© 2023 - Gérard Lavau - https://gerardlavau.fr
Vous avez toute liberté pour télécharger, imprimer, photocopier ce cours et le diffuser gratuitement. Toute diffusion à titre onéreux ou utilisation commerciale est interdite sans accord de l'auteur.
Si vous êtes le gestionnaire d'un site sur Internet, vous avez le droit de créer un lien de votre site vers mon site, à condition que ce lien soit accessible librement et gratuitement. Vous ne pouvez pas télécharger les fichiers de mon site pour les installer sur le vôtre.

PROBABILITES - 1ère partie

Plan

I : Définitions

1- Univers

On considère une expérience aléatoire (ou épreuve) dont le résultat ω (ou l'issue, ou la réalisation) appartient à un ensemble Ω, appelé **univers**. Le choix de Ω suppose une modélisation de l'expérience réalisée. Ω peut être :

❑ *fini* . Exemples :

- Lancer un dé et lire le numéro sorti : $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Jouer au loto : Ω est par exemple l'ensemble de toutes les combinaisons possibles de 5 éléments parmi 49. Il possède $\binom{49}{5}$ éléments, soit 1 906 884.
- Lancer une pièce : Ω = {Pile, Face}

Ces ensembles font l'objet de ce chapitre.

❑ *infini dénombrable*. On désigne ainsi les ensembles en bijection avec $\mathbb{N}$. En général, $\Omega = \mathbb{N}$ lui-même. On peut également choisir un tel Ω lorsque le nombre de résultats est fini, mais borné par un nombre inconnu. Les ensembles finis ou dénombrables sont appelés ensembles discrets. Exemples :

- Lancer une pièce jusqu'à ce qu'il apparaisse P. Au choix, ou bien $\Omega = \{F^nP \mid n \in \mathbb{N}\}$ où F^nP désigne un suite de n Faces suivi d'un Pile ; ou bien $\Omega = \mathbb{N}$ si on considère le résultat comme le nombre de lancers.
- Compter le nombre de voitures passant au péage de l'autoroute à Fontainebleau entre le 30 Juillet 0 h et le 1 Août 24 h. $\Omega = \mathbb{N}$.
- Prendre un épi de blé et compter le nombre de grains. $\Omega = \mathbb{N}$.
- Se promener aléatoirement sur un axe de la façon suivante : i) On lance une pièce. Si elle tombe sur Pile, on avance d'une unité, sinon on recule d'une unité. ii) On relance la pièce. Si elle tombe sur Pile, on s'arrête. Sinon on recommence à l'étape i). Le résultat de l'épreuve est l'abscisse obtenue lorqu'on s'arrête. $\Omega = \mathbb{Z}$.

Ces ensembles font l'objet du chapitre L2/PROBA2.PDF.

❑ *infini non dénombrable*. En général, $\Omega = \mathbb{R}$ ou $\mathbb{R}^2$ ou un intervalle de $\mathbb{R}$. Exemples :

- Mesurer la durée écoulée jusqu'à observer la désintégration d'un élément radioactif que l'on observe. $\Omega = \mathbb{R}^+$.
- Chercher la proportion de pièces défectueuses fabriquées par une usine. $\Omega = [0, 1]$.
- Lancer une pièce indéfiniment. Ω est l'ensemble des suites formées des lettres P et F, qu'on note parfois $\{P, F\}^{\mathbb{N}} = \{(x_i)_{i\in\mathbb{N}} \mid \forall\, i \in \mathbb{N}, x_i = P \text{ ou } x_i = F\}$. Cet ensemble n'est pas dénombrable mais n'est pas non plus un intervalle de $\mathbb{R}$.

Ces ensembles font l'objet du chapitre L3/PROBA3.PDF

Dans la suite de ce chapitre, Ω désigne un ensemble fini.

2- Evénement

a) Définition

On appelle **événement** *une partie* A *de* Ω.

A n'est autre qu'un ensemble de résultats, inclus dans l'ensemble Ω de tous les résultats possibles. L'ensemble des événements est donc l'ensemble des parties de Ω, noté $\mathcal{P}(\Omega)$.

EXEMPLES :

- ❑ Tirer au dé un nombre pair. A = {2, 4, 6}.
- ❑ Tirer au loto cinq nombres successifs. A = {(1, 2, 3, 4, 5) , (2, 3, 4 ,5, 6) , ... , (45, 46, 47, 48, 49)}.
- ❑ Lancer une pièce et tomber sur Face. A = {Face}.

L'événement A est **réalisé** si l'issue de l'épreuve appartient à A. En reprenant les trois exemples ci-dessus, donnons des exemples d'épreuves en indiquant respectivement si l'événement A est réalisé :

4	A est réalisé
(2, 3, 5, 7, 8)	A est non réalisé
Face	A est réalisé

Ω s'appelle **événement certain**. En effet, par définition, l'issue de l'épreuve appartient à Ω et Ω est toujours réalisé au cours d'une épreuve. Pour une raison opposée, $\varnothing$ est appelé **événement impossible**.

Un **événement élémentaire** est constitué d'un singleton. Il est réduit à un élément. Ω est la réunion des événements élémentaires.

b) Opérations sur les événements

Les opérations sur les événements sont celles qu'on peut effectuer sur les parties de Ω, à savoir les opérations usuelles des ensembles.

❑ Réunion : $A \cup B$ est réalisé si et seulement si A ou B est réalisé, ("ou" est pris au sens inclusif, c'est-à-dire que A et B peuvent être réalisés simultanément)

❑ Intersection : $A \cap B$ est réalisé si et seulement si A et B sont réalisés simultanément. Si $A \cap B = \varnothing$, on dit que A et B sont **incompatibles**. Il est impossible de voir se réaliser simultanément ces deux événements.

On pourra vérifier que, I étant une famille d'indice quelconque, on a :

$$\left(\bigcap_{i\in I} A_i\right) \cup B = \bigcap_{i\in I} (A_i \cup B)$$

$$\left(\bigcup_{i\in I} A_i\right) \cap B = \bigcup_{i\in I} (A_i \cap B)$$

❑ Complémentaire : A^c, parfois noté $\overline{A}$, est l'ensemble des éléments de Ω qui ne sont pas dans A. Cet événement est réalisé si et seulement si A n'est pas réalisé. Pour cela, on l'appelle l'événement **contraire** de A. A et A^c sont incompatibles. On notera les propriétés suivantes dont la démonstration est laissée au lecteur :

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$
$$(A^c)^c = A$$

On pourra vérifier que, I étant une famille d'indices quelconque, on a :

$$\left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c$$

$$\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c$$

❑ Différence : A – B ou A \ B est l'ensemble des éléments de A qui ne sont pas dans B. (B n'est pas nécessairement inclus dans A). Cet événement est réalisé si et seulement si A est réalisé, mais pas B. On notera que :

$$A - B = A \cap B^c$$

A – B et B sont incompatibles.

c) Système complet d'événements

$(A_i)_{i \in I}$ forme un **système complet d'événements** si et seulement si les trois conditions suivantes sont vérifiées :

i) $\forall\ i \in I,\ A_i \neq \varnothing$

ii) $\forall\ i \neq j,\ A_i \cap A_j = \varnothing$

iii) $\Omega = \bigcup_{i \in I} A_i$

Du point de vue ensembliste, $(A_i)_{i \in I}$ forme une partition de Ω. La condition (i) n'est pas essentielle en probabilité et est parfois omise.

