Lycée Pierre-Gilles de Gennes **BCPST2**
Mathématiques **2019-2020**

Révisions/Formulaire 02

Révisions Probabilités

Table des matières

A retenir de ce chapitre :
- Les principes de modélisation probabiliste.
- Les notions de variables aléatoires, d'événements, de probabilité d'un événement et les règles de calcul associées. Variables indicatrices (ou de BERNOULLI). Indépendance et conditionnement.
- Représentations graphiques en histogramme, simulations en Python
- La notion d'espérance d'une v.a., de fonction de répartition.
- La formule de transfert.
- Les lois discrètes et leurs caractéristiques du programme : BERNOULLI, uniformes, binomiale et hypergéométrique.

1 Qu'est ce que modéliser

1.1 Le point de départ

La physique, la chimie, la biologie, etc...sont des sciences qui cherchent à modéliser mathématiquement un certain nombre de phénomènes.

Le point de départ est *grosso modo* toujours le même

1. On cherche à identifier les caractéristiques « visibles »du phénomène. Ces caractéristiques peuvent être quantitatives ou qualitatives.
2. Certaines de ces caractéristiques peuvent être *mesurables* ou *observables*. En effectuant une expérience, on pourra évaluer ces caractéristiques sur l'expérience donnée.
3. D'autres caractéristiques peuvent être des paramètres de l'expérience. On peut, au fil des expériences, varier un ou plusieurs de ces paramètres et, à chaque fois, évaluer les caractéristiques observables.
4. Convenons d'appeler ces caractéristiques les *variables* de la modélisation du phénomène observé.

Voici quelques questions qu'un humain aurait pu se poser il y a 10000 ans

- En calorimétrie : je dispose de deux bols d'eau, l'une de la neige tout juste fondue, l'autre tout juste portée à ébullition. Je les mélange. Puis-je boire le tout sans me brûler ?
- En balistique : je lance une sagaie vers le mammouth que j'aimerais manger. Parfois je loupe (trop court ou trop long), parfois je réussis. J'aimerais réussir à coup sûr.
- En halieutique : dans ce lac, quand je lance mon filet, j'en retire deux sortes de poissons. Certaines années, une espèce herbivore semble très majoritaire, d'autres années, c'est l'autre espèce carnivore qui est majoritaire. Pourquoi ?

Pour aborder scientifiquement ces questions, on a besoin de deux concepts : *nombre* et *quantité*.

Le concept de nombre est très ancien et notre ancêtre mangeur de mammouth pouvait très probablement se figurer le troisième problème en termes très similaires (compter le nombre d'animaux).

Concernant les quantités, il est clair que dans le premier problème on peut distinguer que dans un bol il y a plus d'eau que dans le second ou que l'eau froide est plus froide que l'eau chaude. Il n'y a pas de nombres en jeu. Les grecs (et peut-être avant eux les égyptiens ou d'autres dont on a perdu la trace) ont établi un pont entre ces deux concepts en inventant les unités de mesure.

On peut dire que le stade d'Olympie est long comme 600 pieds d'HERCULE ou que le rapport entre l'aire d'un cercle et l'aire du carré ayant pour côté le rayon de cercle, c'est un nombre, est la moitié du rapport entre le périmètre de ce cercle et le périmètre du carré.

Bref, en se servant d'une quantité étalon, on peut évaluer, par des nombres, les quantités de même nature.

Cela implique que les résultats numériques issus de la modélisation d'une situation physique ne sont interprétables qu'avec leur unité de mesure associée. Un changement d'unité se traduit par un changement des résultats numériques.

1.2 Des exemples

Calorimétrie

Concernant le premier problème, il a fallu quelques centaines d'années pour que les concepts de température et de quantité de chaleur reçue ou donnée émergent.

- On a vu comment les longueurs (les aires, les volumes) peuvent être évalués en les comparant à une longueur (aire, volume) étalon.

- Il en est de même pour « quantité de matière », ça se mesure de façon élémentaire, par exemple, en comparant à la balance avec des poids étalonnés. En se donnant une échelle de comparaison, on identifie la masse mesurée à un nombre m.
- la température, ça se mesure, à l'origine, par un système de comparaison avec une matière pouvant se dilater. C'est le principe du thermomètre : une variation de température se mesure par une variation de longueur de la matière du thermomètre. De la sorte, on identifie la température mesurée à un nombre T.

Maintenant que l'on sait mesurer, on peut faire des expériences : dans notre cas, on peut faire varier les masses d'eau froide et d'eau chaude dans notre mélange, varier leurs températures et mesurer la température finale. On a donc 5 variables

$$m_0, m_1, T_0, T_1, T_f$$

respectivement la masse d'eau froide (m_0), celle d'eau chaude (m_1), leurs températures initiales (T_0 et T_1) et la température finale du mélange (T).

On peut constater, expérimentalement, que *grosso modo*, T semble être la moyenne pondérée de T_0 et T_1 avec comme coefficients m_0 et m_1.

$$T \sim \frac{m_0}{m_0+m_1}.T_0 + \frac{m_1}{m_0+m_1}.T_1$$

Une telle relation, affine en T_0 et T_1, est assez facile à constater sur des tables de mesures.

Cette constatation n'est pas une explication ! Une explication doit ressortir d'un niveau plus élémentaire et décrire le mécanisme par lequel on aboutit à cet équilibre.

Si on introduit le concept de quantité de chaleur reçue ou donnée par un corps au cours d'une transformation, la relation obtenue s'explique alors par l'établissement d'un bilan : ce qui a été donné par l'un a été reçu par l'autre et au final, on a un équilibre.

Un tel principe d'échange est suffisamment simple pour être considéré comme explicatif. Mais on pourrait poser la question plus avant : c'est quoi exactement la quantité de chaleur ? comment se produit l'échange, etc...

Là, c'est de la physique !

Balistique

Pour le problème de la sagaie, les variables sont certainement la direction initiale du lancer, la puissance avec laquelle celui-ci est effectué et la distance parcourue par la sagaie.

Je laisse les spécialistes de mécanique établir que, en terrain plat, si la sagaie, de masse m, est lancée avec un angle initial (en radians) de α_0, une énergie cinétique initiale E_0 alors la distance parcourue [1] est

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

On peut douter [2] qu'une telle relation soit directement visible sur une expérience en faisant varier les paramètres E_0 et α_0.

Cette 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 par contre, vérifier 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.

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. personnellement, je n'imagine pas voir/deviner autre chose qu'une relation affine

Halieutique

Dans ce problème, on n'a pas besoin de concepts physiques extrêmement élaborés. On peut se contenter, à chaque pêche, de consigner les proportions de chaque espèce, disons A et B, qui doivent donc être des variables du modèle.

Une variable importante fait son apparition, c'est le temps. Supposons que nous le mesurions en jours, que nous fassions une pêche par jour, et que nous notions t le jour.

Il y a probablement une relation entre A et B au jour $t+1$ et A et B au jour t. Ceci se traduit par le fait que A et B sont fonctions de t et qu'ils satisfont une relation de récurrence du type

$$\left\{\begin{array}{lcl} A_{t+1} & = & \text{fonction}(A_t, B_t) \\ B_{t+1} & = & \text{fonction}(A_t, B_t) \end{array}\right.$$

On peut bien tracer, à partir de données réellles, des graphiques donnant A et B en fonctions de t. Il est douteux que l'on puisse se servir de ces graphiques pour construire une « formule » donnant A et B en fonction de t ;

Proportions de prédateurs (requins, raies, etc.) et de proies débarquées au port de Fiume

Année	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923
Prédateurs	11,9	21,4	22,1	21,2	36,4	27,3	16	15,9	14,8	10,7
Proies	88,1	78,6	77,9	78,8	63,6	72,7	84	84,1	85,2	89,3

FIGURE 1 – Les données de VOLTERRA.

Donner, en Python, une représentation graphique des données[3] de VOLTERRA présentées en fig. 1.

En résumé

Modéliser, c'est déterminer un ensemble de variables intéressantes du phénomène étudié, c'est ensuite trouver quelles relations entre variables sont adéquates pour décrire le phénomène, sur une base rationnelle (quels échanges entre les deux populations, quel taux de reproduction, de quoi dépend-il, etc...) ou sur une base empirique.

C'est aussi être conscient des hypothèses simplificatrices qui ont été faites.

Finalement, il y a toujours une étape de confrontation *critique* des résultats du modèle avec le phénomène réel.

1.3 Et les probabilités ?

Le hasard (et donc le calcul des probabilités) peut intervenir de différentes manières dans un processus de modélisation/confrontation.

Cela signifie que l'on peut considérer certaines variables du modèle comme des *variables aléatoires*, *i.e.* des variables issues du *hasard*.

1. On peut croire que le hasard existe réellement dans la nature et donc, certaines variables du modèle seront par essence aléatoires.
2. Même si l'on n'y croit pas, on peut considérer comme aléatoires certaines variables que l'on ne peut contrôler ou mesurer précisément.

3. *Source :* `http://accromath.uqam.ca/2013/04/des-predateurs-et-leurs-proies/`

Les probabilités interviennent aussi dans le traitement statistique des résultats d'expérience et donc la confrontation du modèle à la réalité.

Reprenons nos exemples

— Calorimétrie. Dans l'expérience proposée, on a la relation

$$T = \frac{m_0}{m_0 + m_1}.T_0 + \frac{m_1}{m_0 + m_1}.T_1$$

On peut se demander, par exemple, ce qui se passe pour T, *i.e.* comment T varie lorsque l'on a une indétermination sur m_0. Si l'on sait que lors de l'expérience m_0 est distribuée d'une certaine façon autour d'une valeur centrale $\overline{m}_0$, on peut se demander si, et comment, T est distribuée autour d'une certaine valeur centrale $\overline{T}$.

