conditionnelles, ni même rappelé ce qui se passe dans le cadre des espaces finis. On peut développer cette théorie directement dans le cadre général, ce que nous faisons maintenant.

Certains points ne seront abordés qu'en fin d'année, dans le chapitre sur les variables discrètes, avec notamment un renforcement « dénombrable » de la formule des probabilités totales énoncée en termes de probabilités conditionnelles.

4.1 Définition

Evaluer la probabilité conditionnelle d'un événement A relativement à un événement B, c'est évaluer la proportion de configurations satisfaisant à la fois les événements A et B parmi celles satisfaisant l'événement B. En formule, si $\mathbb{P}(B) > 0$,

$$\underbrace{\mathbb{P}(A|B)}_{\text{probabilité de } A \textit{ sachant } B} := \frac{\mathbb{P}(A \text{ et } B)}{\mathbb{P}(B)}$$

Une remarque de notation[17] importante : « A sachant B » ou « $(A|B)$ » **n'est pas** un événement, ça n'est pas défini. La notation $\mathbb{P}(A|B)$ ne **signifie pas** que l'on applique la probabilité $\mathbb{P}$ à $A|B$, qui n'a pas de sens ! Pour être plus clair et syntaxiquement correct, on devrait toujours dire « la probabilité, sachant B, de A ».

Proposition 20. *Si $(\Omega, \mathscr{T}, \mathbb{P})$ est un espace probabilisé et $B \in \mathscr{T}$ vérifie $\mathbb{P}(B) > 0$ alors l'application $\mathbb{P}_B$ (notée aussi $\mathbb{P}(.|B)$, la probabilité sachant B) définie par*

$$\forall A \in \mathscr{T}, \mathbb{P}_B(A) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

est une probabilité et $(\Omega, \mathscr{T}, \mathbb{P}_B)$ est un espace probabilisé.

Important ! ! ! Il est crucial de ne pas confondre $\mathbb{P}$, la probabilité « première » et les probabilités conditionnelles du type $\mathbb{P}(.|B)$ qui s'en déduisent. Quand on calcule une probabilité d'événements de type $\mathbb{P}(A)$, c'est « avant » de disposer de la moindre information, les probabilités conditionnelles sont par contre calculées alors que l'on dispose d'informations supplémentaires.

On peut réfléchir au problème suivant : Comparer la « probabilité que je gagne au loto ce samedi sachant que j'ai joué une seule grille simple » avec la « probabilité que je gagne au loto ce samedi sachant que j'ai joué une seule grille simple et que mon petit doigt (qui ne se trompe jamais), m'a dit que la somme des nombres tirés à ce tirage est inférieure ou égale à 61 ».

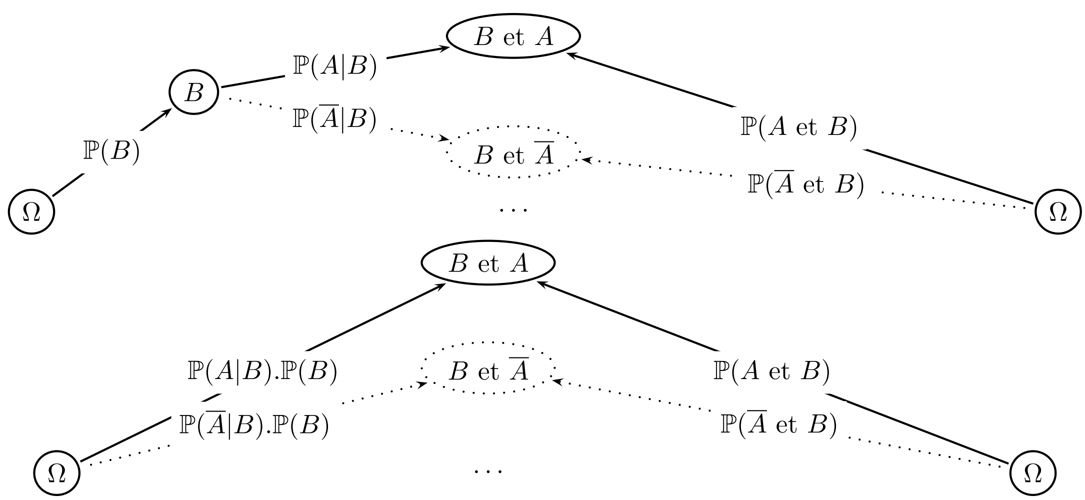


FIGURE 9 – Définition de la probabilité conditionnelle illustrée par deux arbres de choix équivalents. Chaque noeud est un événement. On indique sur la flèche la probabilité conditionnelle du but sachant la source.

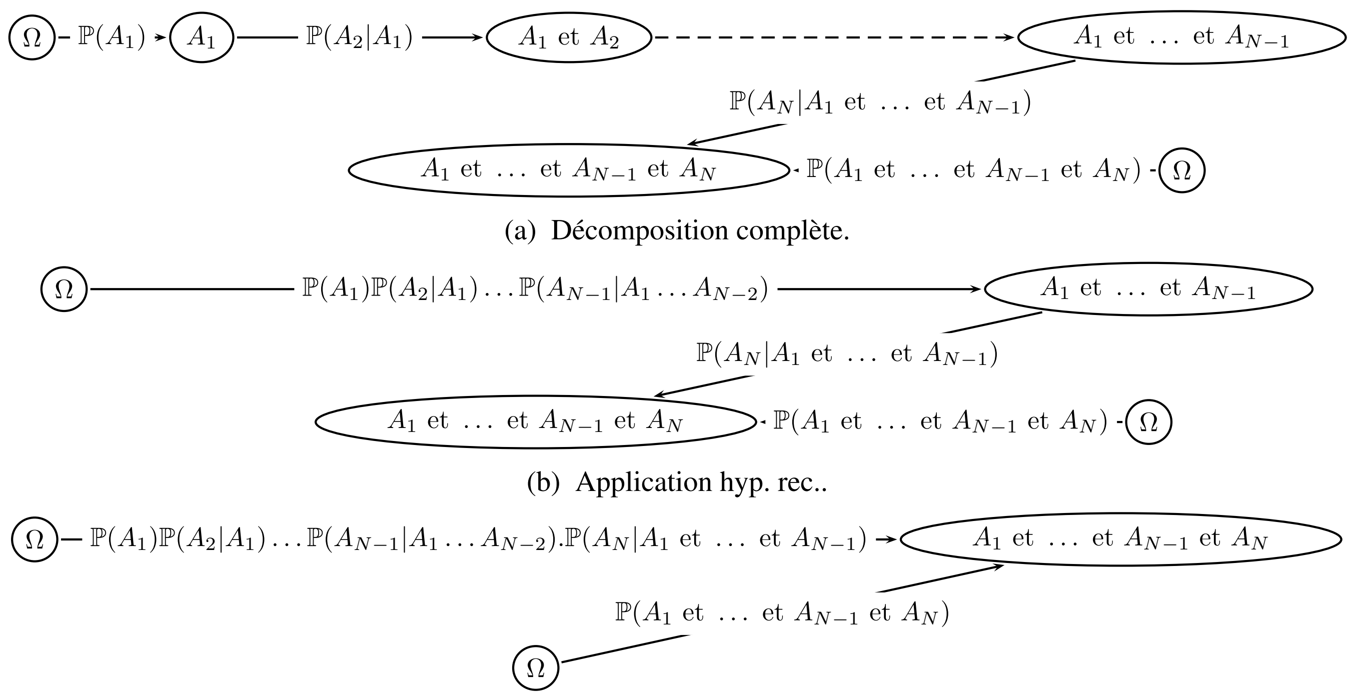
D'un point de vue pratique, donner des probabilités conditionnelles est souvent une façon de spécifier la modélisation à un niveau fin. C'est à dire de spécifier la loi d'une v.a. de modélisation en décomposant suivant la survenue de certains événements.

17. On trouve aussi la notation $\mathbb{P}_B(A)$

4.2 La formule des probabilités composées

Si $A_1,\dots,A_N$ sont $N\geq 2$ événements et $\mathbb{P}(A_1 \text{ et } \dots \text{ et } A_{N_1})\neq 0$ alors

$$\mathbb{P}(A_1 \text{ et } \dots A_N)=\mathbb{P}(\bigcap_{n=1}^{N} A_n)=\mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\dots\mathbb{P}(A_N|A_1 \text{ et } \dots \text{ et } A_{N-1})$$



(c) Application déf. proba. conditionnelle.

FIGURE 10 – La démonstration par récurrence de la formule des probabilités composées.

Exercice 16.— Dans une population de $2.N$ mouches comportant autant de femelles que de mâles, on extrait les individus les uns après les autres jusqu'à obtention d'un couple. On appelle T le nombre de tirages effectués. Donner la loi de T.

Indication:Poser $X_n=1$ si une femelle est tirée au tirage n, $X_n=0$ sinon, conditionner suivant la valeur de X_1 et utiliser la formule des proba. composées.

4.3 Un exemple simple, le BAYES du pauvre

Pour illustrer la signification de ces probabilités conditionnelles, supposons que nous disposions d'une population de 36 individus, composée de $\frac{1}{3}$ de garçons et de $\frac{2}{3}$ de filles. Parmi les garçons, 33% portent des lunettes alors que c'est le cas pour seulement 25% des filles.

Une question simple est la proportion d'individus portant des lunettes.

On peut régler cette question de deux façons.

La première, naïve, consiste à dénombrer les individus :

Il y a 12 garçons dont 4 portent des lunettes, il y a 24 filles dont 6 portent des lunettes. Il y a donc, sur les 36 individus, 10 individus porteurs de lunettes, ce qui fait une proportion de $\frac{10}{36}=27,8\%$ de porteurs de lunettes dans cette population.

On peut faire ce raisonnement sans connaitre le nombre d'individus. Si N est le nombre total d'individus,

Il y a $\frac{1}{3}.N$ garçons et $\frac{1}{3}.\frac{1}{3}.N$ garçons porteurs de lunettes, il y a $\frac{2}{3}.N$ filles et $\frac{1}{4}.\frac{2}{3}.N$ filles porteuses de lunettes, ce qui fait un total de $(\frac{1}{9}+\frac{1}{6}).N = \frac{5}{18}.N$ de porteurs de lunettes et donc une proportion de $27,8\%$ de la population.

La seconde (en apparence plus savante, mais on fait la même chose ! !). On considère l'expérience aléatoire de tirer au sort (uniformément) un individu I et on considère les événements

$A =$« I porte des lunettes », $B =$« I est une fille » et $C =$« I est un garçon ».

La question est la valeur de $\mathbb{P}(A)$.

L'utilisation d'une phrase entre « » pour désigner un événement peut mener à des ambiguités et est une mauvaise pratique. D'un point de vue de la méthode, il est préférable d'introduire immédiatement des variables aléatoires, p.ex. des *indicatrices*, permettant l'écriture usuelle de ces événements. Soit S et M définies par

$$S = \begin{cases} 0 & \text{si } I \text{ est un garçon} \\ 1 & \text{si } I \text{ est une fille} \end{cases} \text{ et } M = \begin{cases} 0 & \text{si } I \text{ sans lunettes} \\ 1 & \text{si } I \text{ à lunettes} \end{cases}$$

On a alors

$$A = \{M = 1\}, B = \{S = 1\} \text{ et } C = \{S = 0\}$$

et les données du problème s'expriment par les faits suivants

$$\mathbb{P}(B) = \frac{2}{3}, \mathbb{P}(C) = \frac{1}{3}, \mathbb{P}(A|B) = \frac{1}{4} \text{ et } \mathbb{P}(A|C) = \frac{1}{3}$$

On a donc

$$\mathbb{P}(A \text{ et } B) = \mathbb{P}(A|B).\mathbb{P}(B) = \frac{1}{4}.\frac{2}{3}, \mathbb{P}(A \text{ et } C) = \mathbb{P}(A|C).\mathbb{P}(C) = \frac{1}{3}.\frac{1}{3}$$

et, sachant que $A = (A \text{ et } B)$ ou $(A \text{ et } C)$, avec une alternative exclusive, on a

$$\mathbb{P}(A) = \mathbb{P}(A \text{ et } B) + \mathbb{P}(A \text{ et } C) = \frac{5}{18}$$

Ces méthodes donnent évidemment le même résultat, la première a l'avantage de la naïveté et de la simplicité conceptuelle, la seconde ouvre la porte à une possibilité de généralisation dans le cas d'une population infinie.

On peut illustrer le calcul fait par une suite de schémas, *cf.* Fig. 11, à interpréter proprement.

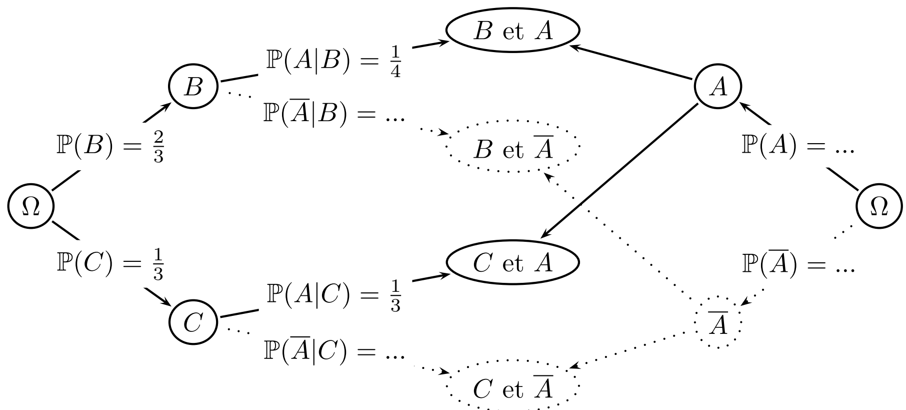
Exercice 17.— Un laboratoire a mis au point un test pour déceler des souris malades. Des essais prouvent que :

— 96 fois sur 100, le test donne un résultat positif quand la souris est effectivement malade.

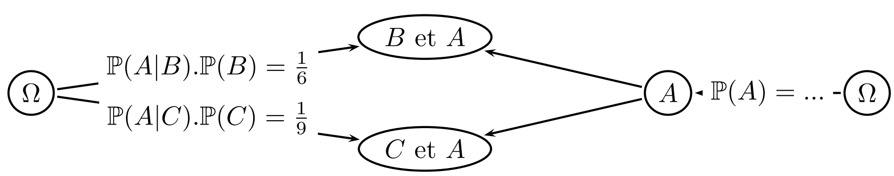
— 94 fois sur 100, le test donne un résultat négatif quand la souris n'est pas malade.