On observe que les mêmes situations sont désignées par des mots différents suivant que l'on parle d'ensembles ou d'événements. Cela est dû au fait que, historiquement, les probabilités se sont développées indépendamment de la théorie des ensembles, à partir du XVIIème. La théorie des ensembles date de la fin du XIXème, et ce n'est qu'à partir du début du XXème que les probabilités ont connu un essor important, grâce au développement des théories des ensembles et de l'intégration. Voici un résumé de la correspondance entre le vocabulaire des ensembles et celui des probabilités.

ENSEMBLES	*PROBABILITES*
partie A de Ω	événement A de l'univers Ω
ensemble Ω	événement certain
ensemble vide ∅	événement impossible
singleton $\{x\}$	événement élémentaire $\{x\}$
ensembles disjoints	événements incompatibles
complémentaire	événement contraire
partition	système complet d'événements

3- Espace probabilisé

a) Equiprobabilité

Soit $\Omega = \{x_1, x_2, \ldots , x_n\}$. On définit une **loi équiprobable** sur Ω, ou une **loi uniforme**, en posant que la probabilité est la même pour tous les événements élémentaires. La probabilité de l'événement certain étant égale à 1, cela nous amène à définir :

$$\mathbf{P}(\{x_i\}) = \frac{1}{n}$$

$\mathbf{P}(A) = \dfrac{\text{Card}(A)}{\text{Card}(\Omega)}$ pour tout événement A (nombre de cas favorables sur nombre de cas possibles), où Card(A) désigne le nombre d'éléments de A.

EXEMPLES :

❑ Pour un dé équilibré, on posera $\mathbf{P}(\{x\}) = \dfrac{1}{6}$ pour $x \in \{1, 2, 3, 4, 5, 6\}$

❑ Pour une pièce équilibrée, $\mathbf{P}(\{\text{Pile}\}) = \mathbf{P}(\{\text{Face}\}) = \dfrac{1}{2}$

❑ Au loto $\mathbf{P}(\{x\}) = \dfrac{1}{1\ 906\ 884}$ pour chaque combinaison x de 5 chiffres parmi 49.

Le choix de l'équiprobabilité relève d'une modélisation d'une situation où une symétrie parfaite joue un rôle important (pièce ou dé parfaitement symétrique, boules jouant des rôles équivalents).

b) Probabilité quelconque

Plus généralement, on peut définir une probabilité quelconque sur $\Omega = \{x_1, ..., x_n\}$ à partir de la probabilité des événements élémentaires. On se donne n valeurs $p_1, ..., p_n$ éléments de [0, 1] tels que $\sum_{i=1}^{n} p_i = 1$, et on pose $p_i = \mathbf{P}(\{x_i\})$ pour tout élément x_i de Ω. Pour toute partie A de Ω, on pose :

$$\mathbf{P}(A) = \sum_{i \text{ tel que } x_i \in A} p_i = \sum_{i=1}^{n} p_i\, \mathbf{1}_A(x_i)$$

où $\mathbf{1}_A$ s'appelle **fonction indicatrice** de la partie A. Elle est définie de la façon suivante :

$$\mathbf{1}_A : \Omega \to \{0,1\}$$
$$x \to \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ si } x \notin A \end{cases}$$

La condition $\sum_{i=1}^{n} p_i = 1$ est imposée, de facon que :

$$\mathbf{P}(\Omega) = 1 = \sum_{i=1}^{n} p_i$$

EXEMPLE :

❑ Pour une pièce truquée, on posera : $\mathbf{P}(\{\text{Pile}\}) = p$ et $\mathbf{P}(\{\text{Face}\}) = q = 1 - p$, $p \in [0, 1]$

c) Définition abstraite

Les probabilités précédentes vérifient la propriété suivante : si A et B sont deux événements incompatibles (i.e. deux parties disjointes de Ω), alors $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$. En effet :

$$\mathbf{P}(A \cup B) = \sum_{i=1}^{n} p_i\, \mathbf{1}_{A \cup B}(x_i)$$

$$= \sum_{i=1}^{n} p_i\, (\mathbf{1}_A(x_i) + \mathbf{1}_B(x_i)) \qquad \text{car A et B sont disjoints}$$

$$= \sum_{i=1}^{n} p_i \mathbf{1}_A(x_i) + \sum_{i=1}^{n} p_i \mathbf{1}_B(x_i)$$

$$= \mathbf{P}(A) + \mathbf{P}(B)$$

Pour des raisons de simplicité de présentation ou de rapidité de démonstration, il est commode de prendre cette propriété comme définition d'une probabilité :

DEFINITION

Soit Ω *un univers fini. On appelle* **probabilité** *sur* Ω *une application* P *de* $\mathcal{P}(\Omega)$ *dans* [0, 1] *telle que* :

$\mathbf{P}(\Omega) = 1$

Pour tout $A \subset \Omega$, $B \subset \Omega$, $A \cap B = \varnothing \Rightarrow \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$

$(\Omega, \mathbf{P})$ *s'appelle alors* **espace probabilisé**.

Cette définition sera généralisée de manière légèrement différente pour un univers infini dans le chapitre L2/PROBA2.PDF.

Dans le cas fini, on généralise aussitôt la deuxième propriété à une réunion finie quelconque. Pour toute famille $(A_i)_{i \in I}$ d'événements incompatibles, indicée par I ensemble fini, on a :

$$\mathbf{P}(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mathbf{P}(A_i)$$

Cela se montre par récurrence sur le nombre d'éléments de I. Si I est réduit à un seul élément, la relation est triviale. Si I possède deux éléments, la relation résulte de la définition d'une probabilité. Supposons la propriété vérifiée pour les ensembles d'indices de cardinal p. Soit I un ensemble d'indice de cardinal $p + 1$. Soit j un élément quelconque de I, soit $J = I - \{j\}$, soit $A = \bigcup_{i \in J} A_i$ et $B = A_j$. A et B sont disjoints, et :

$$\mathbf{P}(\bigcup_{i \in I} A_i) = \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) = \mathbf{P}(\bigcup_{i \in J} A_i) + \mathbf{P}(A_j)$$

$$= \sum_{i \in J} \mathbf{P}(A_i) + \mathbf{P}(A_j) \qquad \text{en appliquant l'hypothèse de récurrence sur J}$$

$$= \sum_{i \in I} \mathbf{P}(A_i)$$

La définition abstraite d'une probabilité permet de retrouver la définition par somme des probabilités des événements élémentaires. En effet, si on note $\Omega = \{x_1, ..., x_n\}$, et si on pose $p_i = \mathbf{P}(\{x_i\})$, alors :

$$A = \bigcup_{i \text{ tel que } x_i \in A} \{x_i\} \qquad \text{union disjointe}$$

donc $$\mathbf{P}(A) = \sum_{i \text{ tel que } x_i \in A} \mathbf{P}(\{x_i\}) = \sum_{i \text{ tel que } x_i \in A} p_i$$

4- Propriétés des probabilités

La définition des probabilités permet d'en déduire les propriétés suivantes :

PROPOSITION :

i) $\mathbf{P}(\varnothing) = 0$

ii) *Pour tout événement* A *et* B, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

iii) *Pour tout événement* A, $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$

iv) *Pour tout événement* A *et* B *tel que* A *soit inclus dans* B, *on a* :

$$\mathbf{P}(A) \leq \mathbf{P}(B) \text{ et } \mathbf{P}(B - A) = \mathbf{P}(B) - \mathbf{P}(A)$$

Démonstration :

Nous donnons les démonstrations en utilisant la définition abstraite des probabilités. On peut également les vérifier en recourant à la définition des probabilités comme somme finie de probabilités élémentaires, mais ces propriétés sont également valides dans un espace probabilisé infini, où la probabilité n'est pas nécessairement définie de cette façon.