— Balistique. Imaginons notre lanceur de sagaie atteint de la maladie de PARKINSON. On peut imaginer que son angle de jet ainsi que l'énergie initiale varient tous les deux et sachant cela, on peut se demander les chances qu'il a d'atteindre(succès) ou pas(échec) une cible ayant une certaine largeur. Cela peut se mesurer expérimentalement en itérant suffisamment l'expérience pour évaluer la proportion de succès. [4]

— Enfin, pour un modèle d'évolution de population, le hasard intervient assez naturellement lorsque l'on considère le phénomène au niveau individuel. Chaque animal va avoir une certaine descendance, une certaine façon de rencontrer ou pas son prédateur ou sa proie suivant un scénario qui lui est propre. Le modèle déterministe ne peut-être que l'expression de la tendance centrale d'un modèle microscopique probabiliste sous-jacent.

2 Modèles probabilistes- la pratique, révisions

2.1 Ce que comporte un modèle

Un modèle probabiliste pour un système physique ou biologique, c'est, d'abord

1. Une famille de variables aléatoires, souvent numériques, les variables du modèle, $X_1,...X_n,...,Y_1,...,Z,...$ qui décrivent l'*état* du système
2. La donnée des relations qu'elles entretiennent entre elles : y a-t-il des relations issues de la physique, de la biologie ? sont elles indépendantes ?...
3. des hypothèses portant sur les *distributions* de certaines de ces variables, p.ex les $X_1,...X_n,..$

Par exemple, dans le problème du lanceur de sagaie tremblotant, les variables *primitives* du système sont α_0, l'angle initial et E_0, l'énergie initiale. m, la masse de la sagaie est une constante, tout comme g. La variable intéressante est d la distance à laquelle atterrit la sagaie.

Les variables α_0 et E_0 sont aléatoires (on verra des exemples), on peut supposer–ou pas–qu'elles sont indépendantes. La variable d est elle aussi aléatoire, elle dépend, via la formule de α_0 et E_0.

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.

4. On pourrait aussi se demander l'effet du vent sur le résultat final. Vent que l'on ne maîtrise pas et que l'on pourrait modéliser de façon probabiliste, c'est compliqué !

2.2 Eléments de première année revisités

Événements, probabilité, distributions

Un événement, c'est le fait qu'une (ou plusieurs) variable soit localisée, ou pas, dans une zone (numérique) donnée, c'est aussi une combinaison logique de tels faits. Typiquement si X, Y sont des variables du modèle, à valeurs réelles, I, J sont des intervalles de $\mathbb{R}$, le fait que X appartienne à I est un événement, le fait que simultanément $X \in I$ et $Y \in J$ l'est aussi.

Un événement est donc donné par une proposition logique faisant intervenir la localisation des valeurs prises par les variables du modèle.

On calcule des probabilités d'événements. $\mathbb{P}(X \in I)$ désignera la probabilité que la valeur prise par la variable X appartienne à l'intervalle I. C'est un nombre entre 0 et 1 décrivant la proportion d'états du système, de cas de figure, qui mènent au résultat $X \in I$.

1. Un/l' événement certain c'est quand dans tous les cas de figure, la proposition logique décrivant l'événement est vérifiée. Sa probabilité est $1 = 100\%$.
2. Un événement *quasi-certain*[5] ou *presque sûr* c'est un évenement de probabilité est $1 = 100\%$.
3. Un/l' événement impossible c'est quand dans tous les cas de figure, la proposition logique décrivant l'événement est fausse. Sa probabilité est $0 = 0\%$.
4. Un événement *quasi-impossible* ou *négligeable* c'est un évenement de probabilité est $0 = 0\%$.

Les probabilités d'événements doivent satisfaire à des règles logiques « naturelles ».

1. (Une conséquence est plus probable que sa cause) Si un événement A implique un événement B, la probabilité de A doit être inférieure à celle de B car la proportion de cas de figure où A est réalisé est inférieure à la proportion de cas de figures où B est réalisé.
2. (Axiome d'additivité) Si deux événements A et B sont incompatibles au sens où aucune configuration du système vérifiant A vérifie B et vice-versa, la probabilité de leur alternative (A ou B est vrai) est la somme de leurs probabilités.
3. (Définition de l'indépendance) Enfin, deux événements A et B sont indépendants, si la probabilité de leur conjonction (A et B est vrai) est le produit de leurs probabilités.

Connaître la famille de nombres $\mathbb{P}(X \in I)$ où I décrit l'ensemble des intervalles de $\mathbb{R}$ c'est connaître la distribution de X ou la loi de X.

Une telle distribution ou, plus modestement, des informations sur cette distribution, peuvent être une donnée de la modélisation ou des résultats recherchés par le traitement mathématique du modèle.

Pour une partie A de $\mathbb{R}$, on dit que X, v.a. réelle, est ($\mathbb{P}$-presque sûrement) à valeurs dans A si $\mathbb{P}(X \in A) = 1$. Cette définition s'étend naturellement si X est une v.a. à valeurs non réelles.

Trouver un ensemble, simple mais raisonnablement resserré, de valeurs prises par une variable d'intérêt est le premier renseignement à rechercher sur cette variable.

On trouve dans les livres la notation $X(\Omega)$ pour désigner *l'*ensemble des valeurs prises par X. Pour une raison de mauvaise définition de cet ensemble, j'évite cette notation.

V.a. de BERNOULLI

Définition 1. *Une v.a X est dite de* BERNOULLI*, de paramètre de succès $p \in [0,1]$ si*

$$\mathbb{P}(X = 0) = 1 - p \text{ et } \mathbb{P}(X = 1) = p.$$

On note $X \sim \mathscr{B}(p)$.

5. Terminologie BCPST

1. Une v.a de BERNOULLI est à valeurs dans $\{0,1\}$.
2. Si A est un événement, $1\!\!1_A$, la v.a indicatrice [6] de A est une variable de BERNOULLI de paramètre $p = \mathbb{P}(A)$.
3. Si une v.a. X est à valeurs dans $\{0,1\}$, c'est une v.a de BERNOULLI de paramètre $p = \mathbb{P}(X = 1)$. C'est la variable indicatrice de l'événement $\{X = 1\}$.
4. Les variables aléatoires de BERNOULLI sont les variables les plus fondamentales. Toute v.a. réelle prenant un nombre fini de valeurs est combinaison linéaire de v.a. de BERNOULLI.

Espérance

L'espérance d'une variable numérique X doit être comprise comme la moyenne des valeurs prises par X lorsque que l'on moyenne sur tous les états possibles du système.

Le mot espérance provient de la théorie des jeux : si dans un jeu (par exemple pile ou face), le succès A rapporte une somme donnée alors que l'échec non(A) ne rapporte rien, l'*espérance de gain* pour ce jeu est

Montant à gagner$\times$ probabilité de gagner

En formules, si A est un événement, on note $1\!\!1_A$ la variable aléatoire valant 1 si l'événement A survient, 0 sinon. Elle représente le gain du joueur si le montant à gagner est de 1. On a

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

Si le montant à gagner en cas de succès est s, le gain du joueur est donné par la v.a.r $s.1\!\!1_A$, son espérance est

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

Si un jeu a trois issues possibles, marquées par la famille complète d'événements incompatibles A_1, A_2, A_3, avec une somme s_k à gagner en cas de survenue de l'événement A_k alors l'espérance de gain est

Montant $s_1\times$ probabilité de A_1 + Montant $s_2\times$ probabilité de A_2+..

En formules, si S est le gain de ce joueur, on a

$$S = s_1.1\!\!1_{A_1} + s_2.1\!\!1_{A_2} + s_3.1\!\!1_{A_3}$$

et

$$\begin{aligned}\mathbb{E}(S) &= s_1.\mathbb{P}(A_1) + s_2.\mathbb{P}(A_2) + s_3.\mathbb{P}(A_3)\\ &= s_1.\mathbb{E}(1\!\!1_{A_1}) + s_2.\mathbb{E}(1\!\!1_{A_2}) + s_3.\mathbb{E}(1\!\!1_{A_3})\end{aligned}$$

Retours sur les calculs de première année

La théorie que vous avez vue en première année est bâtie sur la considération de

1. Un ensemble Ω fini, un événement A étant une partie quelconque de Ω
2. Une suite finie de poids [7] $(p_\omega)_{\omega\in\Omega}$ vérifiant

$$\forall\omega \in \Omega, 0 \leq p_\omega \leq 1 \text{ et } \sum_{\omega\in\Omega} p_\omega = 1$$

permettant de définir la probabilité d'un événement A par la formule

$$\mathbb{P}(A) := \sum_{\omega\in A} p_\omega = \sum_{\omega\in\Omega} 1\!\!1_{\{\omega\in A\}} p_\omega$$

6. la v.a. valant 1 si A survient 0 sinon
7. un « *germe* » de probabilité sur Ω

3. et de considérer qu'une variable aléatoire X à valeurs dans [8] $\mathscr{X}$ est une application $X : \Omega \to \mathscr{X}$
4. Si A est un événement, $1\!\!1_A$ la v.a de BERNOULLI indicatrice de A, *i.e.* la v.a. valant 1 si A survient 0 est définie par
$$\forall \omega \in \Omega,\ 1\!\!1_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases} = 1\!\!1_{\{\omega \in A\}}$$
On a donc
$$\mathbb{E}(1\!\!1_A) = \mathbb{P}(A) = \sum_{\omega \in \Omega} 1\!\!1_A(\omega).p_\omega$$

Ce cadre rattache la théorie des probabilités aux mathématiques « standard ». Du fait du caractère fini de Ω, il ne permet de ne traiter que le cas où un nombre fini de variables prenant un nombre fini de valeurs interviennent. On va chercher à s'affranchir de cette contrainte de finitude en évitant de se référer à la nature intime de Ω.