1. Dans une population de souris comprenant 3% de malades, on pratique le test sur une souris choisie au hasard et on constate que le test donne un résultat positif. Quelle est la probabilité que la souris soit malade ? Donner une formule, un ordre de grandeur numérique et commenter le résultat obtenu.

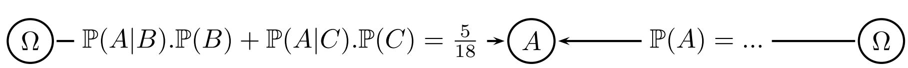
2. La formule de BAYES est la version abstraite de ces calculs. Enoncer cette formule et la redémontrer.



(a) Décomposition complète.



(b) On ne regarde que les branches menant à A, utilisation de la formule d'une probabilité conditionnelle.



(c) Utilisation de l'additivité

FIGURE 11 – Arbre de choix de l'exemple et formule des probabilités totales. Chaque noeud est un événement, que l'on décompose (flèche) en union disjointe d'événements. On indique sur la flèche la probabilité conditionnelle du but sachant la source.

4.4 La formule des probabilités totales

Dans l'exemple précédent, on a illustré la formule fondamentale des probabilités totales : Si $B_1, \dots B_N$ est un système complet d'événements incompatibles (de probabilité > 0), on a, pour tout évenement A,

$$\mathbb{P}(A) = \sum_{n=1}^{N} \mathbb{P}(A|B_n).\mathbb{P}(B_n)$$

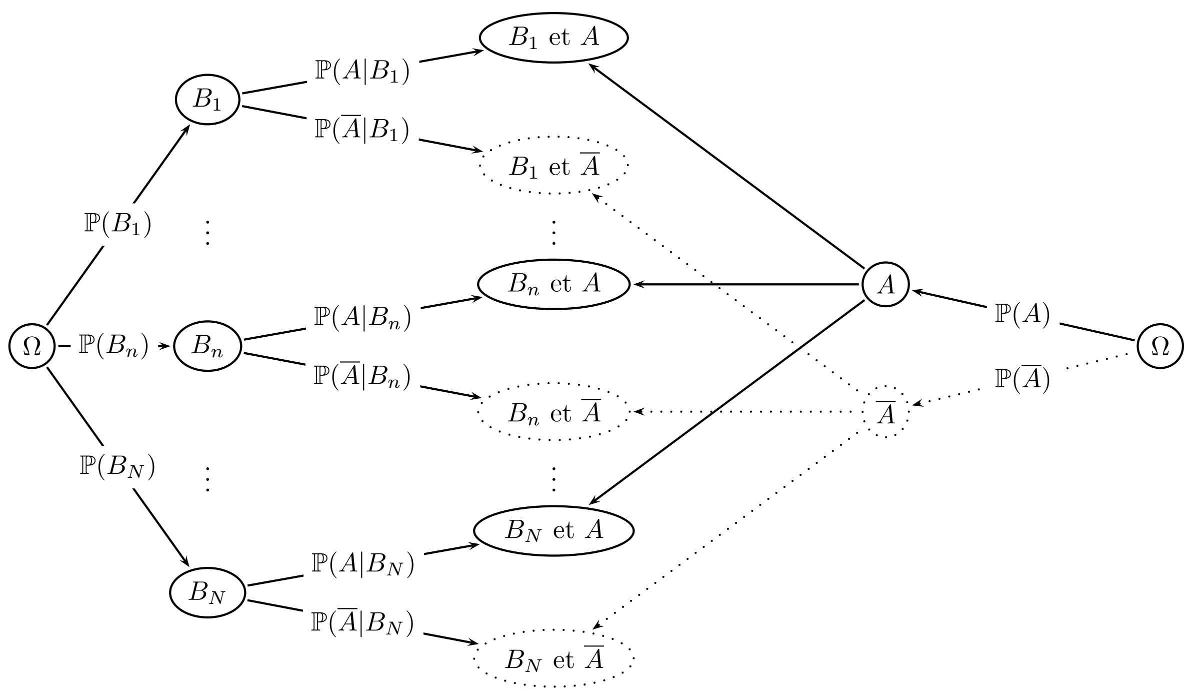


FIGURE 12 – Arbre de choix et formule des probabilités totales. Chaque noeud source est un événement, que l'on décompose (flèches) en union disjointe d'événements. On indique sur la flèche la probabilité conditionnelle du but sachant la source.

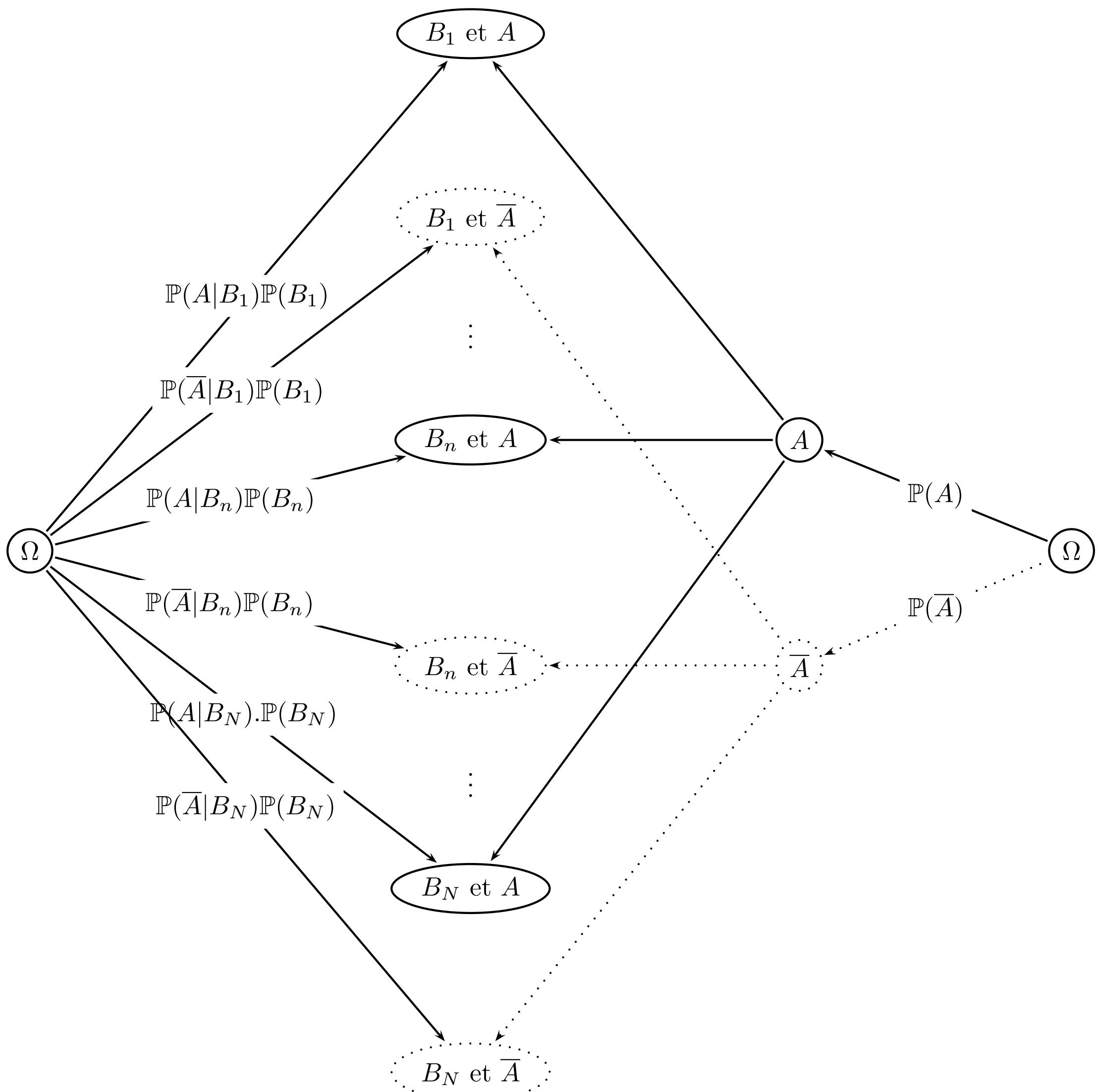


FIGURE 13 – Simplification en utilisant la déf. de proba. conditionnelle.

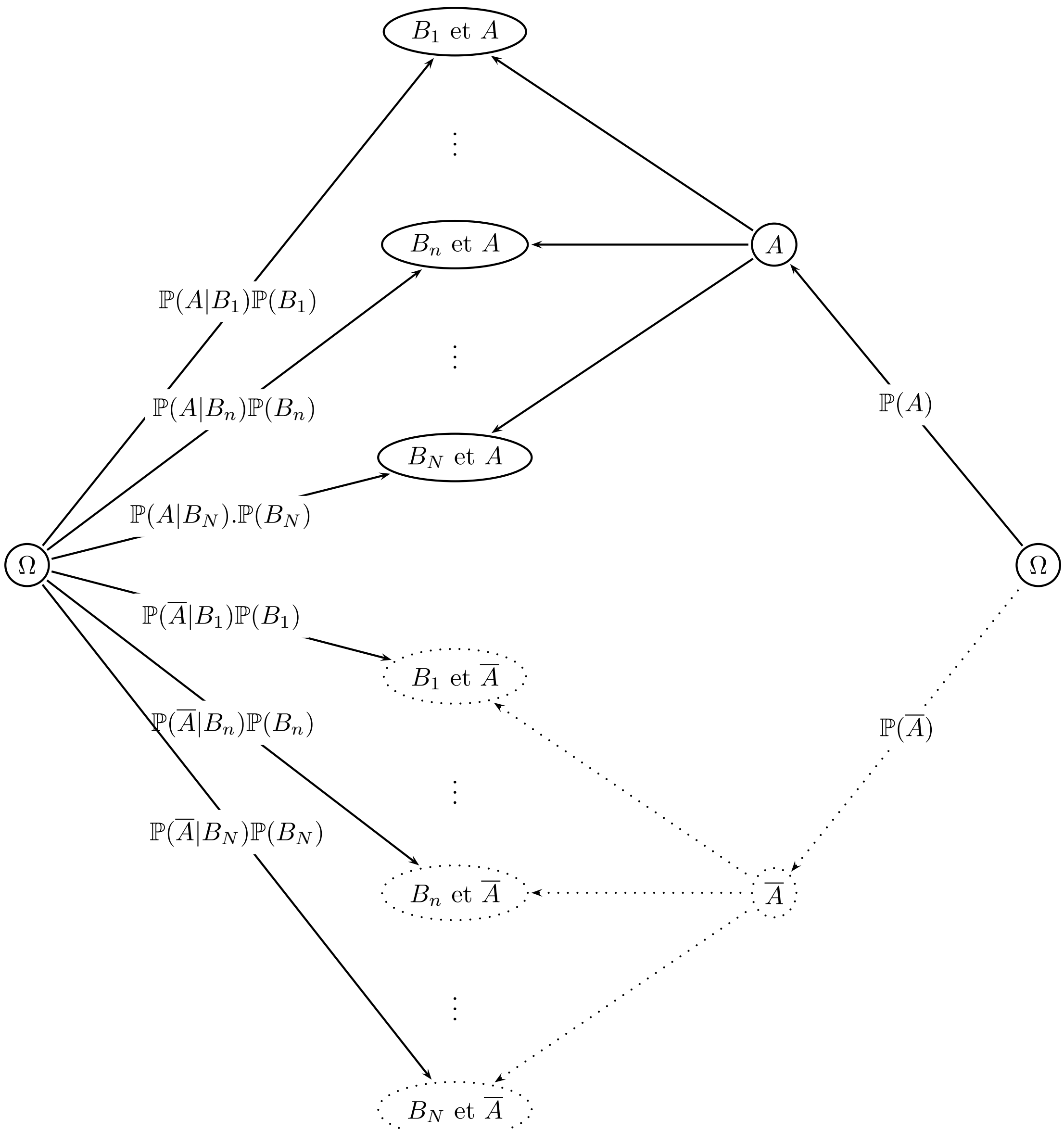


FIGURE 14 – Réordonnancement.

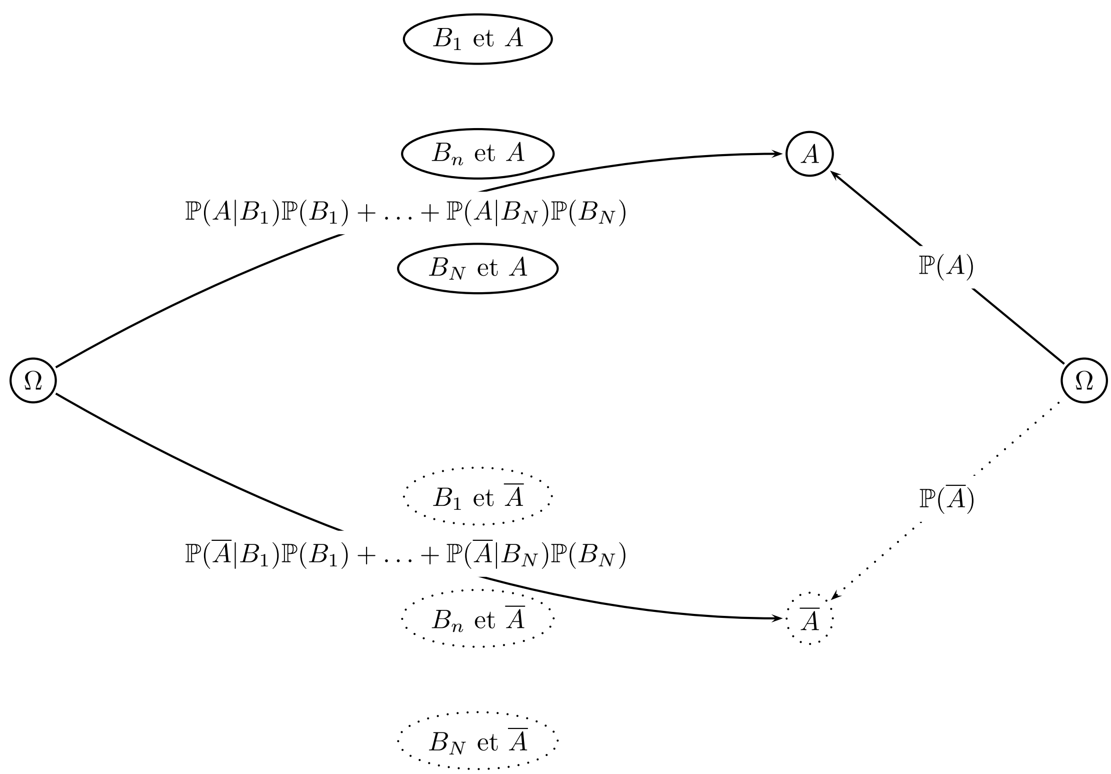


FIGURE 15 – Simplification en utilisant l'axiome d'additivité.