❑ i) Ω et $\varnothing$ sont deux événements incompatibles, donc on a :

$$\mathbf{P}(\Omega) = \mathbf{P}(\Omega \cup \varnothing) = \mathbf{P}(\Omega) + \mathbf{P}(\varnothing)$$

d'où $\mathbf{P}(\varnothing) = 0$.

❑ ii) Pour tout événement A et B, incompatibles ou non, on a :

$$A \cup B = B \cup (A - B)$$
$$A = (A \cap B) \cup (A - B)$$

et les unions des membres de droites sont disjointes. D'où :

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

Le résultat s'en déduit en retranchant membre à membre.

Dans le cas d'une probabilité uniforme, on retrouve la relation :

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B)$$

❑ iii) $\Omega = A \cup A^c$ et cette union est disjointe. D'où :

$$1 = \mathbf{P}(\Omega) = \mathbf{P}(A) + \mathbf{P}(A^c)$$

❑ iv) Si A est inclus dans B, alors $B = A \cup (B - A)$, union disjointe, d'où :

$$\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(B - A)$$

et $\mathbf{P}(B - A)$ est positif ou nul.

II : Probabilités conditionnelles

1- Exemples

On lance un dé. Il sort un nombre pair. Quelle est la probabilité d'avoir sorti un 2 ? Réponse : $\frac{1}{3}$.

Même question avec un dé pipé dont voici la loi de probabilité :

1	2	3	4	5	6
$\frac{2}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{3}{15}$

En moyenne, sur 15 tirages, 8 donnent un nombre pair, et sur ces 8, 3 tirages donnent le 2. On est donc amené à poser la probabilité cherchée égale à $\frac{3}{8}$.

Avec le dé ci-dessus, quelle est la probabilité de sortir un nombre pair, sachant qu'il est sorti un nombre supérieur ou égal à 3 ? Sur 15 tirages, en moyenne 10 donnent un nombre supérieur ou égal à 3, et sur ces 10, 5 donnent un nombre pair. La probabilité cherchée est donc $\frac{1}{2}$.

Quelle est la probabilité de sortir un nombre pair, sachant qu'il est sorti un nombre supérieur ou égal à 2 ? Sur 15 tirages, en moyenne 13 donnent un nombre supérieur ou égal à 2, et sur ces 13, 8 donnent un nombre pair. La probabilité cherchée est donc $\frac{8}{13}$.

Si l'on appelle A l'événement "tirer un nombre pair" et B l'événement tirer un nombre supérieur ou égal à 3 (respectivement 2), on remarque que la probabilité de A sachant que B est réalisé a été calculée comme étant égale à $\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$. Cette situation est à rapprocher de la loi équiprobable sur un ensemble fini, où l'on calcule le quotient du nombre de cas favorable sur le nombre de cas possibles. Ici, on calcule le quotient de la probabilité de l'événement cherché $\mathbf{P}(A \cap B)$ sur la probabilité de l'univers relatif $\mathbf{P}(B)$.

DEFINITION
Soit B *un événement de probabilité non nulle, d'un espace probabilisé* $(\Omega, \mathbf{P})$. *On définit la loi de probabilité sachant* B (*ou* **conditionnelle** *par rapport à* B) *par* :

$$\mathbf{P}_B : \mathcal{P}(\Omega) \to [0,1]$$
$$A \to \mathbf{P}_B(A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

On la note aussi $\mathbf{P}(A \mid B)$.

II est dit que $\mathbf{P}_B$ est une loi de probabilité. Prouvons le :

❑ i) $\mathbf{P}_B$ est clairement définie de $\mathcal{P}(\Omega)$ dans [0, 1] car, puisque $A \cap B \subset B$, $\mathbf{P}(A \cap B) \leq \mathbf{P}(B)$.

❑ ii) $\mathbf{P}_B(\Omega) = \frac{\mathbf{P}(\Omega \cap B)}{\mathbf{P}(B)} = 1$ car $\Omega \cap B = B$

❑ iii) Si A_1 et A_2 sont disjoints, alors $A_1 \cap B$ et $A_2 \cap B$ le sont également et l'on a :

$$\mathbf{P}_B(A_1 \cup A_2) = \frac{\mathbf{P}((A_1 \cup A_2) \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}((A_1 \cap B) \cup (A_2 \cap B))}{\mathbf{P}(B)}$$
$$= \frac{\mathbf{P}(A_1 \cap B) + \mathbf{P}(A_2 \cap B)}{\mathbf{P}(B)} = \mathbf{P}_B(A_1) + \mathbf{P}_B(A_2)$$

Les propriétés i) à iv) du I-4) s'appliquent donc également aux probabilités conditionnelles.

La probabilité $\mathbf{P}_B$ est concentrée sur B. En effet, $\mathbf{P}_B(B) = 1$, ou encore $\mathbf{P}_B(B^c) = 0$.

L'expression de probabilités conditionnelles dans la vie courante conduit parfois à de sérieures confusions. En juin 2006, sur son site internet, un journal de presse d'envergure nationale (et dont on taira le nom) titra : "[...] Pour manger, un salarié sur dix a recours aux associations". Sachant qu'il y avait environ 25 millions de salariés, et que 800.000 personnes recouraient aux associations, la phrase du journal signifiait donc qu'il y avait 2.5 millions de salariés parmi ces 800.000 personnes. On mesure l'absurdité du propos. Le journal voulait probablement dire : "une personne sur dix qui a recours aux associations est un salarié", soit 80.000 salariés ayant recours aux associations (ou encore environ 1 salarié sur 313). Plus précisément, soit A l'événement "être un salarié" et B l'événement "avoir recours à une association". La probabilité correcte est $\mathbf{P}(A \mid B) = \frac{1}{10}$ et le journal a confondu cette probabilité avec $\mathbf{P}(B \mid A) = \frac{1}{313}$. A la suite de remarques de plusieurs internautes, le journal changea le titre de son article sur son site Internet trois jours plus tard en "Pour manger, 16% des retraités ont recours aux associations", commettant exactement la même erreur, cette fois sur les retraités (16% de 13 millions de retraités représentent 2 millions de personnes). Puis, dans l'après-midi du même jour, faute de comprendre les subtilités des probabilités conditionnelles et de trouver une formulation correcte, à bout de ressources, le journal titra : "Pour manger, de plus de plus de personnes ont recours aux associations".