Dans ce cadre, et dans le cadre généralisé, si X, Y sont deux v.a.r, I, J deux intervalles, l'événement $\{X \in I\}$ est une partie de Ω, c'est la partie

$$\{X \in I\} := \{\omega \in \Omega, X(\omega) \in I\} = X^{-1}(I)$$

Sa négation est l'événement $\{X \notin I\}$, i.e

$$\begin{aligned} \{X \notin I\} &:= \{\omega \in \Omega, X(\omega) \notin I\} \\ &= \{\omega \in \Omega, X(\omega) \in \overline{I}\} \end{aligned}$$

L'événement $\{X \in I \text{ et } Y \in J\}$, c'est la partie

$$\begin{aligned} \{X \in I \text{ et } Y \in J\} &:= \{\omega \in \Omega, X(\omega) \in I \text{ et } Y(\omega) \in J\} \\ &= \{\omega \in \Omega, (X(\omega), Y(\omega)) \in I \times J\} \end{aligned}$$

Les opérations logiques d'alternative (ou), de conjonction (et) et de négation $\overline{(.)}$ recherchées pour les évenements probabilistes se traduisent respectivement d'un point de vue ensembliste par les opérations de réunion, d'intersection ou de complémentaire. A titre d'exemples : si X, Y sont deux v.a.r, I, J deux intervalles,

$$\begin{aligned} \{X \in I \text{ ou } Y \in J\} &= \{X \in I\} \cup \{Y \in J\} \\ \{X \in I \text{ et } Y \in J\} &= \{X \in I\} \cap \{Y \in J\} \\ \{X \notin I\} &= \overline{\{X \in I\}} \end{aligned}$$

On a le résultat suivant permettant de fonder mathématiquement les calculs sur un modèle probabiliste :

Théorème 2 (Bonne fondation des modèles probabilistes finis)**.** *Soit $x_1, \ldots, x_K$, K éléments distincts d'un ensemble $\mathscr{X}$, $p_1, \ldots, p_K \in [0,1]$ tels que $\sum_{k=1}^K p_k = 1$ alors il existe un ensemble (fini) Ω, une probabilité $\mathbb{P}$ sur $\mathscr{P}(\Omega)$ et une application $X : \Omega \to \mathscr{X}$ tels que*

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

8. Si $\mathscr{X} = \mathbb{R}$, on parle de variable aléatoire réelle, en abrégé v.a.r.

Définition 3 (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 $\{x_k, k \in \{1,\dots,K\}\}$.
2. Du théorème précédent, on déduit qu'étant 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

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

3. Ici, Ω, $\mathbb{P}$ dépendent de l'ensemble fini donné *a priori* . On verra dans le chapitre suivant des méthodes de simulation (informatique) de variables finies et que l'on peut étendre la théorie à des ensembles Ω de sorte à obtenir un espace Ω (muni d'une probabilité $\mathbb{P}$) universel.

Lorsque l'on modélise un expérience probabiliste, il est primordial de spécifier ce qui est tiré au sort de façon primitive et de bien poser les lois (très souvent une loi uniforme).

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

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 $\frac{\pi}{8}$, $\frac{\pi}{4}$ 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 ?

Indication:On rappelle qu'avec une énergie de lancer E, un angle de tir α, le point d'impact est à distance $d = \frac{2.E.\sin(2\alpha)}{m.g}$

Lien entre probabilités et combinatoire

La combinatoire permet de calculer le nombre d'objets mathématiques satisfaisant certaines propriétés.

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 $\{1,\dots,n\} \to \mathscr{X}$, *i.e.* de numérotations des éléments de $\mathscr{X}$.
- Si $p > 0$, une p-liste[9] 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 $\{1,\dots,p\} \to \mathscr{X}$.
- 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* $\{1,\dots,p\} \to \mathscr{X}$.

9. 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.

— 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.

Exercice 2.— 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 3.— 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 :

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

Évacuer Ω des calculs

La nature exacte de Ω n'intervient pas dans un calcul probabiliste. Seules les distributions (et les quantités associées) des v.a. interviennent et nous intéressent.

Supposons que X soit une variable aléatoire ne pouvant prendre que les valeurs réelles x_1, x_2 et x_3 où $x_1 < x_2 < x_3$. On suppose que l'on connaît les trois probabilités suivantes :

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

les événements $\{X = x_1\}$, $\{X = x_2\}$ et $\{X = x_3\}$ sont mutuellement incompatibles : on rejette le fait que dans un cas de figure donné, la valeur prise par X soit à la fois x_1 et x_2. Il y a plusieurs (une infinité si l'on y réfléchit) modèles probabilistes finis possibles

- Le plus simple : $\Omega = \{x_1,x_2,x_3\} \subset \mathbb{R}$, $\mathbb{P}$ définie sur $\mathscr{P}(\Omega)$ par le système de poids (x_1, p_1), (x_2, p_2) et (x_3, p_3) et enfin $X : \Omega \to \mathbb{R}$ définie par $\forall\omega \in \Omega, X(\omega) = \omega$.
- Variante à numerotation : $\Omega = \{1,2,3\}$, $\mathbb{P}$ définie sur $\mathscr{P}(\Omega)$ par le système de poids $(1, p_1)$, $(2, p_2)$, $(3, p_3)$ et enfin $X : \Omega \to \mathbb{R}$ définie par $\forall\omega \in \Omega, X(\omega) = x_\omega$.
- Variante inutilement compliquée : $\Omega = \{1,2,3\} \times \{0,1\}$, $\mathbb{P}$ définie sur $\mathscr{P}(\Omega)$ par le système de poids $((1,0), p_1/3)$, $((2,0), p_2/3)$, $((3,0), p_3/3)$, $((1,1), 2p_1/3)$, $((2,1), 2p_2/3)$ et $((3,1), 2p_3/3)$ et enfin $X : \Omega \to \mathbb{R}$ définie par $\forall\omega = (\omega_1, \omega_2) \in \Omega, X(\omega) = x_{\omega_1}$.

On peut, dans chaque cas, dessiner les événements $\{X = x_1\}$, $\{X = x_2\}$ et $\{X = x_3\}$. Les calculs suivants ne font pas intervenir la nature de Ω, ils sont valables dans tous les modèles présentés précédemment.

L'alternative de ces trois événements, 2 à 2 incompatibles est certaine : X doit prendre une valeur et, par hypothèse, c'est l'une des trois valeurs x_1, x_2 ou x_3. En conclusion, on doit avoir, en appliquant les règles régissant la probabilité $\mathbb{P}$,

$$\begin{aligned} p_1 + p_2 + p_3 &= \mathbb{P}(X = x_1) + \mathbb{P}(X = x_2) + \mathbb{P}(X = x_3) \\ &= \mathbb{P}((X = x_1) \text{ ou } (X = x_2) \text{ ou } (X = x_3)) = 1 \end{aligned}$$

Si I est un intervalle quelconque de $\mathbb{R}$, la probabilité $\mathbb{P}(X \in I)$ se calcule comme suit car l'événement $\{X \in I\}$ est l'alternative des trois événements incompatibles $\{X = x_1 \text{ et } X \in I\}$, $\{X = x_2 \text{ et } X \in I\}$ et $\{X = x_3 \text{ et } X \in I\}$. (Décomposition suivant les valeurs prises par X)

$$\mathbb{P}(X \in I) = \mathbb{P}(X = x_1 \text{ et } X \in I) + \mathbb{P}(X = x_2 \text{ et } X \in I) + \mathbb{P}(X = x_3 \text{ et } X \in I)$$

Si, par exemple $x_1, x_2 \in I$ et $x_3 \notin I$, on a, par redondance,

$$\{X = x_1 \text{ et } X \in I\} = \{X = x_1\}\,, \{X = x_2 \text{ et } X \in I\} = \{X = x_2\}\,,$$

et $\{X = x_3 \text{ et } X \in I\}$ est impossible et donc

$$\begin{aligned} \mathbb{P}(X \in I) &= \underbrace{\mathbb{P}(X = x_1 \text{ et } X \in I)}_{p_1} + \underbrace{\mathbb{P}(X = x_2 \text{ et } X \in I)}_{p_2} + \underbrace{\mathbb{P}(X = x_3 \text{ et } X \in I)}_{0} \\ &= p_1 + p_2 \end{aligned}$$

On a en général, avec les mêmes principes de calcul :

$$\mathbb{P}(X \in I) = \sum_{i=1}^{3} 1\!\!1_{\{x_i \in I\}} \mathbb{P}(X = x_i) = \mathbb{E}(1\!\!1_{\{X \in I\}})$$

On a, au niveau des espérances,

$$X = x_1 . 1\!\!1_{\{X=x_1\}} + x_2 . 1\!\!1_{\{X=x_2\}} + x_3 . 1\!\!1_{\{X=x_3\}} = \sum_{k=1}^{3} x_k . 1\!\!1_{\{X=x_k\}}$$

et

$$\mathbb{E}(X) = \sum_{k=1}^{3} x_k . \mathbb{P}(X = x_k) = \sum_{k=1}^{3} x_k . \mathbb{E}(1\!\!1_{\{X=x_k\}}) = \sum_{k=1}^{3} x_k . p_k$$

et

$$X^2 = x_1^2 . 1\!\!1_{\{X=x_1\}} + x_2^2 . 1\!\!1_{\{X=x_2\}} + x_3^2 . 1\!\!1_{\{X=x_3\}} = \sum_{k=1}^{3} x_k^2 . 1\!\!1_{\{X=x_k\}}$$

et

$$\mathbb{E}(X^2) = \sum_{k=1}^{3} x_k^2 . \mathbb{E}(1\!\!1_{\{X=x_k\}}) = \sum_{k=1}^{3} x_k^2 . p_k$$

On voit ressortir de ces calculs, la formule de transfert générique pour la v.a. X, si $h : \{x_1, x_2, x_3\} \to \mathbb{R}$ est une fonction, on a

$$h(X) = \sum_{k=1}^{3} h(x_k) . 1\!\!1_{\{X=x_k\}}$$

et

$$\mathbb{E}(h(X)) = \sum_{k=1}^{3} h(x_k) . \mathbb{P}(X = x_k) = \sum_{k=1}^{3} h(x_k) . p_k$$

2.3 Simulation informatique, interprétation fréquentielle

En Python, sous `numpy`, on dispose [10] 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` tiré uniformément (avec probabilité donnée par `p` si `p` est une liste de même taille que `liste`).
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 réel tiré uniformément sur $[0, 1[$.