4.5 Espérance conditionnelle (HP), loi conditionnelle

B étant un événement de probabilité > 0, l'application $\mathbb{P}(.|B) : A \mapsto \mathbb{P}(A|B)$ est une *nouvelle* probabilité sur l'ensemble $\mathscr{T}$ des événements. Si X est une variable aléatoire réelle, on peut tenter de calculer son espérance relativement à cette probabilité. Il s'agit de *l'espérance conditionnelle de X conditionnée à B*[18]. Elle est notée $\mathbb{E}(X|B)$. Si X prend un nombre fini de valeurs $\{x_1, \ldots, x_K\} \subset \mathbb{R}$, on a, *par transcription de la formule usuelle d'espérance*, la formule

$$\mathbb{E}(X|B) = \sum_{k=1}^{K} x_K.\mathbb{P}(X = x_k|B)$$

Exercice 18.— Montrer que, si $B_1, \ldots B_N$ est un système complet d'événements incompatibles (de probabilité > 0), on a,

$$\mathbb{E}(X) = \sum_{n=1}^{N} \mathbb{E}(X|B_n).\mathbb{P}(B_n)$$

Dans le même ordre d'idées, on peut définir la « loi de X sachant B » par la donnée de la famille, indexée par les intervalles I de $\mathbb{R}$,

$$\{\mathbb{P}(X \in I|B), I \text{ intervalle de } \mathbb{R}\}$$

A ce propos, traiter la partie D de l'exercice 15.

Exercice 19.— On dispose de deux urnes U_0 et U_1 chacune remplie de boules rouges et vertes, U_0 contient une proportion p_0 de boules rouges, U_1 une proportion p_1. On procède à l'expérience suivante :

— On choisit « au hasard » une des deux urnes ;

— on tire, avec remise, n boules de l'urne choisie

On note U le numéro de l'urne choisie, X le nombre de boules rouges obtenues.
1. Donner la loi conditionnelle de X sachant $U = 0$, la loi conditionnelle de X sachant $U = 1$.
2. Déterminer la loi de X et donner son espérance.

Indépendance d'événements

Si un événement B est de probabilité > 0 et A est un événement, on traduit l'indépendance de A relativement à B par le fait que

$$\mathbb{P}(A) = \mathbb{P}(A|B)$$

Si on lit cette formule à haute voix, on se rend compte qu'elle signifie que la proportion de configurations vérifiant A parmi celles vérifiant B est la proportion de configurations vérifiant A parmi *toutes* les configurations. C'est bien l'idée que l'on se fait de l'indépendance statistique.

On peut réécrire cette relation de façon à traiter le cas $\mathbb{P}(B) = 0$. Ceci symétrise la relation d'indépendance.

Définition 21. *Deux événements A et B sont indépendants si*

$$\mathbb{P}(A \text{ et } B) = \mathbb{P}(A).\mathbb{P}(B)$$

D'un point de vue pratique, l'indépendance de deux événements (ou de variables aléatoires) est souvent une spécification du modèle, un choix, donc, du modélisateur.

18. cette notion n'est pas formellement au programme, on voit cependant qu'elle n'en est pas disjointe...

Exercice 20.— Soit ε une v.a. uniformément distribuée sur $\{-1,+1\}$

1. Soit $U \sim \mathscr{U}_{[0,1]}$. On suppose que U et ε sont indépendantes et on pose $V = \varepsilon.U$. Montrer que $V \sim \mathscr{U}_{[-1,1]}$.

Indication:Pour un intervalle I de $\mathbb{R}$ quelconque, calculer $\mathbb{P}(V \in I)$ en conditionnant sur les valeurs de ε.

2. Soit W une v.a.r telle que

1. Sachant $\{\varepsilon = +1\}$, W est uniformément distribuée sur $[0,1]$,
2. sachant $\{\varepsilon = -1\}$, W est uniformément distribuée sur $[-1,0]$,

2.a. Montrer que $W \sim \mathscr{U}_{[-1,1]}$.

Indication:Comme précédemment, calculer $\mathbb{P}(W \in I)$ en conditionnant sur les valeurs de ε, la difficulté est de traduire les « sachant », *i.e.* le concept de loi conditionnelle

2.b. Montrer que $|W| = \varepsilon.W \sim \mathscr{U}_{[0,1]}$ et que $|W|$ et ε sont indépendantes.

Lois conditionnelles, Simulation informatique

Les modèles probabilistes décrivant un problème physique sont spécifiés via la donnée de variables aléatoires fondamentales et les relations qu'elles entretiennent entre elles :

1. On peut imposer des relations fonctionnelles, par exemple $Z = f(X,Y)$;
2. on peut imposer de l'indépendance ;
3. on peut donner des lois conditionnelles.

C'est ce dernier point que l'on développe maintenant.

Donnons un exemple simple de mélange de populations. Une population de drosophiles est composée à $p\%$ d'individus ailés et $q\%$ d'individus possédant des ailes légèrement atrophiées et $r\%$ des ailes vestigiales, inutiles. On a $p+q+r = 1$. On s'intéresse à leur durée de vie V.

1. Sachant qu'une drosophile est normalement ailée, sa durée de vie V (en heures) suit une loi uniforme [19] sur $[0,100]$;
2. Sachant qu'une drosophile a des ailes atrophiées, sa durée de vie V (en heures) suit une loi uniforme sur $[0,75]$;
3. Sachant qu'une drosophile a des ailes vestigiales, sa durée de vie V (en heures) suit une loi uniforme sur $[0,50]$;

On prend une drosophile au hasard (uniforme sur la population), quelle est la loi de la durée de vie ? Son espérance ? Comment simule-t-on une telle variable aléatoire ?

Commençons par la simulation informatique. On peut définir des variables aléatoires (informatique), `Vn()`, resp. `Va()`, resp. `Vv()`, retournant une durée de vie dans le cas normalement ailé, resp. atrophié, resp. vestigial. Ce sont des tirages uniformes simples.

Pour définir la variable aléatoire (informatique), `V(p)` retournant une durée de vie de drosophile pour notre population, il suffit de

1. Tirer au sort le type d'aile que la drosophile porte en respectant les paramètres p, q, r.
2. suivant les cas, normalement ailé, resp. atrophié, resp. vestigial, retourner la valeur calculée par `Vn()`,resp. `Va()`, resp. `Vv()`.

19. Cette hypothèse n'est pas biologiquement raisonnable, l'exemple ne vaut que pour illustrer le propos

Listing 3 – python/droso-conditionnelles.py

```
import numpy as np

def Vn(): #Simule temps de vie, cas normal. ailé
    return 100*np.random.rand()

def Va(): #Simule temps de vie, cas atrophié
    return 75*np.random.rand()

def Vv(): #Simule temps de vie, cas vestigial
    return 50*np.random.rand()

global aile_n;aile_n=0 #définit les codes
global aile_a;aile_a=1
global aile_v;aile_v=2

def T(p,q): # retourne le type d'aile dans pop. de param. p,q
    u=np.random.rand()
    if u<=p: #Si celle tirée au sort est ailée
        return aile_n
    elif u<=p+q: #Si celle tirée au sort est atrophiée
        return aile_a
    return aile_v #sinon retourne temps pour vestigiale

def V(p,q): #Simule temps de vie, cas d'une population de param. p,q
    t=T(p,q) #On tire le type d'aile
    #Sachant le type, on peut retourner une valeur tirée suivant loi adéquate
    if t==aile_n:
        return Vn() #temps pour normalement ailee
    if t==aile_a:
        return Va() #temps pour atrophiée
    if t==aile_v:
        return Vv() #temps pour vestigiale
```

Pour mener les calculs, introduisons les v.a. appropriées. Soit T la v.a donnant le type de la drosophile tirée au sort. T est à valeurs dans l'ensemble fini $\{\text{norm. ailé}, \text{atrophié}, \text{vestigial}\} = \{n, a, v\}$. On suppose T distribuée à l'aide des proportions dans la population, *i.e.*

$$\mathbb{P}(T = n) = p, \mathbb{P}(T = a) = q \text{ et } \mathbb{P}(T = v) = r$$

Concernant l'espérance, on peut mener un calcul basé sur le principe des espérances conditionnelles, en conditionnant sur la valeur de T

$$\begin{aligned}\mathbb{E}(V) &= \mathbb{E}(V|T = n)\mathbb{P}(T = n) + \mathbb{E}(V|T = a)\mathbb{P}(T = a) + \mathbb{E}(V|T = v)\mathbb{P}(T = v) \\ &= 100p + 75q + 50r\end{aligned}$$

Si on ne connait pas ce principe (HP), on peut tout de même s'en sortir en suivant ce que fait la simulation informatique. Soit U une v.a. uniforme sur $[0,1]$, indépendante de T et G_n, G_a et G_v trois fonctions telles que $Vn = G_n(U)$, resp. $Va = G_a(U)$, resp. $Vv = G_v(U)$, a pour loi la loi conditionnelle de V sachant la drosophile normale ailée, resp. atrophiée, resp. vestigiale. On pose alors (définition d'une fonction sur Ω en trois morceaux)

$$V = \begin{cases} G_n(U) & \text{sur } \{T = n\} \\ G_a(U) & \text{sur } \{T = a\} \\ G_v(U) & \text{sur } \{T = v\} \end{cases}$$

Cette v.a.r V a les lois conditionnelles voulues. On a alors, en remarquant que

$$\mathbb{1} = \mathbb{1}_{\{T=n\}} + \mathbb{1}_{\{T=a\}} + \mathbb{1}_{\{T=v\}}$$

et que

$$\mathbb{1}_{\{T=n\}}.V = \mathbb{1}_{\{T=n\}}.G_n(U),\ \mathbb{1}_{\{T=a\}}.V = \mathbb{1}_{\{T=a\}}.G_a(U) \text{ et } \mathbb{1}_{\{T=v\}}.V = \mathbb{1}_{\{T=v\}}.G_v(U)$$

que

$$\begin{aligned} \mathbb{E}(V) &= \mathbb{E}(\mathbb{1}.V) = \mathbb{E}(\mathbb{1}_{\{T=n\}}.V + \mathbb{1}_{\{T=a\}}.V + \mathbb{1}_{\{T=v\}}.V) \\ &= \mathbb{E}(\mathbb{1}_{\{T=n\}}.V) + \mathbb{E}(\mathbb{1}_{\{T=a\}}.V) + \mathbb{E}(\mathbb{1}_{\{T=v\}}.V) \\ &= \mathbb{E}(\mathbb{1}_{\{T=n\}}.G_n(U)) + \mathbb{E}(\mathbb{1}_{\{T=a\}}.G_n(U)) + \mathbb{E}(\mathbb{1}_{\{T=v\}}.G_n(U)) \end{aligned}$$

Comme, par indépendance de T et U,

$$\mathbb{E}(\mathbb{1}_{\{T=n\}}.G_n(U)) = \mathbb{E}(\mathbb{1}_{\{T=n\}}).\mathbb{E}(G_n(U)) = p.50$$

et similairement pour les autres termes de la somme, on a le résultat attendu.

Concernant la loi de V, le résultat sera plus facilement exprimable à l'issue du prochain chapitre mais on peut donner la formule de transfert générique pour V en effectuant un calcul similaire au précédent, pour une fonction $h : \mathbb{R} \to \mathbb{R}$ positive ou telle que $\mathbb{E}(h(V))$ admet une espérance, on a

$$\mathbb{E}(h(V)) = \int_{-\infty}^{+\infty} h(v).\delta_V(v)\, dv$$

où

$$\delta_V(v) = \begin{cases} \frac{p}{100} + \frac{q}{75} + \frac{r}{50} & \text{si } 0 \le v \le 50 \\ \frac{p}{100} + \frac{q}{75} & \text{si } 50 < v \le 75 \\ \frac{p}{100} & \text{si } 75 < v \le 100 \\ 0 & \text{sinon} \end{cases}$$

En effet, introduisons, comme dans le calcul de $\mathbb{E}(V)$ les variables T, U et les fonctions G_n, G_a et G_v. On a alors, en remarquant que

$$\mathbb{1}_{\{T=n\}}.h(V) = \mathbb{1}_{\{T=n\}}.h(G_n(U)),\ \mathbb{1}_{\{T=a\}}.h(V) = \mathbb{1}_{\{T=a\}}.h(G_a(U)) \text{ et } \mathbb{1}_{\{T=v\}}.h(V) = \mathbb{1}_{\{T=v\}}.h(G_v(U))$$

que

$$\mathbb{E}(h(V)) = \mathbb{E}(\mathbb{1}_{\{T=n\}}.h(G_n(U))) + \mathbb{E}(\mathbb{1}_{\{T=a\}}.h(G_n(U))) + \mathbb{E}(\mathbb{1}_{\{T=v\}}.h(G_n(U)))$$

Comme, par indépendance de T et U,

$$\mathbb{E}(1\!\!1_{\{T=n\}}.h(G_n(U))) = \mathbb{E}(1\!\!1_{\{T=n\}}).\mathbb{E}(h(G_n(U))) = p.\frac{1}{100}\int_0^{100} h(v)\, dv$$

et similairement pour les autres termes de la somme,

$$\mathbb{E}(1\!\!1_{\{T=a\}}.h(G_a(U))) = q.\frac{1}{75}\int_0^{75} h(v)\, dv = q.\frac{1}{75}\int_0^{100} h(v).1\!\!1_{\{0\le v\le 75\}}\, dv$$

et

$$\mathbb{E}(1\!\!1_{\{T=v\}}.h(G_v(U))) = r.\frac{1}{50}\int_0^{50} h(v)\, dv = r.\frac{1}{50}\int_0^{100} h(v).1\!\!1_{\{0\le v\le 50\}}\, dv$$

En regroupant les termes, on obtient

$$\mathbb{E}(h(V)) = \int_0^{100} h(v)\left(\frac{p}{100} + \frac{q}{75}1\!\!1_{\{0\le v\le 75\}} + \frac{r}{50}1\!\!1_{\{0\le v\le 50\}}\right)\, dv$$

Ce qui peut se traduire facilement en le résultat attendu.

5 Des exemples fondamentaux de modèles

5.1 Deux dés indépendants

On considère deux dés à 6 faces, équilibrés. Une partie, c'est deux joueurs Pierre et Iris, et un tirage de ces deux dés.