2- Formule des probabilités composées

❑ Cas de deux événements : la **formule des probabilités composées** consiste seulement en la formule suivante, qui résulte de la définition des probabilités conditionnelles.

$$\mathbf{P}(A \cap B) = \mathbf{P}(B) \times \mathbf{P}_B(A)$$

Par hypothèse, on suppose $\mathbf{P}(B) > 0$. On peut cependant étendre cette formule au cas où $\mathbf{P}(B) = 0$ sans que $\mathbf{P}_B(A)$ soit défini. En effet, on a également $\mathbf{P}(A \cap B) = 0$ puisque $\mathbf{P}(A \cap B) \leq \mathbf{P}(B)$, et il suffit d'attribuer la valeur nulle au membre de droite.

EXEMPLE :

❑ Dans un groupe d'individus, il y a :

des fumeurs B

des non-fumeurs B^c

des hommes A

des femmes A^c

(On rappelle que fumer nuit gravement à la santé).

On donne :

i) $\mathbf{P}(B) = \frac{1}{4}$ il s'agit de la proportion de fumeurs parmi la population totale (Cette probabilité est très imprécise. En fait, elle varie fortement avec l'âge. Si elle est de 0,4 pour les individus à 20 ans, elle n'est plus que de 0,2 à 60 ans).

ii) $\mathbf{P}(A) = \frac{1}{2}$ il s'agit de la proportion d'hommes dans la population totale

iii) $\mathbf{P}_B(A) = \frac{3}{5}$ il s'agit de la proportion d'hommes parmi les personnes qui fument

Calculer $\mathbf{P}(A \cap B)$, $\mathbf{P}(A^c \cap B)$, $\mathbf{P}(A \cap B^c)$, $\mathbf{P}(A^c \cap B^c)$, $\mathbf{P}_B(A^c)$, $\mathbf{P}_A(B)$, $\mathbf{P}_A(B^c)$, $\mathbf{P}_{B^c}(A)$, $\mathbf{P}_{B^c}(A^c)$, $\mathbf{P}_{A^c}(B)$, $\mathbf{P}_{A^c}(B^c)$ et interpréter les résultats obtenus.

Réponses :

iv) $\mathbf{P}(A \cap B) = \mathbf{P}(B) \times \mathbf{P}_B(A) = \frac{1}{4} \times \frac{3}{5} = \frac{3}{20}$ (proportion de personnes masculines et qui fument, parmi la population totale)

v) $\mathbf{P}(A \cap B^c) = \mathbf{P}(A) - \mathbf{P}(A \cap B) = \frac{1}{2} - \frac{3}{20} = \frac{7}{20}$ (proportion de personnes masculines et qui ne fument pas, parmi la population totale)

vi) $\mathbf{P}(A^c \cap B) = \mathbf{P}(B) - \mathbf{P}(A \cap B) = \frac{1}{4} - \frac{3}{20} = \frac{1}{10}$ (proportion de personnes féminines et qui fument, parmi la population totale)

vii) $\mathbf{P}(A^c \cap B^c) = \mathbf{P}(A^c) - \mathbf{P}(A^c \cap B) = 1 - \mathbf{P}(A) - \mathbf{P}(A^c \cap B) = 1 - \frac{1}{2} - \frac{1}{10} = \frac{2}{5}$

ou bien $= 1 - \mathbf{P}(B) - \mathbf{P}(A \cap B^c) = 1 - \frac{1}{4} - \frac{7}{20} = \frac{2}{5}$ (proportion de personnes féminines et qui ne fument pas, parmi la population totale).

Ces données peuvent être résumées dans un tableau :

	hommes A	femmes A^c	
fumeurs B	$\frac{3}{20}$	$\frac{1}{10} = \frac{2}{20}$	$\frac{5}{20} = \frac{1}{4}$
non fumeurs B^c	$\frac{7}{20}$	$\frac{2}{5} = \frac{8}{20}$	$\frac{15}{20} = \frac{3}{4}$
	$\frac{10}{20} = \frac{1}{2}$	$\frac{10}{20} = \frac{1}{2}$	

Les autres probabilités ne posent plus de problèmes :

viii) $\mathbf{P}_B(A^c) = \frac{2}{5}$ (proportion de femmes parmi les fumeurs)

ix) $\mathbf{P}_A(B) = \frac{3}{10}$ (proportion de fumeurs parmi les hommes)

x) $\mathbf{P}_A(B^c) = \frac{7}{10}$ (proportion de non-fumeurs parmi les hommes)

xi) $\mathbf{P}_{B^c}(A) = \frac{7}{15}$ (proportion d'hommes parmi les non-fumeurs)

xii) $\mathbf{P}_{B^c}(A^c) = \frac{8}{15}$ (proportion de femmes parmi les non-fumeurs)

xiii) $\mathbf{P}_{A^c}(B) = \frac{1}{5}$ (proportions de fumeurs parmi les femmes)

xiv) $\mathbf{P}_{A^c}(B^c) = \frac{4}{5}$ (proportion de non-fumeurs parmi les femmes)

❑ cas de plusieurs événements :

Soient $A_1, A_2, ..., A_n$, n événements. Alors, on a

$$\mathbf{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbf{P}(A_1) \times \mathbf{P}_{A_1}(A_2) \times \mathbf{P}_{A_1 \cap A_2}(A_3) \times ... \times \mathbf{P}_{A_1 \cap A_2 \cap ... \cap A_{n-1}}(A_n).$$

Cela se montre aisément par récurrence en utilisant le fait que :

$$\mathbf{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbf{P}(A_1 \cap A_2 \cap ... \cap A_{n-1}) \times \mathbf{P}_{A_1 \cap A_2 \cap ... \cap A_{n-1}}(A_n).$$

EXEMPLE :

❑ une urne contient a boules blanches et b boules noires. On effectue des tirages avec remise. Lorsqu'on tire une boule noire, on la remplace par une boule blanche. Quelle est la probabilité de tirer n boules noires de suite ?

Soit A_i l'événement tirer une boule noire au i-ème coup. On a :

$$\mathbf{P}(A_1) = \frac{b}{a+b}$$

$\mathbf{P}_{A_1 \cap A_2 \cap ... \cap A_{i-1}}(A_i) = \dfrac{b-i+1}{a+b}$ puisqu'au i-ème tirage, il y a $i - 1$ boules noires qui ont été remplacées par des blanches. La probabilité cherchée est donc :

$$\mathbf{P}(A_1 \cap A_2 \cap ... \cap A_n) = \frac{b(b-1)(b-2)...(b-n+1)}{(a+b)^n}$$

3- Formule des probabilités totales

EXEMPLE :

❑ On considère n usines fabriquant un même produit. Soit p_i la part de marché de la i-ème usine (proportion du nombre de pièces fabriquées par la i-ème usine sur le nombre de pièces fabriquées par toutes les usines). p_i est la probabilité pour que, prenant une pièce au hasard sur l'ensemble du marché, elle provienne de la i-ème usine. Soit A_i cet événement. $\mathbf{P}(A_i) = p_i$. On remarquera également que $(A_1, A_2, ..., A_n)$ forme un système complet d'événements.