Nous verrons plus tard [11] 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.

10. après importation par `import numpy as np`
11. c'est un thème récurrent du cours

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 [12] 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.

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

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

12. De combien est *la* question fondamentale traitée dans le chapitre de statistiques

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=[...],normed=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()
```

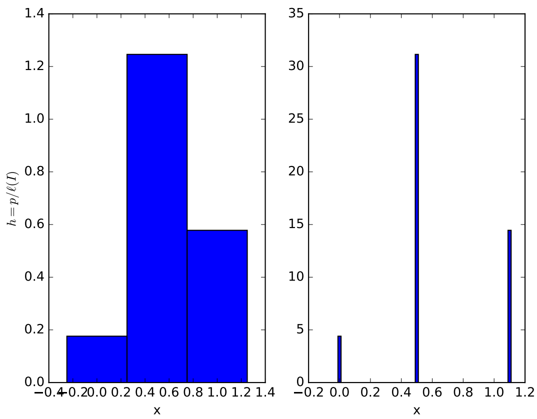


FIGURE 2 – Deux histogrammes d'une simulation de N=1000 valeurs de X, bins de largeur 0.5 autour des valeurs possibles, bins de largeur 0.02. La somme des aires des rectangles vaut 1.

Un apparté sur les bins et les histogrammes

Comme le montre la figure 3, 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 4) ont mieux compris ce problème.

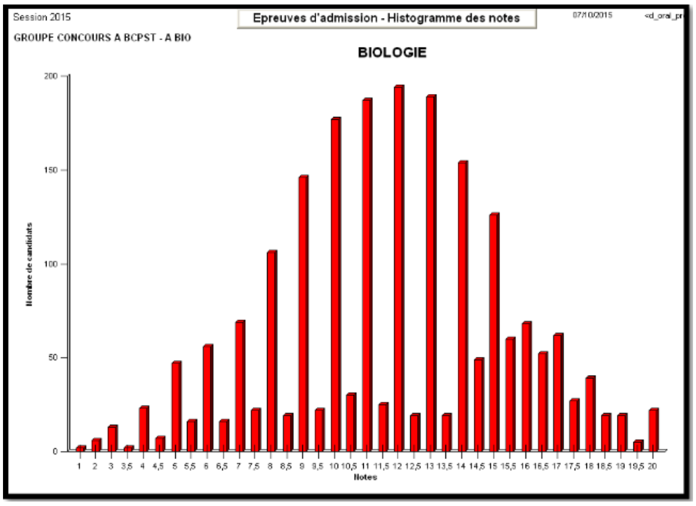


FIGURE 3 – Un histogramme des résultats de l'épreuve de Synthèse-Biologie AV 2015.

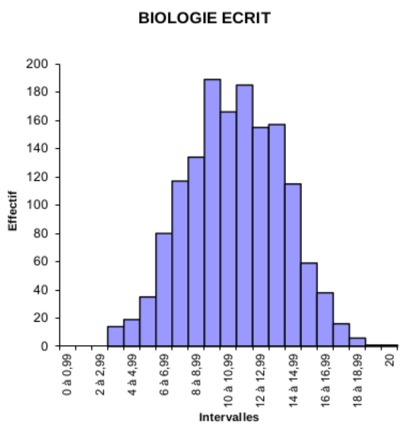


FIGURE 4 – Un histogramme des résultats de l'épreuve de Biologie G2E 2016.

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. *c.f.* figures 5 et 6

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)

Exercice 4.— 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 $Z = Y^2 - 2$.

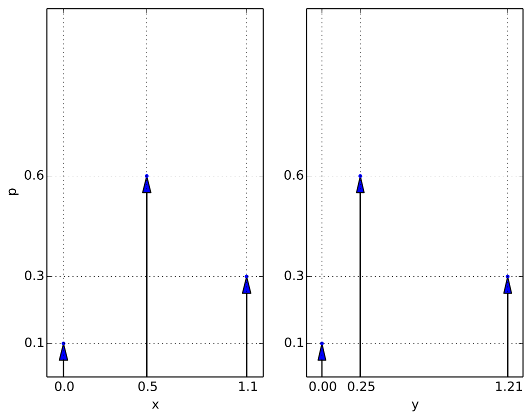


FIGURE 5 – La distribution d'une variable aléatoire X et la distribution de $Y = X^2$.

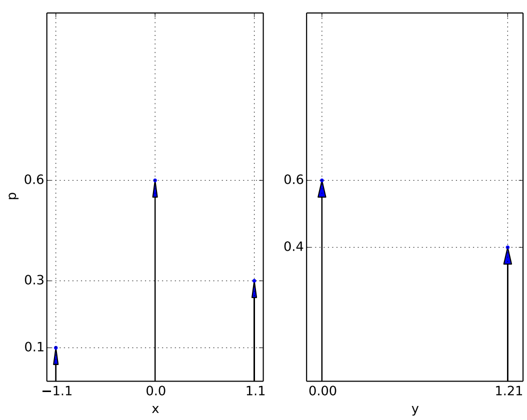


FIGURE 6 – Idem que pour la fig. 5. Noter que $\mathbb{P}(Y = 1.21) = \mathbb{P}(X = -1.1) + \mathbb{P}(X = 1.1)$

2.4 La fonction de répartition d'une v.a. à nombre fini de valeurs réelles

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

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

Les figures 7 et 8 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 constante par morceaux.

Nous définirons la fonction de répartition F_X d'une v.a. réelle X quelconque par la même formule. Cette fonction *caractérise* la distribution de X.

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 $+\infty$.

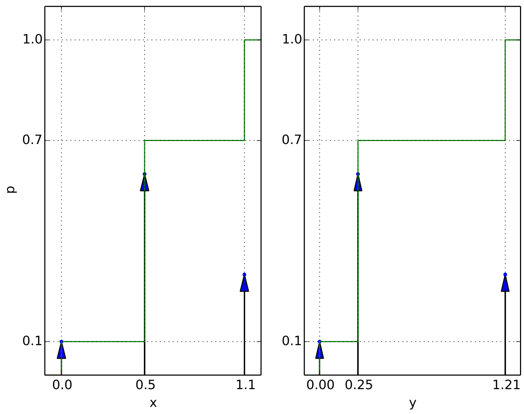


FIGURE 7 – Une variable aléatoire X, $Y = X^2$, distributions et fonctions de répartition.

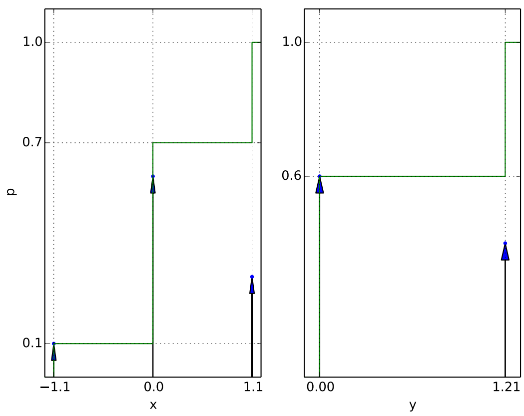


FIGURE 8 – Idem que pour la fig. 7.

2.5 L'espérance et la formule de transfert

Si X est une v.a prenant un nombre fini de valeurs réelles distinctes, $x_1, \ldots, x_K$, on a $X = \sum_{k=1}^{K} x_k 1\!\!1_{\{X=x_k\}}$ et son espérance est

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

Si X est constante, $X = \lambda \in \mathbb{R}$ alors

$$\mathbb{E}(X) = \lambda$$

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

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

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

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

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.