- Si la somme des deux dés est un nombre pair, Pierre reçoit d'Iris un montant de $\alpha_p \times$ somme,
- Si la somme des deux dés est un nombre impair, Iris reçoit de Pierre un montant de $\alpha_i \times$ somme. La question est :comment choisir α_p et α_i de sorte que le jeu soit équitable.

Il faut définir les variables aléatoires fondamentales du modèle, traduire les hypothèses, donner une version mathématique de la question finale et enfin, faire les calculs. On peut poser comme variables fondamentales X_1 et X_2, les chiffres tirés respectivement par le premier dé et le second.

L'hypothèse sur les dés indique que on peut supposer X_1, X_2 *indépendantes* et *uniformément distribuées sur* $\{1,\ldots,6\}$.

La somme est $S = X_1 + X_2$.

La variable qui nous intéresse est G_i, le gain d'Iris. (Noter que $G_i = -G_p$ où G_p est le gain de Pierre).

On a $G_i = \alpha_i.S.1\!\!1_{\{S\ \text{impair}\}} - \alpha_p.S.1\!\!1_{\{S\ \text{pair}\}}$.

Finalement, on dira que le jeu est équitable si $\mathbb{E}(G_i) = 0$.

Il s'agit de calculer $\mathbb{E}(G_i)$ en fonction de α_i et α_p pour trouver la condition cherchée de jeu équitable. On a, par linéarité de l'espérance,

$$\mathbb{E}(G_i) = \alpha_i.\mathbb{E}(S.1\!\!1_{\{S\ \text{impair}\}}) - \alpha_p.\mathbb{E}(S.1\!\!1_{\{S\ \text{pair}\}})$$

Il suffit de calculer $\mathbb{E}(S.1\!\!1_{\{S\ \text{impair}\}})$ et $\mathbb{E}(S.1\!\!1_{\{S\ \text{pair}\}})$. On a

$$\mathbb{E}(S.1\!\!1_{\{S\ \text{impair}\}}) + \mathbb{E}(S.1\!\!1_{\{S\ \text{pair}\}}) = \mathbb{E}(S) = \mathbb{E}(X_1) + \mathbb{E}(X_2) = 7$$

Il suffit de calculer $\mathbb{E}(S.1\!\!1_{\{S\ \text{pair}\}})$. Calculons ce qu'il faut de la loi de $Y = S.1\!\!1_{\{S\ \text{pair}\}}$ pour évaluer $\mathbb{E}(Y)$. C'est une variable à valeurs dans $\{0,2,4,6,8,10,12\}$.

De la formule de l'espérance, il est *a priori* inutile de calculer $\mathbb{P}(Y=0)$.

On calcule par disjonction de cas puis indépendance.

$$(Y=2)=(X_1=1 \text{ et } X_2=1)$$

et, par indépendance

$$\mathbb{P}(Y=2)=\mathbb{P}(X_1=1).\mathbb{P}(X_2=1)=\frac{1}{36}$$

$$(Y=4)=(X_1=1 \text{ et } X_2=3) \text{ ou } (X_1=2 \text{ et } X_2=2) \text{ ou } (X_1=3 \text{ et } X_2=1)$$

par incompatibilité

$$\mathbb{P}(Y=4)=\mathbb{P}(X_1=1 \text{ et } X_2=3)+\mathbb{P}(X_1=2 \text{ et } X_2=2)+\mathbb{P}(X_1=3 \text{ et } X_2=1)$$

et, par indépendance,

$$\begin{aligned}\mathbb{P}(Y=4) &= \mathbb{P}(X_1=1).\mathbb{P}(X_2=3)+\mathbb{P}(X_1=2).\mathbb{P}(X_2=2)+\mathbb{P}(X_1=3).\mathbb{P}(X_2=1)\\ &= \frac{3}{36}=\frac{1}{12}\end{aligned}$$

$$(Y=6)=(X_1=1 \text{ et } X_2=5) \text{ ou } (X_1=2 \text{ et } X_2=4) \text{ ou } (X_1=3 \text{ et } X_2=3) \text{ ou } \dots$$

par incompatibilité,

$$\mathbb{P}(Y=6)=\mathbb{P}(X_1=1 \text{ et } X_2=5)+\mathbb{P}(X_1=2 \text{ et } X_2=4)+\mathbb{P}(X_1=3 \text{ et } X_2=3)+\dots$$

et, par indépendance,

$$\mathbb{P}(Y=6)=\mathbb{P}(X_1=1).\mathbb{P}(X_2=5)+\mathbb{P}(X_1=2).\mathbb{P}(X_2=4)+\mathbb{P}(X_1=3).\mathbb{P}(X_2=3)+\dots=\frac{5}{36}$$

De même,

$$\mathbb{P}(Y=8)=\frac{5}{36},\ \mathbb{P}(Y=10)=\frac{3}{36} \text{ et } \mathbb{P}(Y=12)=\frac{1}{36}$$

et, par la formule de l'espérance,

$$\mathbb{E}(Y)=\frac{1}{36}(2+4\times 3+6\times 5+8\times 5+10\times 3+12)=\frac{126}{36}=\frac{7}{2}$$

Pour finir, on a

$$\mathbb{E}(S.1\!\!1_{\{S \text{ impair}\}})=7-\mathbb{E}(S.1\!\!1_{\{S \text{ pair}\}})=\frac{7}{2}$$

d'où

$$\mathbb{E}(G_i)=\frac{7}{2}(\alpha_i-\alpha_p)$$

et donc, le jeu est équitable si et seulement si $\alpha_i=\alpha_p$.

Remarquons qu'on pouvait faire une variante du calcul de $\mathbb{E}(Y)$ en utilisant directement la formule de transfert sur la variable couplée (X_1,X_2) à valeurs dans $\{1,\ldots,6\}\times\{1,\ldots,6\}$, sachant que pour ces couples, par indépendance, $\mathbb{P}((X_1,X_2)=(x_1,x_2))=\frac{1}{36}$.

$$\begin{aligned}
\mathbb{E}(S.1\!\!1_{\{S\text{ pair}\}}) &= \sum_{x_1=1}^{6}\sum_{x_2=1}^{6}(x_1+x_2).1\!\!1_{\{x_1+x_2\text{ pair}\}}.\mathbb{P}((X_1,X_2)=(x_1,x_2))\\
&= \sum_{x_1,x_2\text{ même parité}}(x_1+x_2).\mathbb{P}((X_1,X_2)=(x_1,x_2))\\
&= \sum_{x_1,x_2\text{ pairs}}\ldots.+\sum_{x_1,x_2\text{ impairs}}\ldots\\
&= \frac{1}{36}\sum_{p_1=1}^{3}\sum_{p_2=1}^{3}(2p_1+2p_2)+\frac{1}{36}\sum_{p_1=0}^{2}\sum_{p_2=0}^{2}(2p_1+1+2p_2+1)\\
\text{(symétrie)} &= \frac{4}{36}\sum_{p_1=1}^{3}\sum_{p_2=1}^{3}p_1+\frac{2}{36}\sum_{p_1=0}^{2}\sum_{p_2=1}^{3}(2p_1+1)\\
&= \frac{4\times 3}{36}\sum_{p_1=1}^{3}p_1+\frac{2\times 3}{36}\sum_{p_1=0}^{2}(2p_1+1)\\
&= \frac{4\times 3\times 6}{36}+\frac{2\times 3\times 9}{36}=\frac{7}{2}
\end{aligned}$$

Exercice 21.— Soit $n\in\mathbb{N}^*$

1. Pour $t\in\mathbb{R}$, écrire sous forme close

$$P(t)=\sum_{k=0}^{n}k.\binom{n}{k}.t^{k-1}$$

2. Soit X une v.a. suivant la loi $\mathscr{B}\left(n,\frac{1}{2}\right)$.

2.a. Quelle est la probabilité que X soit pair ?

2.b. On pose

$$X_0=\begin{cases}X & \text{si } X \text{ est pair}\\ 0 & \text{si } X \text{ est impair}\end{cases}\quad\text{et } X_1=\begin{cases}0 & \text{si } X \text{ est pair}\\ X & \text{si } X \text{ est impair}\end{cases}$$

Vérifier que $X_0+X_1=X$ et $X_0-X_1=(-1)^X.X$. Que valent $\mathbb{E}(X_0+X_1)$, $\mathbb{E}(X_0-X_1)$?

En déduire les valeurs de $\mathbb{E}(X_0)$ et $\mathbb{E}(X_1)$.

Exercice 22.— Soit $n\in\mathbb{N}^*$

1. Pour $t\in\mathbb{R}$, écrire sous forme close

$$P(t)=\sum_{k=0}^{n}k.t^{k-1}$$

2. Soit X une v.a. suivant la loi $\mathscr{U}_{\{0,\ldots,n\}}$.

2.a. Quelle est la probabilité que X soit pair ?

2.b. On pose

$$X_0=\begin{cases}X & \text{si } X \text{ est pair}\\ 0 & \text{si } X \text{ est impair}\end{cases}\quad\text{et } X_1=\begin{cases}0 & \text{si } X \text{ est pair}\\ X & \text{si } X \text{ est impair}\end{cases}$$

Vérifier que $X_0+X_1=X$ et $X_0-X_1=(-1)^X.X$. Que valent $\mathbb{E}(X_0+X_1)$, $\mathbb{E}(X_0-X_1)$?

En déduire les valeurs de $\mathbb{E}(X_0)$ et $\mathbb{E}(X_1)$.

5.2 Tirer des boules avec remise, modèle binomial

Le passionnant phénomène en considération est le suivant : on dispose d'un sac contenant N fruits (prunes et quetsches), indiscernables au toucher, dont une proportion p de prunes et une proportion q de quetsches. On a

$$p+q=1$$

L'opération élémentaire effectuée est la suivante : on prend un fruit dans le sac, on note son type et on le remet dans le sac.

L'opération complète est la suivante : on effectue n opérations élémentaires successives.

On compte au final le nombre S de prunes. La question est donner la distribution de S.

La méthode combinatoire

La méthode combinatoire consiste à considérer les configurations possibles, à évaluer leur probabilité d'apparition et à les dénombrer.

Une configuration pour cette expérience est une suite de n lettres, P ou Q, correspondant au codage d'un tirage. Par exemple $PPQQ....PQ$ code le fait que les deux premiers tirages ont été des prunes, les deux suivants des quetsches, l'avant dernier une prune et le dernier une quetsche.

Une telle configuration à une probabilité $p^{\text{nbre de } P}.q^{\text{nbre de } Q}$ d'apparition. Il y a 2^n telles configurations. Trouver la loi de S, qui prend clairement ses valeurs dans l'ensemble $\{0,\dots,n\}$, consiste, pour chaque $s \in \{0,\dots,n\}$, à trouver et dénombrer les configurations comportant exactement s fois la lettre P. Il y a exactement $\binom{n}{s}$ telles configurations, toutes ayant même probabilité d'apparition $p^s q^{n-s}$ et donc

$$\forall s \in \{0,\dots,n\}, \mathbb{P}(S=s) = \binom{n}{s} p^s q^{n-s}$$

On dit que S suit la loi binomiale $\mathscr{B}(n,p)$.

La modélisation élémentaire

On va maintenant modéliser cette expérience de façon naturelle, en tirant effectivement les fruits les uns après les autres. La question est : est-ce que cette façon de calculer donne les mêmes résultats que la méthode combinatoire classique ?

Les variables du modèle que l'on construit sont (assez naturellement),

1. X_1, X_2,....X_n, X_k désignant le résultat du k-ième tirage. On convient que $X_k = 1$ si on y a tiré une prune et $X_k = 0$ sinon.
2. X_1, X_2,....X_n sont indépendantes. (remise)
3. $S = X_1 + \dots + X_n$, grâce à la convention précédente.
4. La distribution de chaque X_i est connue. X_i suit une loi de BERNOULLI de paramètre de succès p

Modèle binomial : méthode 1

Notons, pour $k \in \{1,\dots,n\}$, $S_k = X_1 + \dots + X_k$. On a, pour $1 \le k \le n-1$, $S_{k+1} = S_k + X_{k+1}$ et S_k et X_{k+1} sont indépendantes. S_{k+1} prend clairement ses valeurs dans $\{0,\dots,k+1\}$. Pour $s \in \{0,\dots,k+1\}$, on a, « en décomposant cet événement suivant les valeurs de X_{k+1} »,

$$\begin{aligned}\mathbb{P}(S_{k+1}=s) &= \mathbb{P}(S_k=s \text{ et } X_{k+1}=0)+\mathbb{P}(S_k=s-1 \text{ et } X_{k+1}=1)\\ &= \mathbb{P}(S_k=s)\mathbb{P}(X_{k+1}=0)\\ &\quad +\mathbb{P}(S_k=s-1)\mathbb{P}(X_{k+1}=1) \text{ par indep.}\\ &= \mathbb{P}(S_k=s).q+\mathbb{P}(S_k=s-1).p\end{aligned}$$

Par récurrence, à l'aide de la formule du triangle de PASCAL, on montre que, pour tout $0\leq s\leq k$, $\mathbb{P}(S_k = s)=\binom{k}{s}p^s.q^{k-s}$

Démonstration. Pour $s=0$, on a clairement par récurrence que $\mathbb{P}(S_k=0)=q^k$. De même, on a clairement par récurrence que $\mathbb{P}(S_k=k)=p^k$.

(Le terme en facteur de p est nul). On a obtenu une formule de récurrence permettant de calculer la loi de S_{k+1} en fonction de celle de S_k, formule qui présente des airs de parenté avec la formule du triangle de PASCAL. Montrons la formule cherchée par récurrence (finie) : pour $k\in\{1,\dots,n\}$, posons l'hypothèse de récurrence

$$P_k : \forall s\in\{0,\dots,k\}, \mathbb{P}(S_k=s)=\binom{k}{s}p^s.q^{k-s}$$

1. P_1 est vraie car $S_1=X_1$
2. Si P_k est vraie pour un certain entier k, $1\leq k\leq n-1$, on a, pour $s\in\{0,\dots,k+1\}$,
   - soit $s=0$, auquel cas, $\mathbb{P}(S_{k+1}=0)=q^{k+1}=\binom{k+1}{0}p^0.q^{k+1-0}$,
   - soit $s=k+1$, auquel cas, $\mathbb{P}(S_{k+1}=k+1)=p^{k+1}=\binom{k+1}{0}p^{k+1}.q^{k+1-(k+1)}$,
   - soit $s\in\{1,\dots,k\}$, auquel cas,

$$\begin{aligned}\mathbb{P}(S_{k+1}=s) &= \mathbb{P}(S_k=s).q+\mathbb{P}(S_k=s-1).p\\ &= \binom{k}{s}p^s.q^{k-s+1}+\binom{k}{s-1}p^s.q^{k-(s-1)}(\text{ par } P_k \text{ car } 0\leq s,s-1\leq k)\\ &= (\binom{k}{s}+\binom{k}{s-1})p^s.q^{k-(s-1)}\\ &= \binom{k+1}{s}p^s.q^{k+1-s}(\text{ PASCAL})\end{aligned}$$

P_{k+1} est donc vraie et le principe de récurrence affirme que P_k est vraie pour tout entier k. □

Modèle binomial : méthode 2

La méthode que l'on présente maintenant est basée sur le principe du polynôme générateur d'une variable aléatoire N prenant un nombre fini de valeurs entières positives. Ce polynôme caractérise complètement la loi de la variable N.