Par ailleurs, chaque usine fabrique un certain nombre de pièces défectueuses. Soit q_i la proportion de pièces défectueuses fabriquées par la i-ème usine, relativement à sa propre production. Si B est l'événement "tirer une pièce défectueuse en choisissant une pièce au hasard", on a $q_i = \mathbf{P}_{A_i}(B)$.

Nous prendrons comme exemple numérique :

$p_1 = 0{,}900$
$p_2 = 0{,}095$
$p_3 = 0{,}005$ (la première usine a quasiment le monopole du marché)
$q_1 = 0{,}01$ (la première usine a la meilleure qualité)
$q_2 = 0{,}05$
$q_3 = 0{,}20$

(la forte proportion de pièces défectueuses fabriquées par la troisième usine explique peut-être sa faible implantation !!)

La question qu'on se pose est la suivante : quelle est la probabilité de tirer une pièce défectueuse en choisissant une pièce au hasard dans la production totale ? Autrement dit, calculer $\mathbf{P}(B)$.

La proportion de pièces défectueuses fabriquées par la i-ème usine, relativement au nombre total de pièces fabriquées par toutes les usines, est $p_i q_i$. La probabilité cherchée est donc $\sum p_i q_i$. L'exemple numérique donne 0,01475 (de l'ordre d'une chance sur 100).

Ceci s'énonce dans le cas général de la façon suivante :

FORMULE DES PROBABILITES TOTALES

Soit $(A_i)_{1\leq i\leq n}$ *un système complet d'événements d'un espace probabilisé* $(\Omega, \mathbf{P})$, *et* B *un autre événement de cet espace. Alors* :

$$\mathbf{P}(B) = \sum_{i=1}^{n} \mathbf{P}(A_i) \times \mathbf{P}_{A_i}(B)$$

<u>Démonstration</u> :

❑ $(A_i)_{1\leq i\leq n}$ formant un système complet d'événements, on a :

$$\Omega = \bigcup_{i=1}^{n} A_i \qquad \text{union disjointe.}$$

D'où
$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(\Omega \cap B)\\ &= \mathbf{P}\Big(\big(\bigcup_{i=1}^{n} A_i\big) \cap B\Big)\\ &= \mathbf{P}\Big(\bigcup_{i=1}^{n} (A_i \cap B)\Big) \qquad \text{cette union étant disjointe}\\ &= \sum_{i=1}^{n} \mathbf{P}(A_i \cap B)\\ &= \sum_{i=1}^{n} \mathbf{P}(A_i) \times \mathbf{P}_{A_i}(B)\end{aligned}$$

4- Formule De Bayes (ou probabilité des causes)

Nous reprenons l'exemple précédent. On tire au hasard une pièce sur l'ensemble du marché. Elle est défectueuse. Quelle est la probabilité qu'elle provienne de la i-ème usine ? On cherche $\mathbf{P}_B(A_i)$. C'est égal à $\dfrac{\mathbf{P}(B \cap A_i)}{\mathbf{P}(B)}$. En remplaçant par les expressions trouvées au paragraphe précédant, on obtient :

PROPOSITION (Formule de Bayes)

Soit $(A_i)_{1\leq i\leq n}$ *un système complet d'événements d'un espace probabilisé* $(\Omega, \mathbf{P})$, *et* B *un autre événement de cet espace tel que* $\mathbf{P}(B) > 0$. *Alors* :

$$\mathbf{P}_B(A_i) = \frac{\mathbf{P}(B \cap A_i)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A_i) \times P_{A_i}(B)}{\sum_{j=1}^{n} \mathbf{P}(A_j) \times P_{A_j}(B)}$$

Dans notre exemple des usines, le calcul donne :

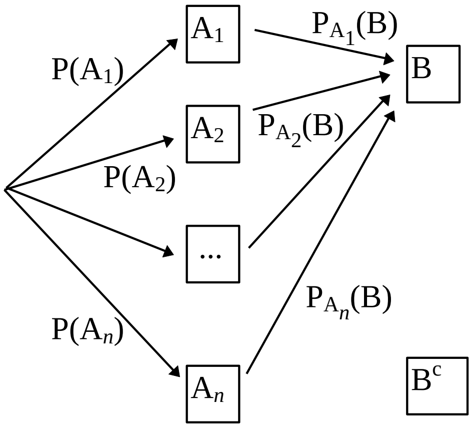
❑ pour la première usine 0,61

❑ pour la deuxième usine 0,32

❑ pour la troisième usine 0,07

II peut paraître étonnant que la probabilité que la pièce défectueuse provienne de la troisième usine soit si faible, alors que cette usine est de mauvaise qualité, mais elle n'intervient que fort peu sur le marché. Par contre, la probabilité que la pièce défectueuse provienne de la première usine est très forte, car elle a presque le monopole de fabrication, malgré la qualité de sa fabrication.

La situation générale peut être visualisée par le schéma suivant :



La probabilité d'aller en B en passant par A_i est $\mathbf{P}(B \cap A_i) = \mathbf{P}(A_i) \times \mathbf{P}_{A_i}(B)$

La probabilité d'aller en B est la somme des probabilités sur tous les chemins possibles. Cela donne la formule des probabilités totales :

$$\mathbf{P}(B) = \sum_{i=1}^{n} \mathbf{P}(A_i) \times \mathbf{P}_{A_i}(B)$$

La probabilité d'être passé par A_i sachant qu'on est arrivé en B est le quotient de la probabilité du i-ème chemin sur la somme des probabilités de tous les chemins menant à B (un analogue du nombre de cas favorables sur le nombre de cas possibles). Cela donne la formule de Bayes :

$$\mathbf{P}_B(A_i) = \frac{\mathbf{P}(B \cap A_i)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A_i) \times P_{A_i}(B)}{\sum_{j=1}^{n} \mathbf{P}(A_j) \times P_{A_j}(B)}$$

5- Evénements indépendants

DEFINITION

Deux événements A *et* B *sont dits* **indépendants** *si* $\mathbf{P}(A \cap B) = \mathbf{P}(A) \times \mathbf{P}(B)$.

La justification de cette définition est la suivante. Si $\mathbf{P}(B)$ est non nul, et si A et B sont indépendants, alors on doit s'attendre à ce que $\mathbf{P}(A)$ et $\mathbf{P}_B(A)$ soient égales. En effet, l'indépendance de A et B a pour conséquence que la connaissance de la réalisation de B n'a aucune influence sur celle de A. Or $\mathbf{P}(A) = \mathbf{P}_B(A) \Leftrightarrow \mathbf{P}(A) = \dfrac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$, d'où la définition.

EXEMPLE :

❑ On considère généralement que les tirages successifs avec une pièce, un dé ou au loto sont indépendants. Les tirages précédents n'ont aucune influence sur les tirages suivants. Si la pièce est équilibrée, il y aura toujours une probabilité $\dfrac{1}{2}$ de tirer Pile. En particulier, le fait d'avoir joué toutes les semaines au loto depuis sa création sans jamais gagner n'augmente pas ses chances de gagner au prochain tirage (cela ne les diminue pas non plus d'ailleurs).