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

$$\mathbb{E}(Z) = \mathbb{E}(\sum_{i=1}^{n} \lambda_i.X_i) = \sum_{i=1}^{n} \lambda_i.\mathbb{E}(X_i)$$

Exercice 5.—
1. Donner l'espérance d'une v.a. uniforme sur $\{0, \ldots, n\}$, d'une v.a uniforme sur $\{1, \ldots, n\}$.
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 $\{0, \ldots, 250\}$, d'une v.a. uniforme sur l'ensemble des entiers naturels impairs inférieurs à 501 (se ramener au cas précédent en retranchant 1).

Démonstration. (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

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

Ce qui est la formule de transfert annoncée. □

En complément, 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),

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

Exercice 6.— 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 7.— 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 $h(Z) = h(X,Y) = X.Y$.

Exercice 8.—
1. Retrouver la formule $\frac{(n+1)^2-1}{12}$ donnant la variance d'une variable uniforme sur $\{0,\ldots,n\}$. Donner la variance d'une v.a. uniforme sur $\{1,\ldots,n\}$.

Indication:On rappelle que

$$\sum_{k=1}^{n} k = \frac{n.(n+1)}{2} \text{ et } \sum_{k=1}^{n} k^2 = \frac{n.(n+1).(2n+1)}{6}$$

2. Donner la variance d'une v.a. uniforme sur l'ensemble des entiers naturels pairs inférieurs à 500, d'une v.a. uniforme sur l'ensemble des entiers naturels impairs inférieurs à 501.

Exercice 9.— 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 $\{0,\ldots,n\}$.

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

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

2.7 Probabilités conditionnelles et indépendance

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 [13] importante : « A sachant B » ou « $(A|B)$ » **n'est pas** un événément. 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 !

D'un point de vue pratique, donner des probabilités conditionnelles est souvent une façon de spécifier la modélisation à un niveau fin.

13. On trouve aussi la nottion $\mathbb{P}_B(A)$

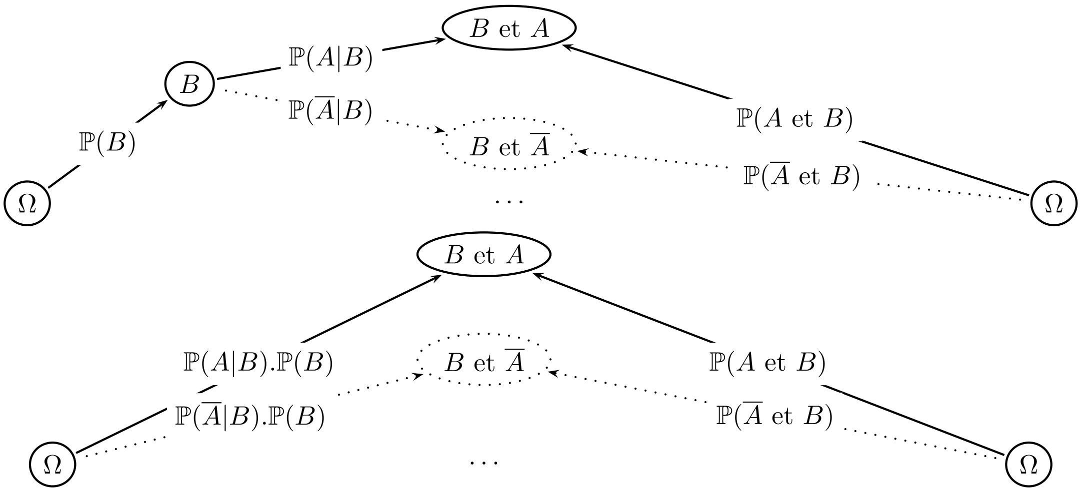
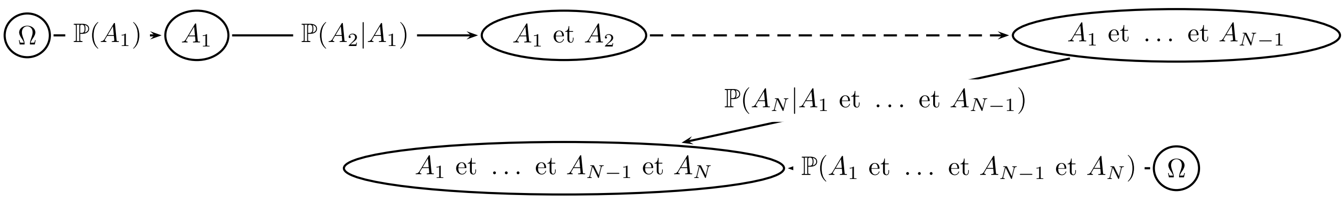


FIGURE 9 – Définition de la probabilité conditionnelle illustrée par deux arbres de choix équivalents. Chaque noeud est un événement. On indique sur la flèche la probabilité conditionnelle du but sachant la source.

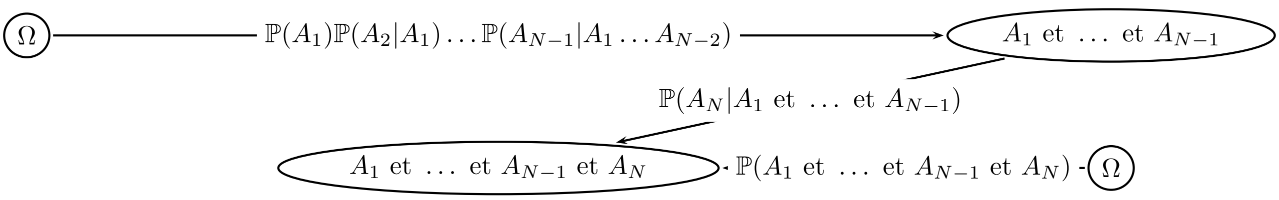
La formule des probabilités composées

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

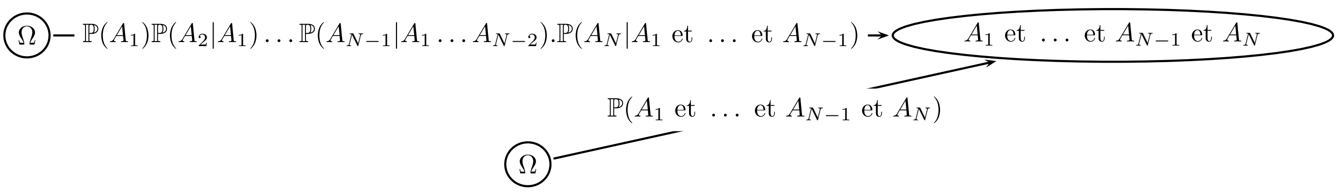
$$\mathbb{P}(A_1 \text{ et } \dots A_N) = \mathbb{P}(\bigcap_{n=1}^{N} A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\dots\mathbb{P}(A_N|A_1 \text{ et } \dots \text{ et } A_{N-1})$$



(a) Décomposition complète.



(b) Application hyp. rec..



(c) Application déf. proba. conditionnelle.

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

Exercice 10.—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 valeurs de X_1 et utiliser le formule des proba. composées.

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

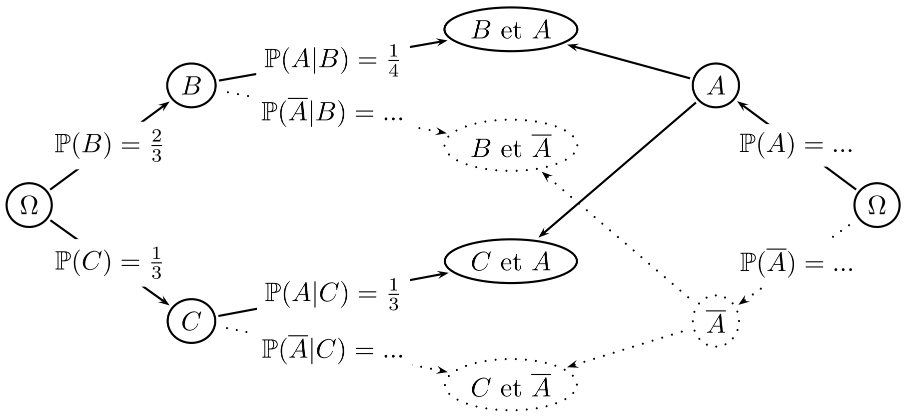