On définit donc, pour $x\in\mathbb{R}$,

$$\pi_N(x):=\mathbb{E}(x^N)=\sum_{n=0}^{*}\mathbb{P}(N=n)x^n$$

Connaissant le polynôme, on retrouve ses coefficients qui donnent la loi de N et réciproquement, connaissant la loi de N, on peut construire le polynôme.

Si $X \sim \mathscr{B}(p)$, on a, pour $t \in \mathbb{R}$,

$$\pi_X(t) = \mathbb{E}(t^X) = q + p.t = 1 - p + p.t$$

Pour $S = X_1 + \cdots + X_n$, on a, pour $t \in \mathbb{R}$,

$$\begin{aligned}
\pi_S(t) &= \mathbb{E}(t^S) = \mathbb{E}(t^{X_1} \ldots t^{X_n}) \\
&= \mathbb{E}(t^{X_1}) \ldots \mathbb{E}(t^{X_n}) (\text{ indépendance }) \\
&= (q + p.t)^n \\
&= \sum_{s=0}^{n} \binom{n}{s} q^{n-s} p^s . t^s (\text{ Newton})
\end{aligned}$$

On a donc que pour tout $s \in \{0, \ldots, n\}$, $\mathbb{P}(S = s) = \binom{n}{s} q^{n-s} p^s$ et $S \sim \mathscr{B}(n, p)$.

Simulation informatique

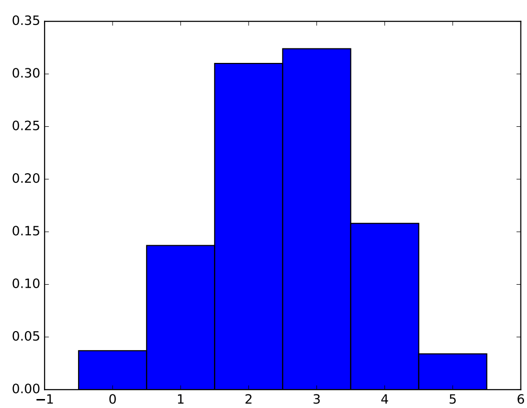


FIGURE 16 – Simulation d'une expérience de BERNOULLI, $n = 5$, $p = 0.5$. simulation sur $N = 1000$ expériences

Utiliser le script `simulation-binomiale.py`.

5.3 Tirer des boules sans remise, modèle hypergéométrique (HP)

Le passionnant phénomène, variante du précédent, en considération est le suivant : on dispose d'un sac contenant $N \geq 2$ fruits (prunes et quetsches), indiscernables au toucher, dont une quantité $p.N$ de prunes et une quantité $q.N$ de quetsches. On a

$$p + q = 1$$

L'opération élémentaire effectuée est la suivante : on prend un fruit dans le sac, on note son type et on *ne le remet pas* dans le sac.

L'opération complète est la suivante : on effectue $n \leq N$ opérations élémentaires successives.

On compte au final le nombre S_n de prunes. La question est donner la distribution de S_n.

Il s'agit de la loi hypergéométrique $\mathscr{H}(n,p,N)$. Vous devez savoir exprimer la loi d'une telle variable aléatoire et connaitre son espérance [20]

$$\mathbb{E}(S_n) = n.p$$

La méthode combinatoire

La méthode combinatoire consiste à considérer les configurations possibles, à évaluer leur probabilité d'apparition et à les dénombrer.

Pour décrire cette expérience, on numérote les objets en plaçant les P en premier et on met un x sous ceux qui ont été tirés, un o sous les autres. Par exemple pour $p.N = 3$, $q.N = 2$, $N = 5$, $n = 3$

$$\left|\begin{array}{c|c|c|c|c} P1 & P2 & P3 & Q4 & Q5 \\ o & x & x & o & x \end{array}\right|$$

signifie que l'on a tiré les prunes 2 et 3 et la quetsche 5. Le nombre de prunes tirées est $S = 2$. Une configuration pour cette expérience est une suite de N lettres, x ou o comportant n x, correspondant au codage d'un tirage.

Le nombre total de configurations est $\binom{N}{n}$. L'expérience aléatoire consiste à tirer uniformément l'une de ces configurations et on peut noter C la v.a aléatoire valant la configuration tirée.

Il est clair que S est le nombre de prunes codé dans la configuration C et que S peut prendre toutes les valeurs s entre $\max(0, n - q.N)$ et $\min(n, p.N)$

On voit que le nombre de configurations comportant s prunes et $n-s$ quetsches est $\binom{p.N}{s}.\binom{q.N}{n-s}$.

On a donc

$$\mathbb{P}(S = s) = \mathbb{P}(C \text{ comporte } s \text{ prunes}) = \sum_{\text{configs } c \text{ comportant } s \text{ prunes}} \mathbb{P}(C = c) = \frac{\binom{p.N}{s}.\binom{q.N}{n-s}}{\binom{N}{n}}.$$

Personnellement, je trouve cette combinatoire un peu douteuse : comment être certain que la succession de tirages est exactement équivalente à un tirage (uniforme) d'une configuration de x et de o ?

Modèle hypergéométrique

On modélise cette expérience de la même façon naturelle que pour le modèle avec remise. On verra que de la sorte on retrouve le résultat combinatoire.

–Les variables du modèle que l'on construit sont (assez naturellement),

1. $X_1, X_2, ..., X_n$, On convient que $X_k = 1$ si on y a tiré une prune au k-ième tirage et $X_k = 0$ sinon.
2. $S_k = X_1 + \cdots + X_k$, le nombre total de prunes obtenues au k-ième tirage. On a $S_{k+1} = S_k + X_{k+1}$.

–La distribution [21] de chaque X_k : $X_1 \sim \mathscr{B}(p)$ et, si $s \in \{0, \ldots, p.N\}$, $k \in \{1, \ldots, n-1\}$,

$$\begin{aligned} \mathbb{P}(X_{k+1} = 1 | S_k = s) &= \frac{p.N - s}{N - k}, \\ \mathbb{P}(X_{k+1} = 0 | S_k = s) &= 1 - \frac{p.N - s}{N - k} = \frac{q.N + s - k}{N - k} \end{aligned}$$

20. La formule de la variance est HP

$$\mathbb{V}(S_n) = \frac{N - n}{N - 1}.n.p.(1 - p)$$

21. sauf pour le premier, elle n'est connue que conditionnellement

–De ceci, on déduit par récurrence sur $1 \leq k \leq N$ que

$$Q_k : S_k \text{ suit la loi } \mathscr{H}(k, p, N)$$

est vraie pour tout k, $1 \leq k \leq N$.

Démonstration. 1. Si $k = 1$, $S_1 = X_1$,

$$\begin{aligned} P(S_1 = 1) &= p = \frac{\binom{p.N}{1}\binom{q.N}{1-1}}{\binom{N}{1}} \\ P(S_1 = 0) &= q = \frac{\binom{p.N}{0}\binom{q.N}{1-0}}{\binom{N}{1}} \end{aligned}$$

2. Supposons Q_k vraie au rang $k \leq N-1$ et déduisons en Q_{k+1}.
   Soit $s \in \{\max(0, k+1-q.N), \ldots, \min(k+1, p.N)\}$.
   — Si $s, s-1 \in \{\max(0, k-q.N), \ldots, \min(k, p.N)\}$, on a, par la formule des probabilités totales & loi conditionnelle puis par Q_k,

$$\begin{aligned} \mathbb{P}(S_{k+1} = s) &= \mathbb{P}(S_k = s \text{ et } X_{k+1} = 0) + \mathbb{P}(S_k = s-1 \text{ et } X_{k+1} = 1) \\ &= \mathbb{P}(X_{k+1} = 0 | S_k = s)\mathbb{P}(S_k = s) + \mathbb{P}(X_{k+1} = 1 | S_k = s-1)\mathbb{P}(S_k = s-1) \\ &= \frac{q.N - (k-s)}{N-k}\mathbb{P}(S_k = s) + \frac{p.N - (s-1)}{N-k}\mathbb{P}(S_k = s-1) \\ &= \frac{1}{(N-k).\binom{N}{k}}\left(\underbrace{(q.N - (k-s))\binom{q.N}{k-s}}_{(k+1-s)\binom{q.N}{k+1-s}}\binom{p.N}{s} + \right. \\ &\qquad \left. + \underbrace{(p.N - (s-1))\binom{p.N}{s-1}}_{s.\binom{p.N}{s}} . \binom{q.N}{k-(s-1)}\right) \\ &= \frac{k+1}{(N-k).\binom{N}{k}}\binom{q.N}{k+1-s}\binom{p.N}{s} = \frac{\binom{q.N}{k+1-s}\binom{p.N}{s}}{\binom{N}{k+1}} \end{aligned}$$

   — On traite de manière similaire les cas extrêmes où s ou $s-1 \notin \{\max(0, k-q.N), \ldots, \min(k, p.N)\}$ pour obtenir que Q_{k+1} est vraie.
3. Par récurrence, Q_k est vraie pour tout k, $0 \leq k \leq N$.

$\square$

Simulation informatique

Utiliser le script `simulation-sansremise.py`.

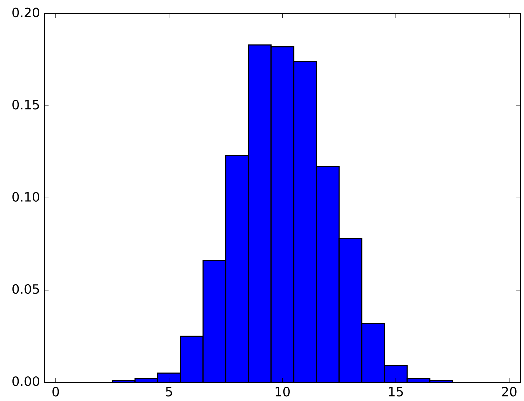


FIGURE 17 – Simulation de tirages sans remise $N = 100$, $n = 50$, $p.N = 20$, simulation sur $NS = 1000$ expériences

Exercice 23.— Le but est de démontrer les formules annoncées d'espérance et de variance pour la loi hypergéométrique. On conserve les notations du cours et notamment le fait que X_k est l'indicatrice d'un tirage de prune à l'étape $k(\leq N)$ et que

$$S_k = X_1 + \cdots + X_k$$

indique le nombre de prunes obtenues à l'étape k. On rappelle que chaque S_k est à valeurs dans $\mathscr{S} = \{0, \ldots, p.N\}$ et que pour un tel s, pour $1 \leq k \leq N-1$

$$\mathbb{P}(X_{k+1} = 1 | S_k = s) = \frac{p.N - s}{N - k}$$

1. On cherche à montrer que pour $k \leq N$, X_k est une variable de BERNOULLI de paramètre p et, conjointement, que $\mathbb{E}(S_k) = k.p$.
1.a. Montrer, en conditionnant sur les valeurs de S_k, que pour $1 \leq k \leq N-1$,

$$\mathbb{P}(X_{k+1} = 1) = \sum_{s \in \mathscr{S}} \mathbb{P}(X_{k+1} = 1 | S_k = s)\mathbb{P}(S_k = s) = \frac{p.N}{N-k} - \frac{1}{N-k}\mathbb{E}(S_k)$$

1.b. Conclure la récurrence.
2. (Difficile) On veut, sur le même principe, montrer que pour $1 \leq k \leq N$,

$$\mathbb{V}(S_k) = \frac{N-k}{N-1} k.p(1-p)$$

2.a. Vérifier que pour $1 \leq k \leq N-1$,

$$S_{k+1} = S_k + X_{k+1} \text{ et } \mathbb{V}(S_{k+1}) = \mathbb{V}(S_k) + \mathbb{V}(X_{k+1}) + 2\,\mathbb{C}ov(S_k, X_{k+1})$$

2.b. Montrer, en conditionnant sur les valeurs de S_k et en utilisant la formule de l'espérance conditionnelle,

$$\mathbb{E}(S_k.X_{k+1}) = \sum_{s \in \mathscr{S}} \mathbb{E}(S_k.X_{k+1} | S_k = s)\mathbb{P}(S_k = s) = \frac{p.N}{N-k}\mathbb{E}(S_k) - \frac{1}{N-k}\mathbb{E}(S_k^2)$$

puis que

$$\mathbb{C}ov(S_k, X_{k+1}) = -\frac{1}{N-k}\mathbb{V}(S_k)$$

2.c. Conclure en rédigeant la récurrence.

5.4 Chaînes de MARKOV

Il s'agit d'une méthode de modélisation pour des phénomènes évoluant stochastiquement dans le temps. On considère le temps discret. Chaque étape dans l'évolution se déduit de la précédente en faisant intervenir le hasard.

Le système que l'on considère peut être dans un certain nombre (fini) d'états. Par exemple, supposons que (pour une raison que j'ignore) un champignon puisse prendre trois couleurs : Rouge (R), Vert(V) ou Bleu(B).

On constate que ce champignon, s'il est d'un certaine couleur à un instant donné, il passe à une autre couleur, au top chrono suivant, avec certaines probabilités

Pour fixer les idées

1. S'il est R, il devient R avec probabilité 0.5, V avec proba 0.3 et B avec proba 0.2,
2. S'il est V, il devient R avec probabilité 0.2, V avec proba 0.4 et B avec proba 0.4,
3. S'il est B, il devient R avec probabilité 0.3, V avec proba 0.3 et B avec proba 0.4,

La question que l'on se pose est sachant qu'au départ $n = 0$ le champignon est R, quelle est la proba qu'à l'instant $n = N(= 10)$, il soit B ?