DEFINITION

Plusieurs événements A_1, A_2, ..., A_n *sont indépendants si, pour toute sous-famille de p événements* A_{i_1}, A_{i_2}, ... , A_{i_p} (*les indices* i_1, ..., i_p *étant distincts*), *on a* :

$$\mathbf{P}(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_p}) = \mathbf{P}(A_{i_1})\mathbf{P}(A_{i_2})...\mathbf{P}(A_{i_p})$$

Insistons sur le fait que la propriété doit être vérifiée pour TOUTE sous-famille de p événements. On ne peut se contenter de vérifier l'indépendance des événements deux à deux.

EXEMPLE :

❑ On lance deux dés. On considère les événements suivants :

A : le lancer du premier dé est pair

B : le lancer du deuxième dé est pair

C : la somme des deux lancers est paire

II est clair que les événements A et B, issus de deux dés différents sont indépendants, et que les trois événements A, B et C ne le sont pas. En effet la réalisation de A et B entraîne celle de C. Or :

$\mathbf{P}(A) = \mathbf{P}(B) = \mathbf{P}(C) = \frac{1}{2}$ (Pour $\mathbf{P}(C)$, réfléchir un peu)

$\mathbf{P}(A \cap B) = \frac{1}{4} = \mathbf{P}(A \cap C) = \mathbf{P}(B \cap C)$

En effet, les trois événements $A \cap B$, $B \cap C$ et $A \cap C$ sont identiques. Ainsi les événements sont indépendants deux à deux. On a également :

$\mathbf{P}(A \cap B \cap C) = \frac{1}{4}$

car $A \cap B \cap C = A \cap B$. C n'apporte rien de plus. Donc, les trois événements ne sont pas indépendants, bien qu'ils le soient deux à deux.

6- Probabilité produit

On considère n univers finis Ω_i, $1 \leq i \leq n$. On réalise alors n expériences aléatoires x_i supposées indépendantes, la i-ème ayant son issue dans Ω_i. Cela revient à faire une épreuve unique $(x_1, x_2, ..., x_n)$ dont l'issue appartient à l'espace $\Omega = \Omega_1 \times \Omega_2 \times ... \times \Omega_n$. Cet espace s'appelle l'**espace produit** des Ω_i. C'est un espace fini. Si Ω_i est munie d'une probabilité $\mathbf{P}_i$, il est possible de munir Ω d'une probabilité $\mathbf{P}$ définie de la façon naturelle suivante :

$\mathbf{P}(\{(x_1, x_2, ..., x_n)\}) = \mathbf{P}_1\{\{x_1\}) \times \mathbf{P}_2(\{x_2\}) \times ... \times \mathbf{P}_n(\{x_n\})$

$\mathbf{P}$ s'appelle la **probabilité produit** des $\mathbf{P}_i$.

EXEMPLES :

❑ On lance un dé, et on tire une carte d'un jeu de 32 cartes. Alors, $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$, $\Omega_2 = \{c \mid c$ est une carte d'un jeu de 32 cartes$\}$. Ω est l'ensemble des couples formés d'un lancer de dé et d'un tirage de cartes : $\Omega = \{(d, c) \mid d \in \Omega_1, c \in \Omega_2\}$. On a $\mathbf{P}_1(\{d\}) = \frac{1}{6}$, et $\mathbf{P}_2(\{c\}) = \frac{1}{32}$. D'où $\mathbf{P}(\{c, d\}) = \frac{1}{192}$.

❑ On lance un dé n fois, (ou bien n dés). Tous les espaces Ω_i sont égaux à {Pile, Face}. Ω = {Pile, Face}^{n}. Toutes les lois $\mathbf{P}_i$ sont les mêmes : $\mathbf{P}_i(\{\text{Pile}\}) = \frac{1}{2} = \mathbf{P}_i(\{\text{Face}\})$. Pour chaque x élément de Ω, on a $\mathbf{P}(\{x\}) = \frac{1}{2^n}$.

III : Variables aléatoires

l- Loi d'une variable aléatoire

Soit (Ω , **P**) un espace probabilisé fini et E un ensemble quelconque. On appelle **variable aléatoire** une application X de Ω dans E. Le plus souvent, E = $\mathbb{R}$ et X est appelée **variable aléatoire réelle**.

EXEMPLES :

❑ On lance un dé. On gagne un franc s'il tombe sur 5 ou 6, et on perd un franc sinon. Soit X le gain. On a Ω = {1, 2, 3, 4, 5, 6}, et X est une application de Ω dans $\mathbb{R}$ définie par :

$$X(l) = X(2) = X(3) = X(4) = -1$$
$$X(5) = X(6) = 1$$

La probabilité que le gain soit 1 est $\frac{2}{6}$ ou $\frac{1}{3}$; la probabilité qu'il soit –1 est $\frac{4}{6}$ ou $\frac{2}{3}$. Initialement nous avons une probabilité **P** sur Ω, or nous voyons que, naturellement, nous en déduisons une sur {–1, 1} = X(Ω). Cette dernière probabilité s'appelle loi de X et sera notée $\mathbf{P}_X$ (ne pas confondre la notation avec celle d'une probabilité conditionnelle).

❑ On lance deux dés. Soit X la somme des numéros tirés. On a ici comme espace probabilisé Ω^2, où Ω = {1, 2, 3, 4, 5, 6}, muni de la loi uniforme. X peut prendre les valeurs {2, 3, 4, 5, 6, ..., 11, 12}. Chercher la loi de X, c'est chercher la probabilité de sortie de chacun de ces totaux. Le lecteur vérifiera que l'on a :

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}_X(\{k\})$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Par exemple, $\mathbf{P}(\{3\}) = \frac{2}{36}$ car, sur les 36 sorties possibles des deux dés, seules les sorties (1, 2) et (2, 1) ont pour somme 3.
On a donc défini une probabilité $\mathbf{P}_X$ sur {2, 3, 4, ..., 12}. Par exemple, la probabilité que X appartienne à A = {6, 7, 8} est $\frac{16}{36}$. Une curiosité : si les deux dés portent respectivement les numéros {1, 2, 2, 3, 3, 4} et {1, 3, 4, 5, 6, 8} (dés de Sicherman), on obtiendra exactement la même probabilité $\mathbf{P}_X$ pour la somme des deux dés. On peut jouer avec ces deux dés au lieu de deux dés normaux sans jamais voir la différence (pourvu qu'on se limite à faire la somme des deux dés).

Regardons, dans les exemples précédents, comment calculer $\mathbf{P}_X(A)$, pour A inclus dans X(Ω), ou plus généralement, pour un intervalle A de $\mathbb{R}$. On réalise des expériences (tirer un dé, ou deux dés, ...) dont le résultat ω est dans Ω. On applique X à ω, et on regarde si X(ω) appartient à A ou non. Le premier cas correspond à l'événement recherché. La probabilité $\mathbf{P}_X(A)$ de cet événement est donc égal à la probabilité de l'ensemble des ω tels que $X(\omega)$ appartienne à A. Or $\{\omega \mid X(\omega) \in A\}$ n'est autre que l'image réciproque de A par X, $X^{-1}(A)$.