$$\mathbb{P}(A \text{ et } B) = \mathbb{P}(A|B).\mathbb{P}(B) = \frac{1}{4}.\frac{2}{3}, \mathbb{P}(A \text{ et } C) = \mathbb{P}(A|C).\mathbb{P}(C) = \frac{1}{3}.\frac{1}{3}$$

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

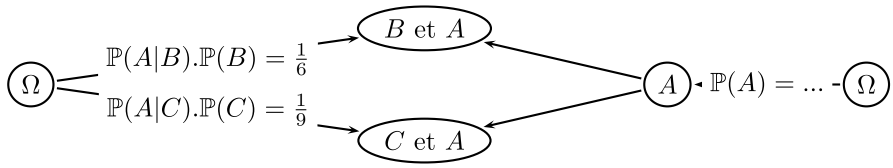
$$\mathbb{P}(A) = \mathbb{P}(A \text{ et } B) + \mathbb{P}(A \text{ et } C) = \frac{5}{18}$$

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.

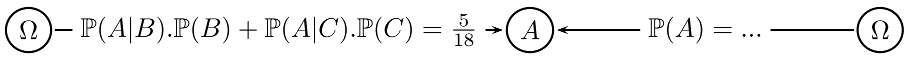
On peut illustrer le calcul fait par une suite de schémas, *c.f.* Fig. 11, à interpréter proprement.



(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.

Exercice 11.— 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.

La formule des probabilités totales

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

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

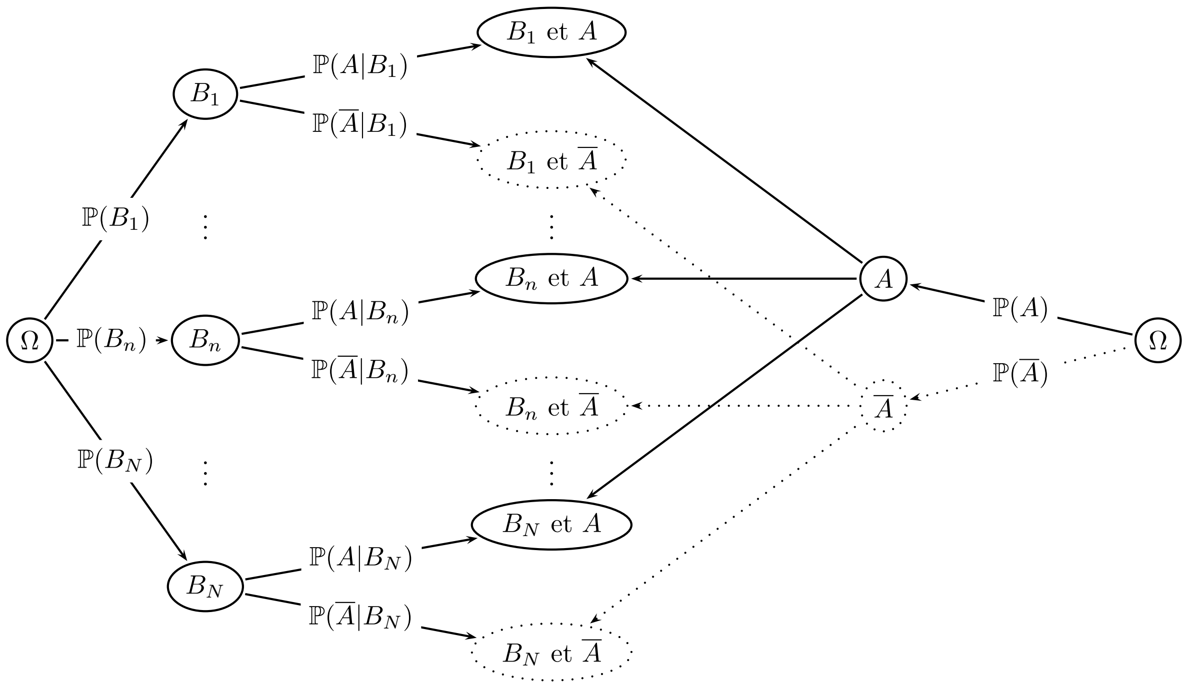


FIGURE 12 – Arbre de choix et formule des probabilités totales. Chaque noeud source est un événement, que l'on décompose (flèches) en union disjointe d'événements. On indique sur la flèche la probabilité conditionnelle du but sachant la source.

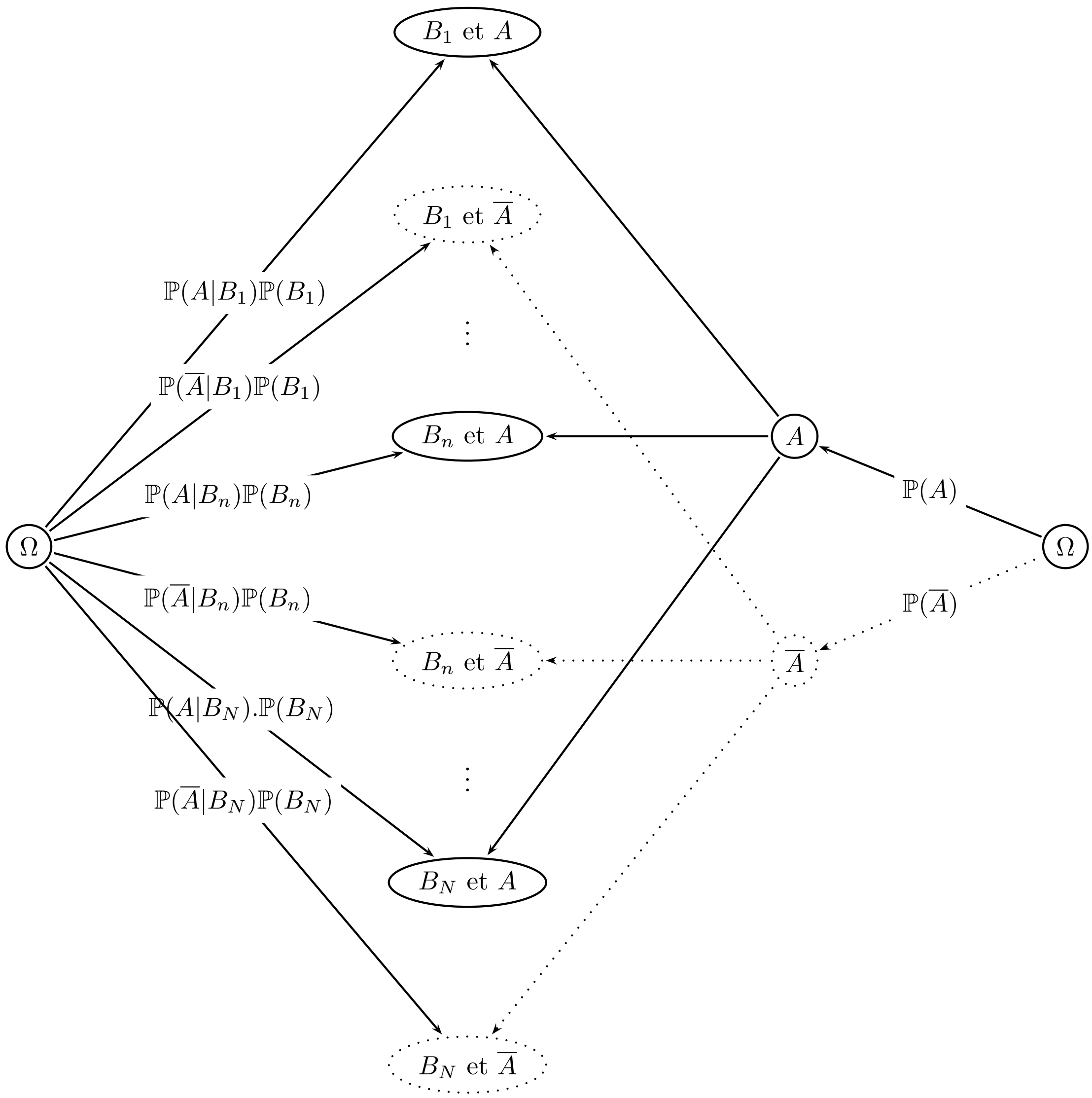


FIGURE 13 – Simplification en utilisant la déf. de proba. conditionnelle.

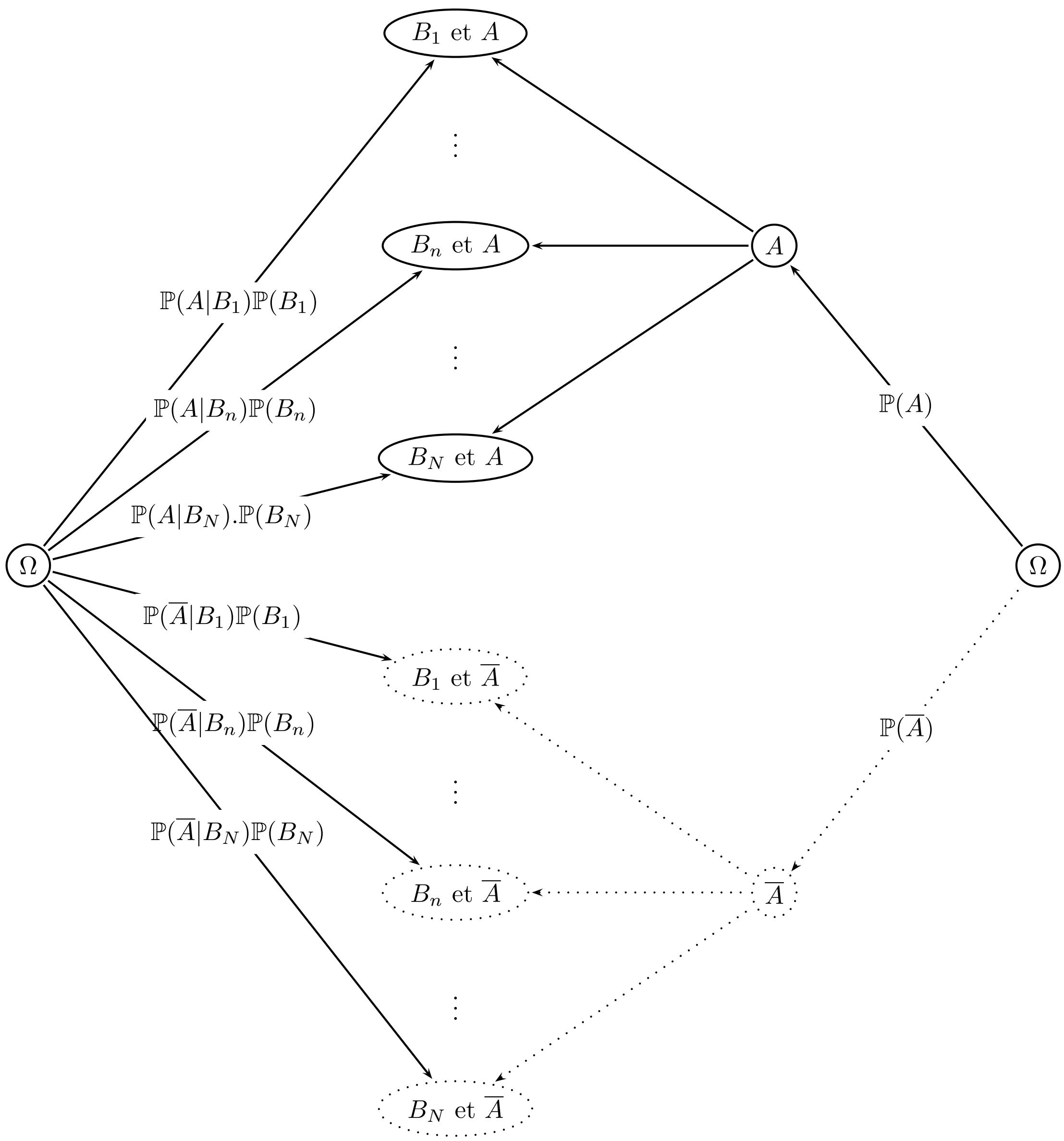


FIGURE 14 – Réordonnancement.

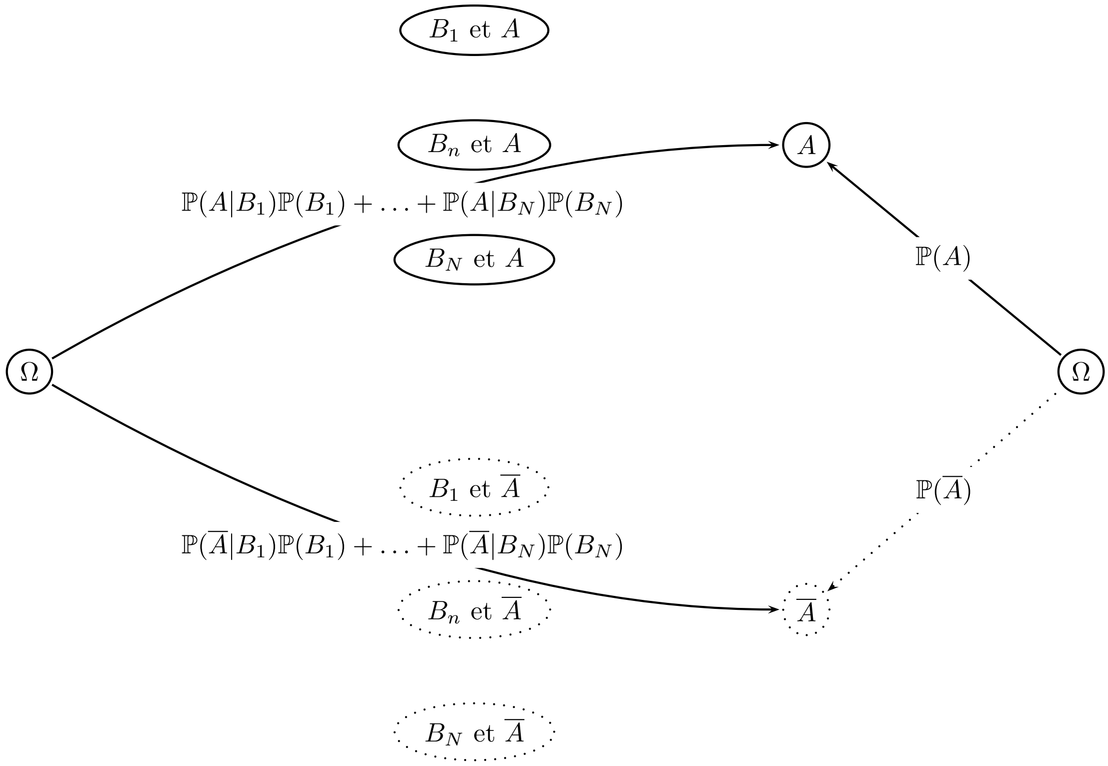


FIGURE 15 – Simplification en utilisant l'axiome d'additivité.

Espérance conditionnelle, 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 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* [14]. 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 12.— 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 3.

Exercice 13.—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 4. *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.

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

Indépendance de variables aléatoires

Définition 5. *Soient X et Y deux variables aléatoires réelles.*

— *X et Y sont dites indépendantes si pour tous intervalles I et J, les événements $\{X \in I\}$ et $\{Y \in J\}$ sont indépendants i.e.*
$$\mathbb{P}(X \in I \text{ et } Y \in J) = \mathbb{P}(X \in I).\mathbb{P}(Y \in J)$$

— *X et Y sont indépendantes ssi pour toutes fonctions $f, g : \mathbb{R} \to \mathbb{R}$,*
$$\mathbb{E}(f(X).g(Y)) = \mathbb{E}(f(X)).\mathbb{E}(g(Y))$$

Rq : On a une définition similaire pour définir l'indépendance mutuelle d'une famille de n variables aléatoires :

$X_1, \dots, X_n$ sont mutuellement indépendantes ssi pour toutes fonctions $f_1 \dots, f_n : \mathbb{R} \to \mathbb{R}$,
$$\mathbb{E}(f_1(X_1).\dots.f_n(X_n)) = \mathbb{E}(f_1(X_1)).\dots.\mathbb{E}(f_n(X_n))$$

Exercice 14.— On dispose de 3 urnes U_1, U_2 et U_3 comportant chacune n boules numérotées de 1 à n. On tire une boule dans chaque urne et on note X_i le numéro tiré dans U_i.
1. Donner la loi de $X_1 + X_2$.
2. Quelle est la probabilité pour que le numéro d'une boule tirée soit la somme des deux autres.

3 Des exemples fondamentaux de modèles

3.1 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* $\{1, \dots, 6\}$.

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,