Modélisation

Si on suit notre méthode de modélisation, on voit que les variables qui importent sont $X_0, X_1, \dots, X_n, \dots, X_N$, X_n donnant la couleur du champignon à l'instant n. On sait que $X_0 = R$ et ce qu'on cherche c'est

$$\mathbb{P}(X_N = R), \mathbb{P}(X_N = V) \text{ et } \mathbb{P}(X_N = B)$$

Ce que donne l'énoncé, c'est la *loi conditionnelle* de X_{n+1} sachant X_n. On a

1. (R) : $\mathbb{P}(X_{n+1} = R|X_n = R) = 0.5$, $\mathbb{P}(X_{n+1} = V|X_n = R) = 0.3$ et $\mathbb{P}(X_{n+1} = B|X_n = R) = 0.2$,
2. (V) : $\mathbb{P}(X_{n+1} = R|X_n = V) = 0.2$, $\mathbb{P}(X_{n+1} = V|X_n = V) = 0.4$ et $\mathbb{P}(X_{n+1} = B|X_n = V) = 0.4$,
3. (B) : $\mathbb{P}(X_{n+1} = R|X_n = B) = 0.3$, $\mathbb{P}(X_{n+1} = V|X_n = B) = 0.3$ et $\mathbb{P}(X_{n+1} = B|X_n = B) = 0.4$.

Pour calculer la loi de X_{n+1}, appliquons la formule des probabilités totales, les définitions des probabilités conditionnelles en « décomposant les événements suivant les valeurs de X_n »,

$$\begin{array}{rcl}
\mathbb{P}(X_{n+1} = R) & = & \mathbb{P}(X_{n+1} = R|X_n = R)\mathbb{P}(X_n = R) + \mathbb{P}(X_{n+1} = R|X_n = V)\mathbb{P}(X_n = V) + \mathbb{P}(X_{n+1} = R|X_n = B)\mathbb{P}(X_n = B) \\
\mathbb{P}(X_{n+1} = V) & = & \mathbb{P}(X_{n+1} = V|X_n = R)\mathbb{P}(X_n = R) + \mathbb{P}(X_{n+1} = V|X_n = V)\mathbb{P}(X_n = V) + \mathbb{P}(X_{n+1} = V|X_n = B)\mathbb{P}(X_n = B) \\
\mathbb{P}(X_{n+1} = B) & = & \mathbb{P}(X_{n+1} = B|X_n = R)\mathbb{P}(X_n = R) + \mathbb{P}(X_{n+1} = B|X_n = V)\mathbb{P}(X_n = V) + \mathbb{P}(X_{n+1} = B|X_n = B)\mathbb{P}(X_n = B)
\end{array}$$

i.e.

$$\begin{array}{rcl}
\mathbb{P}(X_{n+1} = R) & = & 0.5\mathbb{P}(X_n = R) + 0.2\mathbb{P}(X_n = V) + 0.3\mathbb{P}(X_n = B) \\
\mathbb{P}(X_{n+1} = V) & = & 0.3\mathbb{P}(X_n = R) + 0.4\mathbb{P}(X_n = V) + 0.3\mathbb{P}(X_n = B) \\
\mathbb{P}(X_{n+1} = B) & = & 0.2\mathbb{P}(X_n = R) + 0.4\mathbb{P}(X_n = V) + 0.4\mathbb{P}(X_n = B)
\end{array}$$

Ce qui n'est pas sans rappeler le calcul matriciel.

Représentation graphique de chaîne

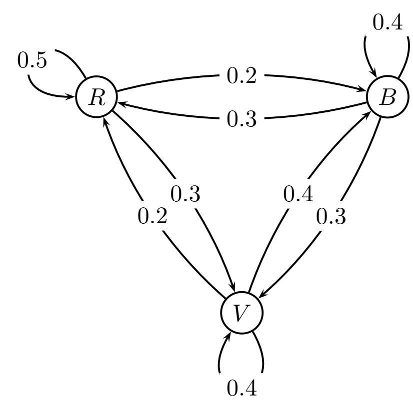


FIGURE 18 – Représentation graphique de l'exemple : on place les états et, sur des flèches, les probabilités de transition d'un état à l'autre

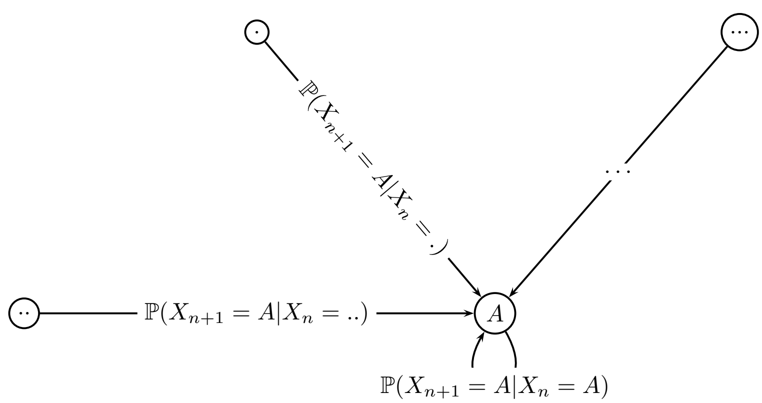


FIGURE 19 – Un noeud, probabilités de l'atteindre, la ligne *A* de la matrice de transition

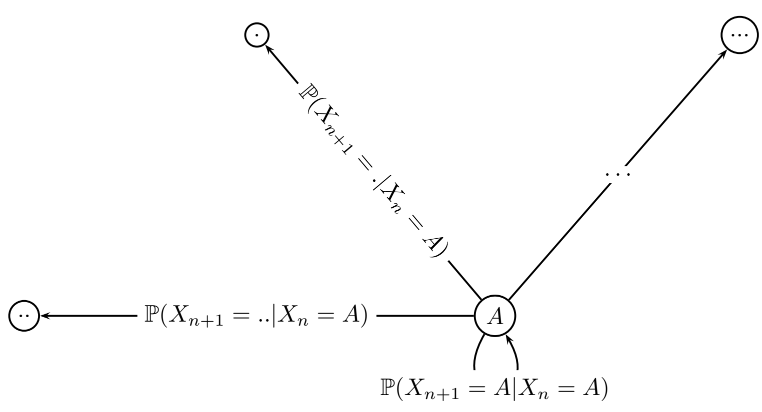


FIGURE 20 – Un noeud, probabilités des destinations, la colonne *A* de la matrice de transition, leur somme vaut 1

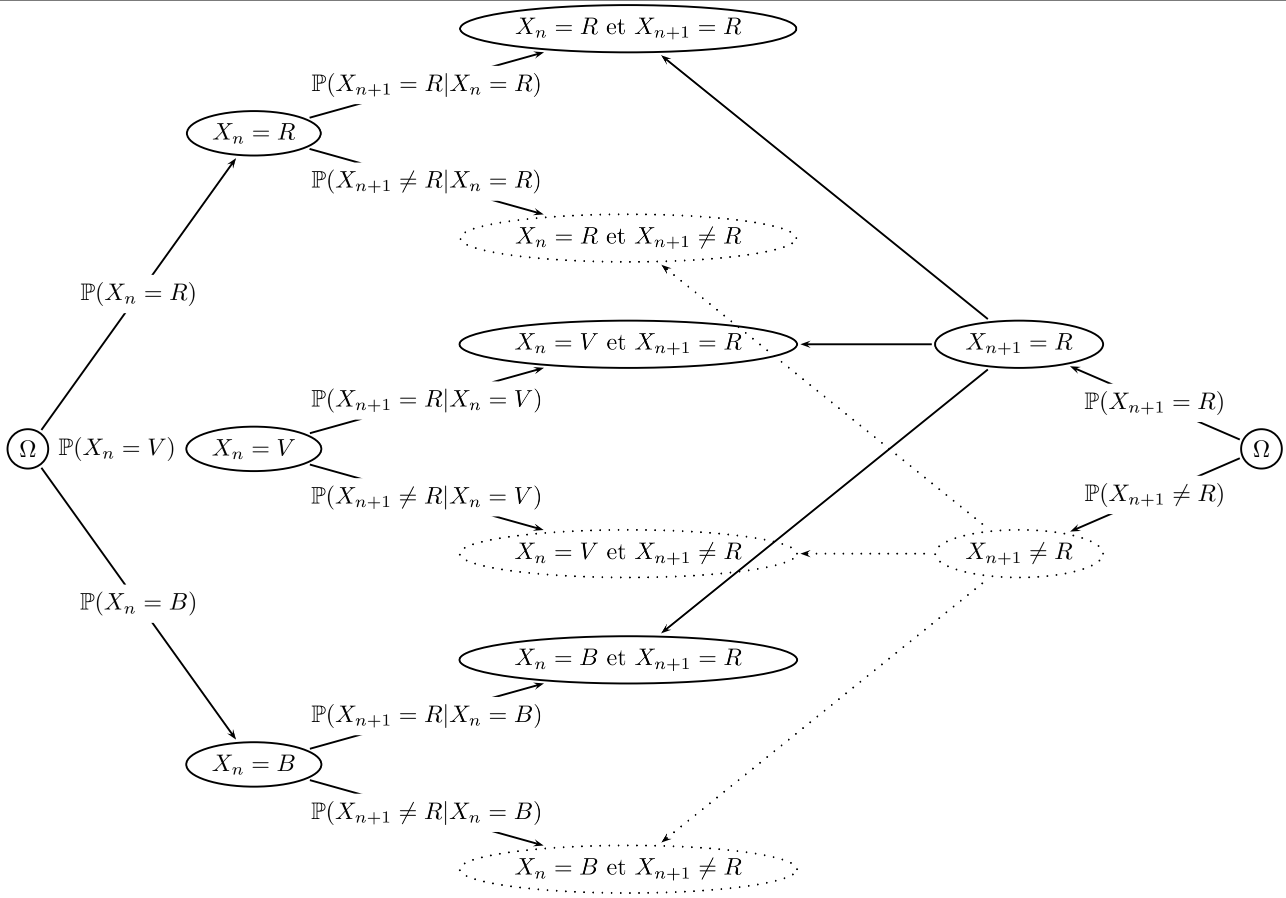


FIGURE 21 – Représentation de l'arbre des choix (initial) pour le calcul de $\mathbb{P}(X_{n+1} = R)$. Cet arbre se traite ensuite comme celui des Fig. 12,13,14 et 15

Calcul matriciel et matrice de transition

Posons

$$A = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}, P_n = \begin{pmatrix} \mathbb{P}(X_n = R) \\ \mathbb{P}(X_n = V) \\ \mathbb{P}(X_n = B) \end{pmatrix}$$

La relation précédente se traduit par

$$\forall n \in \{0, \dots, N-1\}, P_{n+1} = A.P_n$$

La récurrence des suites géométriques donne que $P_N = A^N.P_0$ où le vecteur P_0 traduit la position initiale.

Simulation informatique

Avec

$$A' = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.1 & 0.4 & 0.3 \\ 0.4 & 0.4 & 0.4 \end{pmatrix}$$

Utiliser le script `simulation-markov.py`.

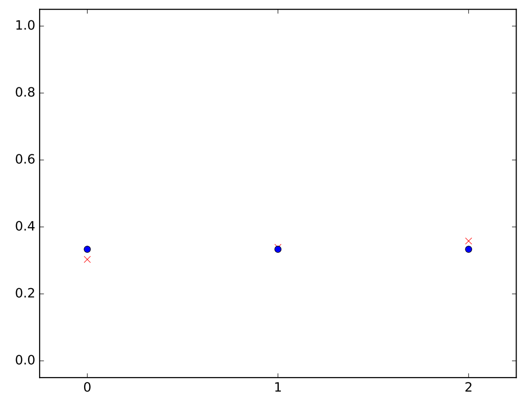


FIGURE 22 – Simulation d'une chaîne de MARKOV, $N = 10$, A et P_0 donné dans le texte. Simulation sur $NS = 100$ expériences. ○=théorique, ×=simulé

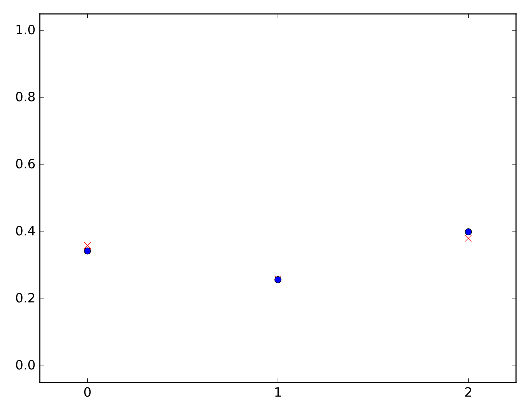


FIGURE 23 – Simulation d'une chaîne de MARKOV, $N = 10$, A', P_0 donné dans le texte. Simulation sur $NS = 1000$ expériences. ○=théorique, ×=simulé

Exercice 24.— On jette n fois une pièce, pile sortant avec probabilité p à chaque tirage. Soit P_n la probabilité pour que le nombre de piles soit pair dans n jets (comptant 0 comme nombre pair). Montrer que

$$P_n = (q-p)P_{n-1} + p$$

où $q = 1 - p$.

En déduire que $P_n = \frac{1}{2}(1 + (q-p)^n)$ pour $n \geq 0$.

Exercice 25.— Le prof de maths, après avoir lu une série de statistiques effrayantes sur les risques de cancer, problèmes cardiovasculaires liés au tabac, décide d'arrêter de fumer; Va-t-il finir par s'arrêter ?

On estime les probabilités suivantes :

1. sachant que cette personne n'a pas fumé le jour J_n, alors la probabilité pour qu'elle ne fume pas le jour suivant J_{n+1} est $0,3$;
2. si sachant qu'elle a fumé le jour J_n, alors la probabilité pour qu'elle ne fume pas le jour suivant J_{n+1} est $0,9$;

Exercice 26.— On dispose d'une urne qui contient trois boules numérotées de 0 à 2.

On s'intéresse à une suite d'épreuves définies de la manière suivante :

— On choisit au hasard une boule dans cette urne.

— Si j est le numéro de la boule tirée, on enlève de l'urne toutes les boules dont le numéro est strictement supérieur à j. Pour le tirage suivant, l'urne ne contient plus que les boules numérotées de 0 à j.

On considère alors la variable aléatoire réelle X_k égale au numéro de la boule obtenue à la $k^{ème}$ épreuve ($k \geq 0$)

On note alors U_k la matrice unicolonne définie par :

$$U_k = \begin{pmatrix} \mathbb{P}[X_k = 0] \\ \mathbb{P}[X_k = 1] \\ \mathbb{P}[X_k = 2] \end{pmatrix}$$

où $\mathbb{P}[X_k = j]$ est la probabilité de tirer la boule numéro j à la $k^{ème}$ épreuve.