DEFINITION

Soit X *une variable aléatoire définie sur un espace probabilisé* $(\Omega, \mathbf{P})$ *et* A *une partie de* $X(\Omega)$. *On définit la* **loi** *de* X *par* :

$$\mathbf{P}_X(A) = \mathbf{P}(X^{-1}(A)) = \mathbf{P}(\{\omega \in \Omega \mid X(\omega) \in A\})$$

PROPOSITION

La loi $\mathbf{P}_X$ *d'une variable aléatoire définie sur un espace probabilisé* $(\Omega, \mathbf{P})$ *est une loi de probabilité sur* $X(\Omega)$.

<u>Démonstration</u> :

❑ Il faut vérifier les trois propriétés des probabilités.

i) Il est clair que $\mathbf{P}_X$ est à valeurs dans [0, 1], puisque c'est le cas de P.

ii) $\mathbf{P}_X(X(\Omega)) = \mathbf{P}(X^{-1}(X(\Omega))) = \mathbf{P}(\Omega) = 1$

iii) Soit A et B deux événements disjoints de $X(\Omega)$. Alors $\mathbf{P}_X(A \cup B) = \mathbf{P}(X^{-1}(A \cup B))$. Or :

$$\begin{aligned} X^{-1}(A \cup B) &= \{\omega \mid X(\omega) \in A \cup B\} \\ &= \{\omega \mid X(\omega) \in A\} \cup \{\omega \mid X(\omega) \in B\} \\ &= X^{-1}(A) \cup X^{-1}(B) \end{aligned}$$

donc :

$$\mathbf{P}_X(A \cup B)) = \mathbf{P}(X^{-1}(A) \cup X^{-1}(B))$$

Par ailleurs, A et B étant disjoints, il en est de même de $X^{-1}(A)$ et $X^{-1}(B)$ puisqu'un élément ω commun à ces deux parties serait tel que $X(\omega) \in A \cap B$ qui est vide. Donc :

$$\mathbf{P}_X(A \cup B)) = \mathbf{P}(X^{-1}(A)) + \mathbf{P}(X^{-1}(B)) = \mathbf{P}_X(A) + \mathbf{P}_X(B)$$

On abrège $\mathbf{P}_X(A) = \mathbf{P}(X^{-1}(A)) = \mathbf{P}(\{\omega \mid X(\omega) \in A\})$ en $\mathbf{P}(X \in A)$. De même, si x est un élément de E, $\mathbf{P}(X = x)$ désigne $\mathbf{P}(\{\omega \in \Omega \mid X(\omega) = x) = \mathbf{P}(X^{-1}(\{x\})) = \mathbf{P}_X(\{x\})$. Comme toute probabilité sur un espace fini, les valeurs $\mathbf{P}(X = x) = \mathbf{P}_X(\{x\})$ définissent sans ambiguïté la probabilité $\mathbf{P}_X$ sur $X(\Omega)$. Enfin, dans le cas d'une variable aléatoire réelle, on note, pour tout x réel :

$$\mathbf{P}(X \leq x) = \mathbf{P}(\{\omega \in \Omega \mid X(\omega) \leq x) = \mathbf{P}(X^{-1}(]-\infty, x])) = \mathbf{P}_X(]-\infty, x] \cap X(\Omega))$$

La variable aléatoire X permet de transférer la probabilité $\mathbf{P}$ sur Ω en une probabilité $\mathbf{P}_X$ sur $X(\Omega)$. Considérons maintenant une variable aléatoire X de Ω dans un ensemble E, et une application f de E dans un ensemble F.

❑ X permet de transférer la probabilité $\mathbf{P}$ sur Ω en une probabilité $\mathbf{P}_X$ sur $X(\Omega)$. $(X(\Omega), \mathbf{P}_X)$ devient alors un espace probabilisé et f une variable aléatoire sur $X(\Omega)$ à valeurs dans F. f permet donc de transférer la probabilité $\mathbf{P}_X$ en une probabilité sur $f(X(\Omega))$.

❑ $Y = f \circ X$ est une variable aléatoire définie sur Ω à valeurs dans F, qui transfère la probabilité $\mathbf{P}$ de Ω en une probabilité $\mathbf{P}_Y$ sur $Y(\Omega) = f(X(\Omega))$

Vérifions que les deux procédés donnent le même résultat. Pour toute partie A de $Y(\Omega)$, on a :

$$\mathbf{P}_Y(A) = \mathbf{P}(Y^{-1}(A))$$

Or $Y^{-1}(A) = X^{-1}(f^{-1}(A))$ car :

$$\omega \in Y^{-1}(A) \Leftrightarrow Y(\omega) \in A \Leftrightarrow f(X(\omega)) \in A \Leftrightarrow X(\omega) \in f^{-1}(A) \Leftrightarrow \omega \in X^{-1}(f^{-1}(A))$$

Donc :

$\mathbf{P}_Y(A) = \mathbf{P}(X^{-1}(f^{-1}(A))) = \mathbf{P}_X(f^{-1}(A))$

et cette dernière est bien la probabilité définie sur $f(X(\Omega))$ par le premier procédé. On note habituellement $Y = f(X)$ au lieu de $f \circ X$.

2- Lois usuelles discrètes

a) Variable certaine

Soit $(\Omega, \mathbf{P})$ un espace probabilisé et k un réel donné. On appelle **variable certaine** une variable aléatoire de la forme :

$$X : \Omega \to \mathbb{R}$$
$$\omega \to k$$

Sa loi est concentrée sur la valeur k. $\mathbf{P}(X = k) = 1$.

EXEMPLE :

❑ On lance une pièce. Si on tombe sur P, on gagne un franc (G = l), sinon on perd un franc (G = – l). La variable $X = G^2$ est une variable certaine, égale à l.

b) Loi de Bernoulli

La **loi de Bernoulli** est la loi notée $\mathcal{B}(p)$ d'une variable aléatoire qui ne prend que deux valeurs, habituellement 0 ou 1. Par exemple, si on lance une pièce, le fait de tomber sur Pile attribue à X la valeur 0, et sur Face, la valeur 1. On a :

$$\mathbf{P}(X = 1) = p$$
$$\mathbf{P}(X = 0) = 1 - p$$

Généralement, la valeur 1 de X est appelée **succès**, et la valeur 0 **échec**. p est élément de $]0, 1[$. Si $p = 0$ ou 1, on retrouve une loi certaine.

c) Loi binomiale

La **loi binomiale** est la loi des variables aléatoires suivantes :

❑ On lance n pièces. On compte le nombre de Piles.

❑ On tire avec remise n boules d'une urne contenant des boules blanches et des boules noires. On compte le nombre de boules blanches tirées.

❑ Schéma général : on répète n expériences de Bernoulli indépendantes, et on compte le nombre X de succès. Si X_i est le résultat 0 ou 1 de la i-ème expérience, $1 \leq i \leq n$, alors X est la somme des X_i.

On a :

$$\mathbf{P}_X(\{k\}) = \mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$\binom{n}{k}$ représente en effet le nombre de combinaisons de k éléments parmi n, correspondant aux k succès. La probabilité de chacune de ces combinaisons est $p^k (1 - p)^{n-k}$.