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

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

par incompatibilité,

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

et, par indépendance,

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

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, $\mathbb{P}((X_1,X_2)=(x_1,x_2))=\frac{1}{36}$.

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

Exercice 15.— 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 16.— 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 $\mathbb{E}(X_1)$.

3.2 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. $S = X_1 + \cdots + X_n$, 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 $k \in \{1,\dots,n\}$, $S_k = X_1 + \cdots + X_k$. On a, pour $1 \leq k \leq n-1$, $S_{k+1} = S_k + X_{k+1}$ 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 $t \in \mathbb{R}$,

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

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

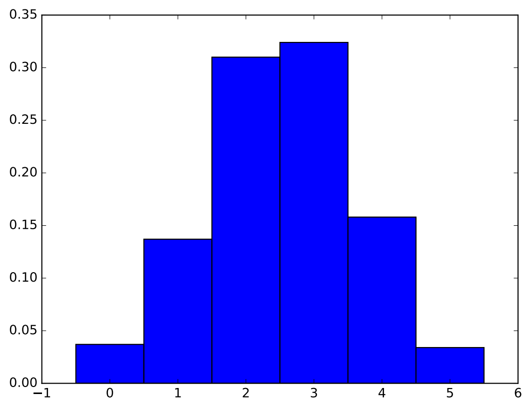


FIGURE 16 – Simulation d'une expérience de BERNOULLI, $n = 5$, $p = 0.5$. simulation sur $N = 1000$ expériences

Utiliser le script `simulation-binomiale.py`.

3.3 Tirer des boules sans remise, modèle hypergéométrique

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 [15]

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

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 $p.N = 3$, $q.N = 2$, $N = 5$, $n = 3$

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

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

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

15. La formule de la variance est HP

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

16. 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`.

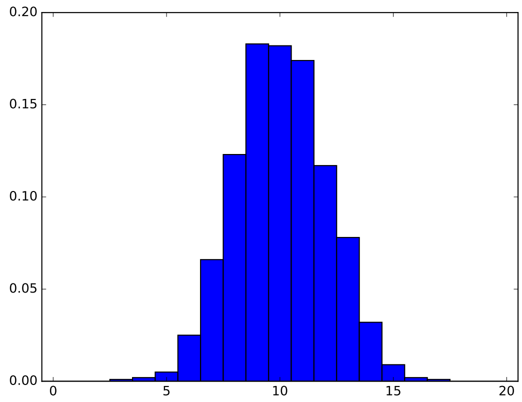


FIGURE 17 – Simulation de tirages sans remise $N = 100$, $n = 50$, $p.N = 20$, simulation sur $NS = 1000$ expériences

Exercice 17.— 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.

3.4 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 »,

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

i.e.

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

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

Représentation graphique de chaîne

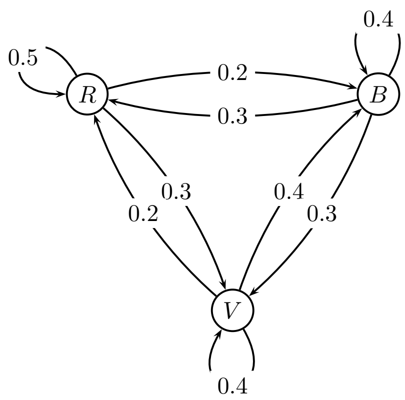


FIGURE 18 – Représentation graphique de l'exemple : on place les états et, sur des flèches, les probabilités de transition d'un état à l'autre

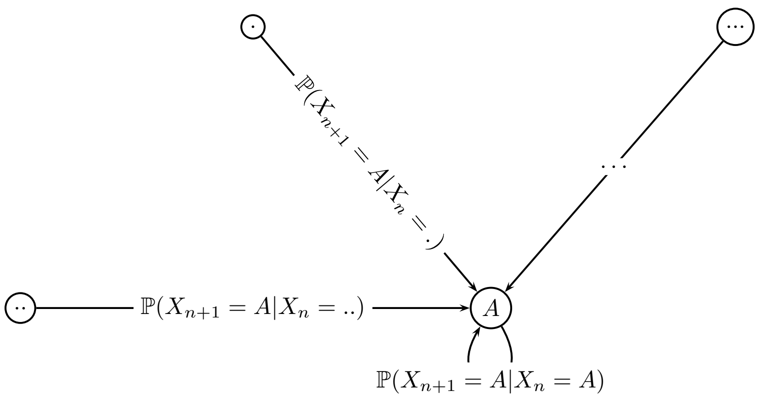


FIGURE 19 – Un noeud, probabilités de l'atteindre, la ligne A de la matrice de transition

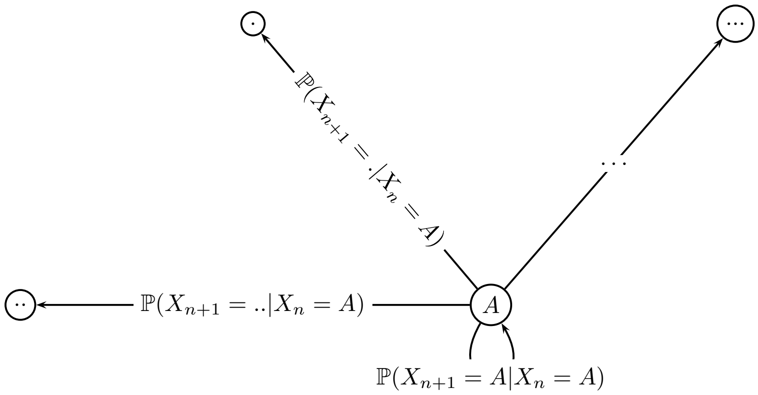


FIGURE 20 – Un noeud, probabilités des destinations, la colonne A de la matrice de transition, leur somme vaut 1

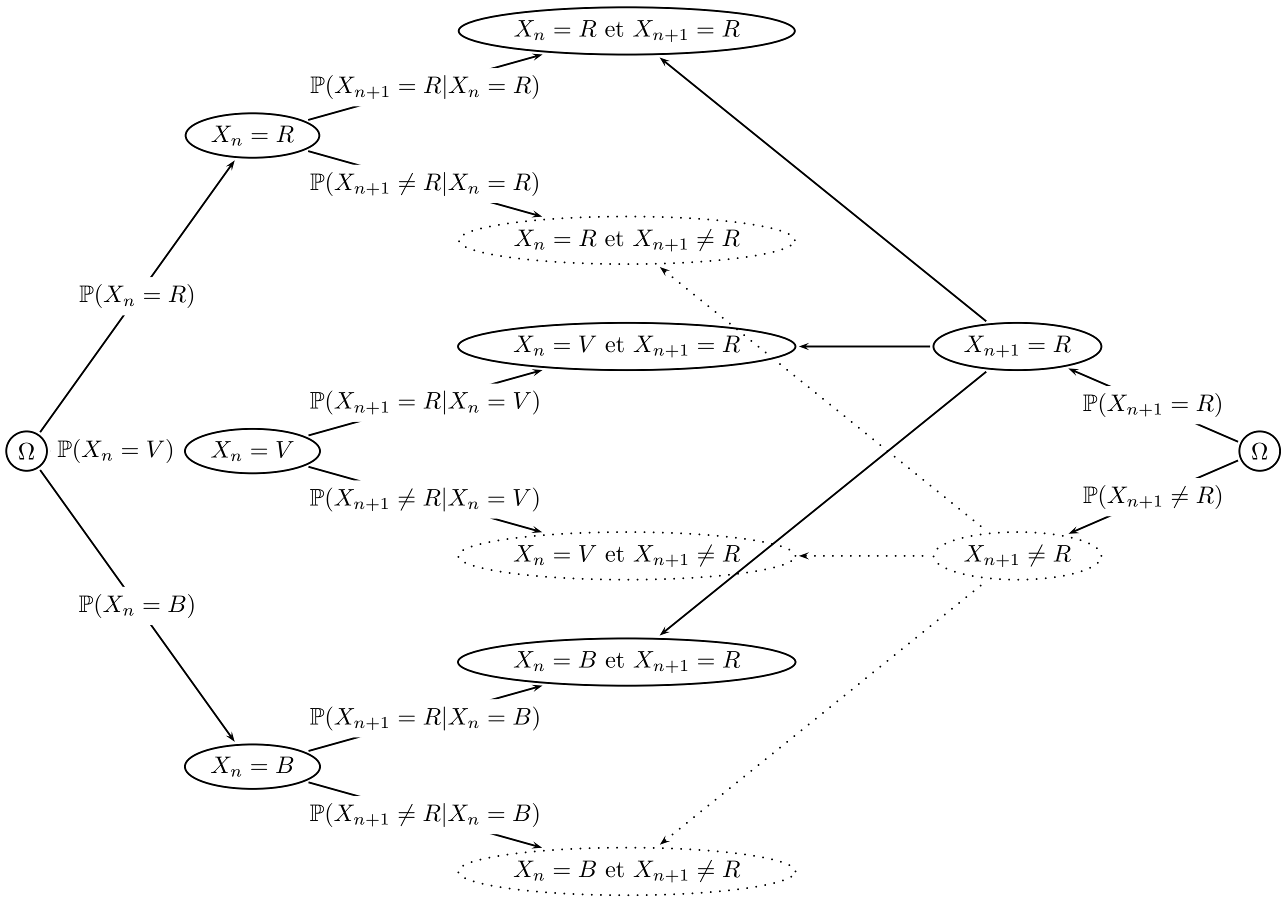


FIGURE 21 – Représentation de l'arbre des choix (initial) pour le calcul de $\mathbb{P}(X_{n+1} = R)$. Cet arbre se traite ensuite comme celui des Fig. 12,13,14 et 15

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`.

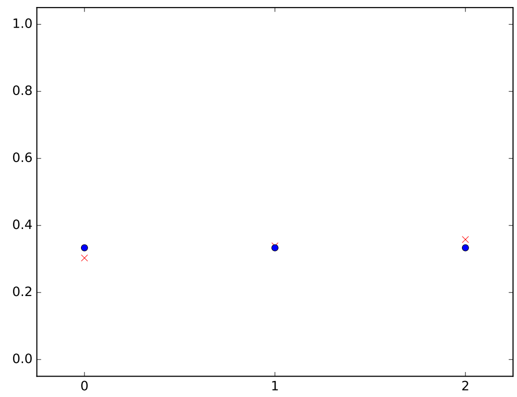