On convient de définir la matrice U_0 par :

$$U_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

1. Préliminaire algèbrique. Soient

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \end{pmatrix} \text{ et } P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

1.a. Calculer P^2 et déterminer l'inverse P^{-1} de la matrice P.

1.b. Montrer que $A = P.D.P^{-1}$ où

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

1.c. En déduire, pour $k \in \mathbb{N}^*$, une formule explicite donnant A^k.

2. **2.a.** Vérifier que $U_1 = A.U_0$.

2.b. Déterminer la loi de X_2. (On pourra montrer que $U_2 = A.U_1$.) Calculer l'espérance et la variance de X_2

3. Par utilisation de la formule des probabilités totales, prouver que pour tout entier naturel $k \geq 1$:

$$U_{k+1} = A.U_k$$

4. Écrire, pour un entier naturel k, U_k en fonction de A et U_0

5. Pour tout k de $\mathbb{N}^*$, donner la loi de X_k et vérifier que l'on a :

$$\lim_{k \to +\infty} \mathbb{P}[X_k = 0] = 1, \quad \lim_{k \to +\infty} \mathbb{P}[X_k = 1] = 0, \quad \lim_{k \to +\infty} \mathbb{P}[X_k = 2] = 0$$

Annexe : variables et fonctions indicatrices

Partie A

Définition et expérimentations.

Si E est un ensemble, A une partie (ou un sous-ensemble) de E, on appelle fonction indicatrice de A la fonction, notée $1\!\!1_A : E \to \mathbb{R}$ définie par

$$\forall x \in E,\ 1\!\!1_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

Plus généralement, si P est une proposition logique (pouvant inclure des variables libres), on pose $1\!\!1_{\{P\}}$ valant 1 si P est vraie et 0 sinon. On a donc,

$$\forall x \in E,\ 1\!\!1_A(x) = 1\!\!1_{\{x \in A\}}.$$

On peut représenter graphiquement une fonction indicatrice d'ensemble sur un diagramme en « patate » via les ensembles de niveau. *i.e.* On colorie avec des couleurs différentes, l'une codant 1, l'autre codant 0 les points de E suivant la valeur prise par la fonction au point.

L'expression $1\!\!1_{\{P\}}$ peut être utilisée pour créer une fonction de toutes les (ou partie des) variables libres de P. Les variables libres dans l'écriture d'une proposition P sont les variables (identifiants de variable) qui ne sont pas "internes" ou muettes dans la proposition, *i.e.* quantifiées par un signe $\forall$ ou $\exists$. On signale les variables libres en les plaçant entre parenthèses derrière le symbole de la proposition.

En exemples :

A.1. si P désigne la proposition : $\exists k \in \mathbb{Z}, x = 2k.\pi$ La véracité de P dépend de la valeur de x qui est une variable libre dans l'écriture de P (la variable k est liée à P par l'utilisation d'un quantificateur). On peut considérer la fonction

$$c : x \in \mathbb{R} \mapsto 1\!\!1_{\{P(x)\}} = \begin{cases} 1 & \text{si } x \text{ est multiple entier relatif de } 2.\pi \\ 0 & \text{sinon} \end{cases}$$

Sur cet exemple, montrer que la fonction $x \mapsto \cos(x.c(x))$ vaut constamment 1 ;

A.2. si P désigne la proposition : $x - t \geq 0$ et $t \geq 0$ Les deux variables x et t sont libres dans P. On peut alors considérer la fonction de deux variables réelles :

$$(x,t) \in \mathbb{R}^2 \mapsto 1\!\!1_{\{P(x,t)\}} = \begin{cases} 1 & \text{si } t \leq x \text{ et } t \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Faire une représenation graphique de cette fonction en ensembles de niveau.

On peut aussi, en considérant x comme un paramètre, considérer la fonction d'une variable réelle

$$c_x : t \in \mathbb{R} \mapsto 1\!\!1_{\{P(x,t)\}} = \begin{cases} 1 & \text{si } t \leq x \text{ et } t \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Montrer qu'en fait

$$c_x = \begin{cases} 0 (\text{la fonction nulle sur } \mathbb{R}) & \text{si } x < 0 \\ 1\!\!1_{[0,x]} & \text{si } x \geq 0 \end{cases}$$

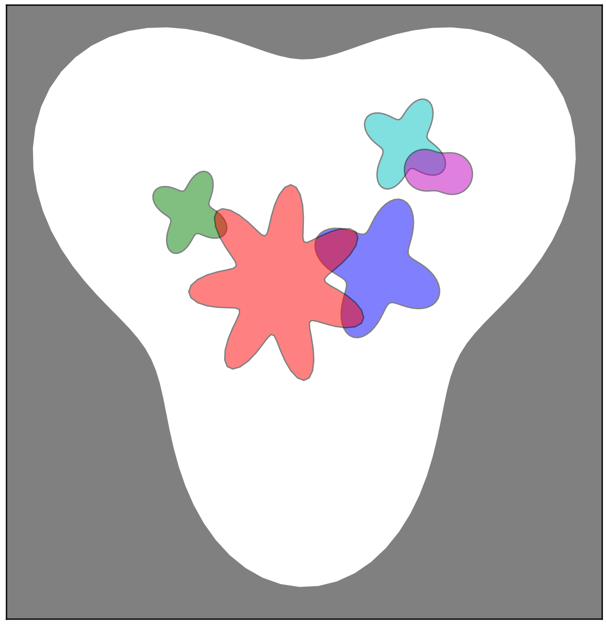


FIGURE 24 – Représentation graphique de fonctions indicatrices. L'univers E est en blanc, chaque « étoile » colorée représente une partie de E. Les nommer $A_1,\dots,A_5$ de gauche à droite, et, en reproduisant partiellement le dessin, représenter successivement $1\!\!1_{A_1}$, $1\!\!1_{A_1\cup A_2\cup A_3}$, $1\!\!1_{A_2\cap A_3}$, $1\!\!1_{E\setminus(A_1\cup A_5)}$

A.3. Traiter la figure 24.

Partie B

Définir des fonctions, discuter dans une intégrale

B.1. Représenter graphiquement les fonctions

B.1.a. $f_1 : x\in\mathbb{R}\mapsto 1\!\!1_{\{x\in[0,1]\}}$, $f_2 : x\in\mathbb{R}\mapsto 1\!\!1_{\{x\in[0,1[\}}$, $f_3 : x\in\mathbb{R}\mapsto 1\!\!1_{\{x\in]-1,+\infty[\}}$, $f_4 : x\in\mathbb{R}\mapsto 1\!\!1_{\{x\in]-\infty,2[\}}$ ainsi que f_3+f_4, $\max(f_3,f_4)$, $f_3+f_4-f_3.f_4$, $f_3.f_1$, $\min(f_3,f_1)$,..

On prend comme convention que la valeur ZERO$=0$ prise par une fonction indicatrice est « super absorbante ». Cela signifie que dans une formule, un problème de définition du type non def $\times$ ZERO est levé et vaut 0. Par exemple, pour $x=-1$,

$$\ln(x).1\!\!1_{\{x\in]0,+\infty[\}}=0$$

B.1.b. $g_1 : x\mapsto e^{-x}1\!\!1_{\{x\in]0,+\infty[\}}$, $g_2 : x\mapsto \ln x.1\!\!1_{\{x\in]1,+\infty[\}}+\sin(x).1\!\!1_{\{x\in]-\infty,-1[\}}$ Que vaut le produit $g_1.g_2$?

B.2. Si $f:\mathbb{R}\to\mathbb{R}$ est une fonction continue, sauf, peut-être, aux points $x_1<x_2\cdots<x_N$, on dit que l'intégrale généralisée $\int_{-\infty}^{+\infty} f(x)\,dx$ est convergente si *chacune* des intégrales généralisées

$$\int_{-\infty}^{x_1} f(x)\,dx,\ \int_{x_1}^{x_2} f(x)\,dx,\ \dots,\ \int_{x_{N-1}}^{x_N} f(x)\,dx \text{ et } \int_{x_N}^{+\infty} f(x)\,dx$$

est convergente. En cas de cas convergence, on définit la valeur de cette intégrale comme

$$\int_{-\infty}^{+\infty} f(x)\,dx := \int_{-\infty}^{x_1} f(x)\,dx+\int_{x_1}^{x_2} f(x)\,dx+\cdots+\int_{x_{N-1}}^{x_N} f(x)\,dx+\int_{x_N}^{+\infty} f(x)\,dx.$$

Calculer les intégrales suivantes (dire un mot de leur nature). Discuter suivant les valeurs de x au besoin.

$$I_1 = \int_{-\infty}^{+\infty} 1\!\!1_{\{t\in]0,1[\}} \frac{dt}{\sqrt{1-t^2}},\ I_2 = \int_{-\infty}^{+\infty} 1\!\!1_{\{t\in]0,+\infty[\}} t^2.e^{-3t}\ dt$$

$$I_3(x) = \int_{-\infty}^{+\infty} 1\!\!1_{\{0\leq t\leq 1\}}.1\!\!1_{\{0\leq x-t\leq 1\}}\ dt,\ I_4(x) = \int_{-\infty}^{+\infty} 1\!\!1_{\{0\leq t\}} e^{-t}.1\!\!1_{\{0\leq x-t\}} e^{-(x-t)}\ dt$$

$$I_5(x) = \int_{-\infty}^{+\infty} 1\!\!1_{\{t\leq 0\}} e^{+t}.1\!\!1_{\{0\leq x-t\}} e^{-(x-t)}\ dt$$

B.3. Expliquer en quoi [22] l'additivité de l'intégrale généralisée de fonctions sur $\mathbb{R}$ implique formellement la relation de CHASLES.

Partie C

Généralités ensemblistes

On devra, pour chaque question, représenter la situation typique de la question sur un diagramme « en patate ».

C.1.a. Montrer si A et B sont deux parties de E, on a équivalence entre

1. $A \subset B$
2. $1\!\!1_A \leq 1\!\!1_B$, inégalité à prendre au sens des fonctions, *i.e.* $\forall x \in E$, $1\!\!1_A(x) \leq 1\!\!1_B(x)$.

C.1.b. Comment caractériser l'égalité de parties de E avec l'égalité de leurs fonctions indicatrices ?

C.1.c. Montrer que si $f : E \to \mathbb{R}$ est une fonction ne pouvant prendre que les valeurs 0 et 1 alors f est fonction indicatrice d'une certaine partie de A que l'on écrira à l'aide de f. Quelle partie de E a pour fonction indicatrice la fonction nulle, que l'on notera 0 ? Quelle partie de E a pour fonction indicatrice la fonction constante égale à 1 sur E, que l'on notera $1\!\!1$?

C.2. On suppose que A, B, etc... sont des parties de E.

C.2.a. Démontrer, par disjonction de cas, rapidement les relations entre fonctions $E \to \mathbb{R}$,

$$1\!\!1_{\overline{A}} = 1\!\!1 - 1\!\!1_A,\ 1\!\!1_{A\cap B} = 1\!\!1_A.1\!\!1_B = \min(1\!\!1_A, 1\!\!1_B),\ 1\!\!1_{A\cup B} = 1\!\!1_A + 1\!\!1_B - 1\!\!1_A.1\!\!1_B = \max(1\!\!1_A, 1\!\!1_B).$$

C.2.b. En se servant uniquement des relations précédentes, montrer que

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \text{ et } A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

C.2.c. On rappelle que $A\Delta B$, la différence symétrique des deux parties A et B est définie par

$$A\Delta B = (A \cup B) \setminus (A \cap B)$$

Montrer [23] que $1\!\!1_{A\Delta B} = (1\!\!1_A - 1\!\!1_B)^2$. En déduire, « en une ligne de calcul », que $A \cap (B\Delta C) = (A \cap B)\Delta(A \cap C)$.

22. On pourra écrire pour $a \leq b \leq c$, $1\!\!1_{]a,c[} = 1\!\!1_{]a,b]} + 1\!\!1_{[b,c[}$
23. On remarquera que $1\!\!1_A^2 = 1\!\!1_A$

C.3. On se donne deux familles finies de nombres réels $(f_i)_{i\in\{1,\ldots,I\}}$ et $(g_j)_{j\in\{1,\ldots,J\}}$, deux familles de parties de E, $(A_i)_{i\in\{1,\ldots,I\}}$ et $(B_j)_{j\in\{1,\ldots,J\}}$ et on définit les fonctions f et $g : E \to \mathbb{R}$ par

$$f = \sum_{i=1}^{I} f_i.1\!\!1_{A_i} \text{ et } g = \sum_{j=1}^{J} g_j.1\!\!1_{B_j}$$

C.3.a. Donner une formule pour le produit $f.g$ où les ensembles $C_{i,j} = A_i \cap B_j$ interviennent. Quel est le nombre maximum de valeurs prises par $f.g$?
C.3.b. Illustrer graphiquement la situation dans le cas où f prend les 3 valeurs distinctes $2,3,5$ et g prend les valeurs 7 et 11. Indication:Prendre pour E le carré unité, pour les A_i des bandes verticales, pour les B_j des bandes horizontales.

Partie D

Applications aux sommes (cf. début BCPST1).

D.1. Montrer que, pour $N \in \mathbb{N}^*$, les coefficients $a_{i,j}$ étant des nombres complexes,

$$\sum_{i=1}^{N}\sum_{j=i}^{N} a_{i,j} = \sum_{1\le i,j\le N} 1\!\!1_{\{1\le i\le j\le N\}} a_{i,j}$$

et en déduire la formule d'interversion de sommes pour la première somme.

Interpréter graphiquement cette formule en plaçant les nombres $a_{i,j}$ dans un tableau carré $N \times N$.
D.2. Appliquer ce principe au type d'interversion de sommes apparaissant naturellement dans la démonstration de la formule suivante. Si $x \in \mathbb{C}$, $N \in \mathbb{N}$,

$$\sum_{n=0}^{N}\sum_{k=-n}^{+n} x^k = \sum_{k=-N}^{+N} (N - |k| + 1)x^k$$

Partie E

POINCARÉ

Soit $N \in \mathbb{N}$, $N \geq 2$. On devra commencer à répondre à chacune de ces questions avec $N = 2$, $N = 3$, $N = 4$.
E.1. Montrer que si $a_1, \ldots a_N$ sont des nombres complexes,

$$\prod_{n=1}^{N}(1 - a_n) = \sum_{P\subset\{1,\ldots,N\}} (-1)^{\#P} \left(\prod_{n\in P} a_n\right)$$

la somme précédente ayant lieu sur toutes [24] les parties P de $\{1,\ldots,N\}$. $\#P$ désigne le cardinal, ou nombre d'éléments, de la partie P. Dans cette somme, combien vaut le terme correspondant à P, la partie vide ? (c'est la convention du produit sur un ensemble vide d'indices)

24. au nombre de 2^N

E.2. En appliquant le principe que, pour une partie A de E, si E est fini, on a

$$\#A = \sum_{x\in E} 1\!\!1_{\{x\in A\}},$$

montrer que si $A_1,\dots,A_N$ est une famille de parties de E, on a

$$\begin{aligned}
\#\left(\cup_{n=1}^N A_n\right) &= \#E - \#\left(\cap_{n=1}^N \overline{A_n}\right)\\
&= \sum_{P\subset\{1,\dots,N\},P\neq\emptyset} (-1)^{(\#P+1)}\#(\cap_{n\in P}A_n)\\
&= \sum_{k=1}^N (-1)^{k+1} \sum_{1\le i_1<\cdots<i_k\le N} \#(A_{i_1}\cap\cdots\cap A_{i_k})
\end{aligned}$$

Combien de termes comporte la somme la plus intérieure dans cette dernière expression ?

E.3. En appliquant le principe que, pour un événement A d'un univers probabilisé $(\Omega,\mathbb{P})$,

$$\mathbb{P}(A) = \mathbb{E}(1\!\!1_A),$$

où $1\!\!1_A$ est la variable aléatoire indicatrice de l'événement A, montrer que si $A_1,\dots,A_N$ est une famille d'événements, on a

$$\begin{aligned}
\mathbb{P}\left(\cup_{n=1}^N A_n\right) &= 1-\mathbb{P}\left(\cap_{n=1}^N \overline{A_n}\right)\\
&= \sum_{P\subset\{1,\dots,N\},P\neq\emptyset} (-1)^{(\#P+1)}\mathbb{P}\left(\cap_{n\in P}A_n\right)\\
&= \sum_{k=1}^N (-1)^{k+1} \sum_{1\le i_1<\cdots<i_k\le N} \mathbb{P}\left(A_{i_1}\cap\cdots\cap A_{i_k}\right)
\end{aligned}$$

E.4. Dénombrement des surjections. On se donne X et Y deux ensembles finis, $\#X \geq \#Y$, et l'on cherche à dénombrer le nombre de surjections de X sur Y.

E.4.a. E.4.a.i. Rappeler ce qu'est une surjection, faire un schéma représentant une application surjective et une application qui ne l'est pas.

E.4.a.ii. Combien y a-t-il d'applications $X \to Y$?

E.4.a.iii. Combien y a-t-il de surjections $X \to Y$ au cas où $\#X = \#Y$?

E.4.b. On note, pour $y \in Y$, A_y l'ensemble des applications de $X \to Y$ qui « évitent » la valeur y, *i.e.*

$$A_y = \{f \in Y^X, \forall x \in X, f(x) \neq y\}$$

E.4.b.i. Représenter un élément de A_y générique, comment construire un tel élément ? Quel est le cardinal de A_y ?

E.4.b.ii. Si $y_1,y_2 \in Y$, $y_1 \neq y_2$, quel est le cardinal de $A_{y_1}\cap A_{y_2}$?

E.4.b.iii. Généraliser au cas de k éléments de Y.

E.4.c. Quel est le lien entre l'ensemble des surjections $X \to Y$ et les ensembles A_y ? Conclure avec la formule de POINCARÉ.

Compléments hors-programme : tribus et probabilités

Exemples graphiques : Carré (HP)

Considérons le carré $\Omega = [0,1[\times [0,1[$ subdivisé en petits carrés $C_{i,j} = \left[\frac{i}{4}, \frac{i+1}{4}\right[\times \left[\frac{j}{4}, \frac{j+1}{4}\right[$, $i, j \in \{0, \ldots, 3\}$, de côté $\frac{1}{4}$ comme sur la figure 25

L'ensemble $\mathscr{C} = \{C_{i,j}, i, j \in \{0, \ldots, 3\}\}$ est un ensemble de parties de Ω. Il comporte 16 éléments. Ce **n'est pas** une tribu de parties de Ω.

Par contre, l'ensemble $\mathscr{T}_0$ où chaque élément est une figure T formée d'une union finie de carrés $C_{i,j}$ est une tribu de parties de Ω. Il comporte 2^{16} éléments. Une probabilité $\mathbb{P}_0$ sur cette tribu est donnée par l'application définie par

$$\forall T \in \mathscr{T}_0, \mathbb{P}_0(T) = \frac{\text{aire}(T)}{\text{aire}(\Omega)} = \text{aire}(T)$$

Sur la figure 25, on a

$$\mathbb{P}_0(T_1) = \frac{4}{16}, \mathbb{P}_0(T_2) = \frac{3}{16}, \mathbb{P}_0(T_1 \cap T_2) = \frac{1}{16} \text{ et } \mathbb{P}_0(\overline{(T_1 \cup T_2)}) = \frac{10}{16}$$

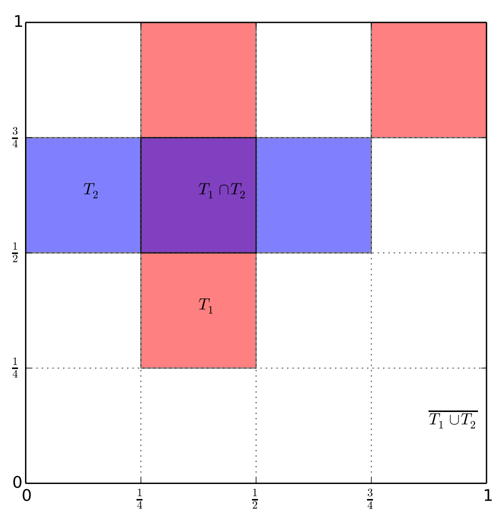


FIGURE 25 – Le carré, T_1 en rouge, T_2 en bleu, des éléments de la tribu $\mathscr{T}_0$

Sur ce même Ω, on peut définir de façon un peu vague, une tribu $\mathscr{T}$ beaucoup plus large, à savoir, l'ensemble des parties du carré admettant une aire [25] et définir sur cette tribu $\mathscr{T}$ la probabilité $\mathbb{P}$ définie par

$$\forall T \in \mathscr{T}, \mathbb{P}(T) = \frac{\text{aire}(T)}{\text{aire}(\Omega)} = \text{aire}(T)$$

Spécifier la probabilité d'un événement A, c'est imposer une mesure de la proportion de A relativement à Ω.

25. C'est beaucoup plus délicat que ça n'en a l'air !

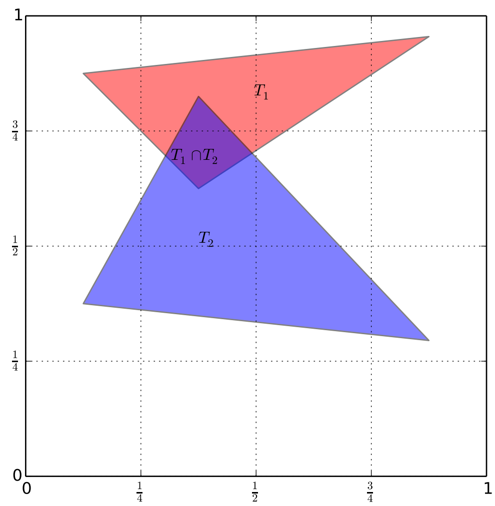


FIGURE 26 – Le carré Ω, quelques éléments de la tribu $\mathscr{T}$

Exemples graphiques : Segment (HP)

Considérons maintenant le segment $\Omega = [0, 1[$ subdivisé en petits segments $S_k = \left[\frac{k}{16}, \frac{k+1}{16}\right[, k \in \{0, \ldots, 15\}$, de longueur $\frac{1}{16}$ comme sur la figure 27. On a épaissi les segments en hauteur pour pouvoir voir les couleurs.

L'ensemble $\mathscr{S} = \{S_k, k \in \{0, \ldots, 15\}\}$ est un ensemble de parties de Ω. Il comporte 16 éléments. Ce **n'est pas** une tribu de parties de Ω.

Par contre, l'ensemble $\mathscr{T}_0$ où chaque élément est une figure T formée d'une union finie de segments S_k est une tribu de parties de Ω. Il comporte 2^{16} éléments. Une probabilité $\mathbb{P}_0$ sur cette tribu est donnée par l'application définie par

$$\forall T \in \mathscr{T}_0, \mathbb{P}_0(T) = \frac{\text{longueur}(T)}{\text{longueur}(\Omega)} = \text{longueur}(T)$$

Sur la figure 27, on a

$$\mathbb{P}_0(T_1) = \frac{4}{16}, \mathbb{P}_0(T_2) = \frac{3}{16}, \mathbb{P}_0(T_1 \cap T_2) = \frac{1}{16} \text{ et } \mathbb{P}_0(\overline{(T_1 \cup T_2)}) = \frac{10}{16}$$

Comme pour le carré, on peut aussi considérer la tribu $\mathscr{T}$ de toutes les parties T du segment unité admettant une longueur. La longueur longueur(T) d'une telle partie peut-être interprétée comme la probabilité $\mathbb{P}(T)$ de l'événement T.

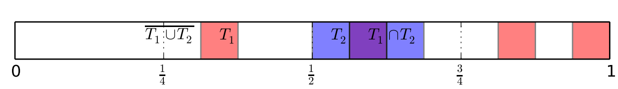


FIGURE 27 – Le segment, T_1 en rouge, T_2 en bleu, des éléments de la tribu $\mathscr{T}_0$

Pouvez vous deviner la transformation qui fait passer du carré au segment et réciproquement ?

Un exemple graphique (HP)

Reprenons le cas où Ω est le carré $[0,1[\times[0,1[$, $\mathbb{P}$ l'aire et prenons comme v.a., à valeurs dans $[0,1[$, X_1 et X_2 définies par

$$\forall\omega=(\omega_1,\omega_2)\in[0,1[\times[0,1[,\, X_1(\omega)=\omega_1,\, X_2(\omega)=\omega_2$$

Sur la figure 28, on a représenté

$$T_1=\left\{X_2\in\left[\frac{1}{2},\frac{3}{4}\right[\right\}\text{ et }T_2=\left\{X_1\in\left[\frac{1}{4},\frac{1}{2}\right[\right\}$$

On a

$$\mathbb{P}(\frac{1}{2}\leq X_1<\frac{1}{4})=\frac{1}{4},\mathbb{P}(\frac{1}{4}\leq X_2<\frac{1}{2})=\frac{1}{4},$$
$$\mathbb{P}(\frac{1}{2}\leq X_1<\frac{1}{4}\text{ et }\frac{1}{4}\leq X_2<\frac{1}{2})=\frac{1}{16}$$

Une interprétation graphique plus imaginative montre que, si I_1, I_2 sont deux intervalles quelconques de $\mathbb{R}$, alors

— $\{X_1\in I_1\}$ est un rectangle « vertical » de largeur $\mathbb{P}(X_1\in I_1)$, de hauteur 1
— $\{X_2\in I_2\}$ est un rectangle « horizontal » de largeur 1, de hauteur $\mathbb{P}(X_2\in I_2)$,
— $\{X_1\in I_1\}\cap\{X_2\in I_2\}$ est un rectangle de largeur $\mathbb{P}(X_1\in I_1)$, de hauteur $\mathbb{P}(X_2\in I_2)$,
— On a donc $\mathbb{P}(X_1\in I_1\text{ et }X_2\in I_2)=\mathbb{P}(X_1\in I_1).\mathbb{P}(X_2\in I_2)$.
— Ceci étant vrai pour tout couple d'intervalle (I_1,I_2), les v.a X_1 et X_2 sont indépendantes ?

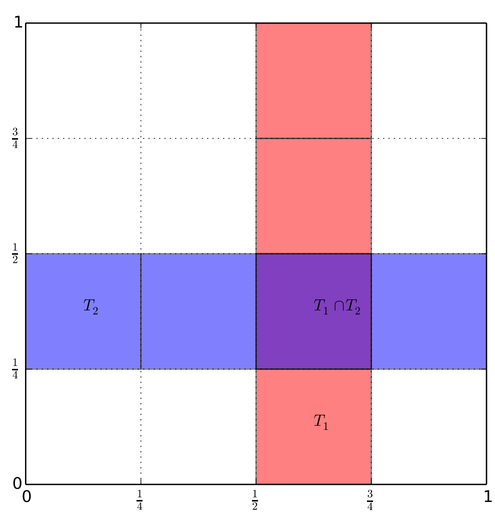


FIGURE 28 – Le carré, T_1 en rouge, T_2 en bleu, $T_1\cap T_2$

Posons maintenant les v.a. D_1 et D_2 définies par

$$D_1=\begin{cases}0 & \text{si } 0\leq X_1<\frac{1}{4}\\ 1 & \text{si } \frac{1}{4}\leq X_1<\frac{1}{2}\\ 2 & \text{si } \frac{1}{2}\leq X_1<\frac{3}{4}\\ 3 & \text{si } \frac{3}{4}\leq X_1<1\end{cases}\text{ et }D_2=\begin{cases}0 & \text{si } 0\leq X_2<\frac{1}{4}\\ 1 & \text{si } \frac{1}{4}\leq X_2<\frac{1}{2}\\ 2 & \text{si } \frac{1}{2}\leq X_2<\frac{3}{4}\\ 3 & \text{si } \frac{3}{4}\leq X_2<1\end{cases}$$

Pour résumer, pour $(i,j) \in \{0,\ldots,3\} \times \{0,\ldots,3\}$, $\omega \in [0,1[^2$, $(D_1(\omega), D_2(\omega)) = (i,j)$ si et seulement si $\omega \in C_{i,j}$. D_1 et D_2 sont des variables prenant un nombre fini de valeurs, uniformément distribuées sur $\{0,\ldots,3\}$ et indépendantes.

En d'autres termes, on a « construit » un espace [26] pour la modélisation de l'expérience de tirer indépendamment deux dés à 4 faces numérotées de 0 à 3 [27]. Cet espace est différent de l'espace naturel, fini, où l'on prend $\Omega' = \{0,\ldots,3\}^2$ et $\mathbb{P}'$ la probabilité uniforme sur Ω'.

Les calculs relativement à un espace ou à l'autre donnent les mêmes résultats concernant les expériences à deux dés, le nouvel espace permet en plus de se garder la possibilité de construire d'autres v.a. n'ayant rien à voir avec l'expérience des deux dés.

26. En fait deux : $(\Omega, \mathscr{T}_0, \mathbb{P}_0)$ et $(\Omega, \mathscr{T}, \mathbb{P})$, ce dernier n'ayant été que très incomplètement défini.
27. Les amateurs de jeux de rôles apprécieront