FIGURE 22 – Simulation d'une chaîne de MARKOV, $N = 10$, A et P_0 donné dans le texte. Simulation sur $NS = 100$ expériences. ○=théorique, ×=simulé

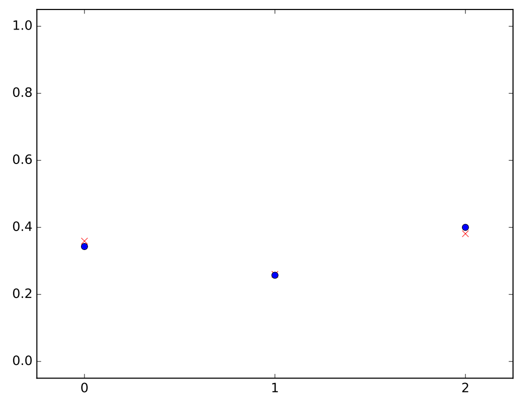


FIGURE 23 – Simulation d'une chaîne de MARKOV, $N = 10$, A', P_0 donné dans le texte. Simulation sur $NS = 1000$ expériences. ○=théorique, ×=simulé

Exercice 18.—

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 19.— 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; toujours d'après des statistiques, on estime les probabilités suivantes :

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

Va-t-il finir par s'arrêter ?

Exercice 20.— 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 :

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

4 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

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

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

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

A.3.

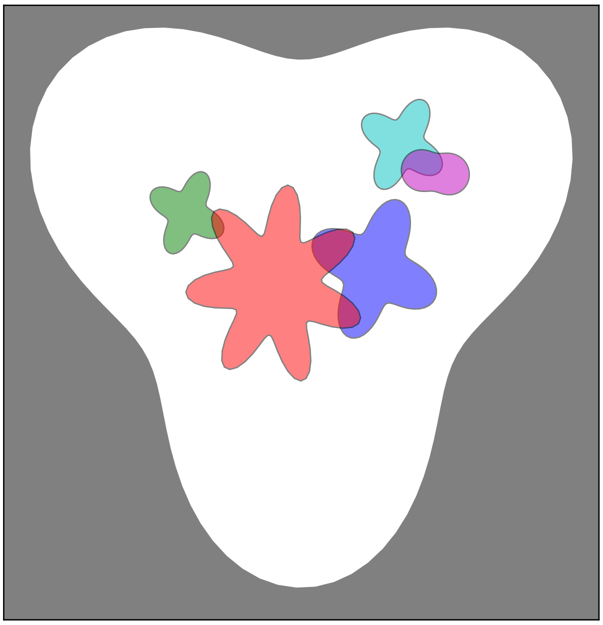


FIGURE 24 – Représentation graphique de fonctions indicatrices. L'univers E est en blanc, chaque « étoile » colorée représente une partie de E. Les nommer $A_1,\dots,A_5$ de gauche à droite, et, en reproduisant partiellement le dessin, représenter successivement $1\!\!1_{A_1}$, $1\!\!1_{A_1\cup A_2\cup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

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.

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

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

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

B.3. Expliquer en quoi [17] 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. $1\!\!1_A \le 1\!\!1_B$, inégalité à prendre au sens des fonctions, *i.e.* $\forall x \in E,\ 1\!\!1_A(x) \le 1\!\!1_B(x)$.

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

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

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

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

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 $a_1,\dots a_N$ sont des nombres complexes,

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

la somme précédente ayant lieu sur toutes [19] les parties P de $\{1,\dots,N\}$. $\#P$ 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)
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

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

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

19. au nombre de 2^N

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 $A_1, \ldots, A_N$ est une famille d'événements, on a

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

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.

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

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

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.*

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

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

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

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.