Cette loi dépend des deux paramètres n et p et est notée $\mathcal{B}(n, p)$. La loi $\mathcal{B}(1, p)$ est identique à la loi de Bernoulli $\mathcal{B}(p)$.

d) Loi hypergéométrique

La loi binomiale, intervenant lors des tirages avec remise, peut également parfois servir d'approximation lors de tirages sans remise. Considérons le cas de tirages de n boules sans remise dans une urne contenant N boules, dont m blanches. La loi de la variable aléatoire X égale au nombre de boules blanches tirées se trouve en calculant le quotient du nombre de tirages favorables sur le nombre de tirages possibles. k est un entier pouvant varier de Max(0, $m + n - \mathrm{N}$) à Min(m, n).

$$\mathbf{P}_{\mathrm{X}}(\{k\}) = \mathbf{P}(\mathrm{X} = k) = \frac{\binom{m}{k}\binom{\mathrm{N}-m}{n-k}}{\binom{\mathrm{N}}{n}} = \frac{m!\ n!\ (\mathrm{N}-m)!\ (\mathrm{N}-n)!}{k!\ (m-k)!\ (n-k)!\ (\mathrm{N}-m-n+k)!\ \mathrm{N}!}$$

Curieusement, cette expression est symétrique en n et m. Cette symétrie s'explique aisément si l'on imagine l'expérience suivante : une urne contient N boules. Deux personnes sont en présence. La première personne en choisit m sur lesquelles elle dessine une marque personnelle. Puis elle les remet dans l'urne. La deuxième personne en choisit alors n sur lesquelles elle dessine une marque différente de la première personne. On examine alors les N boules. Quelle est la probabilité qu'il y en ait k disposant des deux marques ? Il s'agit de la loi précédente (dite **loi hypergéométrique**). Il est évident que l'expérience donne la même loi si on permute les deux personnes, c'est-à-dire si on échange le rôle de n et m.

Remarquons par ailleurs que la formule $\sum_k \mathbf{P}(\mathrm{X} = k) = 1$ donne la relation non triviale suivante :

$$\sum_k \binom{m}{k}\binom{\mathrm{N}-m}{n-k} = \binom{\mathrm{N}}{n}$$

qu'on pourra justifier directement en remarquant que le membre de gauche peut aussi exprimer le nombre de façons de choisir une partie de n éléments dans un ensemble à N éléments. Pour une valeur quelconque de k, on en prend k parmi m boules particulières et $n - k$ parmi celles qui restent.

Cette loi intervient également dans les sondages de n personnes différentes, dans une population totale de N personnes, dont m possèdent un caractère donné. On dénombre le nombre k de personnes dans l'échantillon choisi ayant le dit caractère.

Lorsque le nombre N de boules est de plus en plus grand, on peut s'attendre à ce qu'un tirage sans remise donne à peu près le même résultat qu'un tirage avec remise car il est de moins en moins probable de tirer deux fois la même boule. La loi hypergéométrique doit donc permettre de retrouver une loi binomiale. Cherchons donc la limite de $\mathbf{P}(\mathrm{X} = k)$ lorsque N tend vers l'infini et $\dfrac{m}{\mathrm{N}}$ (proportion de boules blanches) tend vers p, élément de [0,1]. On a :

$$\begin{aligned}\mathbf{P}(\mathrm{X} = k) &= \frac{m!\ n!\ (\mathrm{N}-m)!\ (\mathrm{N}-n)!}{k!\ (m-k)!\ (n-k)!\ (\mathrm{N}-m-n+k)!\ \mathrm{N}!} \\ &= \binom{n}{k}\frac{m!\ (\mathrm{N}-m)!\ (\mathrm{N}-n)!}{(m-k)!\ (\mathrm{N}-m-n+k)!\ \mathrm{N}!} \\ &= \binom{n}{k}\frac{(\mathrm{N}-m)(\mathrm{N}-m-1)...(\mathrm{N}-m-n+k+1)}{\mathrm{N}(\mathrm{N}-1)(\mathrm{N}-2)...(\mathrm{N}-n+k+1)}\ \frac{m(m-1)...(m-k+1)}{(\mathrm{N}-n+k)...\ (\mathrm{N}-n+1)}\end{aligned}$$

Le terme général du premier quotient vaut, pour r variant de 0 à $n - k - 1$:

$$\frac{\mathrm{N}-m-r}{\mathrm{N}-r} = \frac{1 - m/\mathrm{N} - r/\mathrm{N}}{1 - r/\mathrm{N}}$$

or r est borné donc $\dfrac{r}{\mathrm{N}}$ tend vers 0, et $\dfrac{m}{\mathrm{N}}$ est supposé tendre vers p. La limite de chacun de ces termes est donc $1 - p$, et ils sont au nombre de $n - k$. On prouve de même que chacun des termes du second

quotient tend vers p, et ils sont au nombre de k. La limite de $\mathbf{P}(X = k)$ est donc $\binom{n}{k} p^k (1 - p)^{n-k}$, terme général de la loi binomiale de paramètres n et p.
Cette propriété explique que, dans le cas d'un sondage où n est, en France, de l'ordre de 1000, alors que N est de l'ordre de 60 000 000, bien que ne procédant pas a priori à des tirages avec remise, la loi en présence est considérée comme une loi binomiale.

EXEMPLE :
❑ 46% des gens sont de groupe sanguin O. Sur 10 000 personnes, on en prend 20. Quelle est la probabilité d'en trouver 9 de groupe O ? Si on considère qu'on est dans le cadre d'une loi binomiale $\mathcal{B}(20, \frac{46}{100})$, alors :

$$\mathbf{P}(X = 9) = \binom{20}{9} \times 0.46^9 \times 0.54^{11} \approx 0{,}17634$$

alors que la valeur donnée par la loi hypergéométrique est 0,17651.

3- Espérance et variance

a) Espérance

On considère une loi de probabilité d'une variable aléatoire réelle X à valeur dans $\{x_i, i \in I\}$ où I est fini. Soit p_i la probabilité $\mathbf{P}_X(\{x_i\}) = \mathbf{P}(X = x_i)$. On s'intéresse à la valeur moyenne de X. On peut s'attendre à ce que, sur un grand nombre d'expérience, X prenne la valeur x_i avec une fréquence de plus en plus proche de p_i. Il est donc naturel de poser, comme valeur moyenne de X, la quantité :

$$\mathbf{E}(X) = \sum_{i \in I} p_i\, x_i = \sum_{i \in I} x_i\, \mathbf{P}(X = x_i)$$

Cette quantité s'appelle **espérance** de X.

EXEMPLE :
❑ Soit X la somme de deux dés. La loi de X est la suivante :

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}(X = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Le calcul conduit à $\mathbf{E}(X) = 7$, ce qui est naturel, vu la symétrie du tableau.

On remarquera que la définition de $\mathbf{E}(X)$ ne fait intervenir que les valeurs que prend X et sa loi $\mathbf{P}_X$ sur $\mathbb{R}$, et que l'on peut oublier totalement l'espace Ω sur lequel est défini X. Nous pouvons cependant donner une expression de $\mathbf{E}(X)$ faisant intervenir Ω et la loi de probabilité $\mathbf{P}$ sur Ω.

$$\mathbf{E}(X) = \sum_{\omega \in \Omega} X(\omega)\, \mathbf{P}(\{\omega\})$$

En effet, soit $x_1, x_2, ..., x_n$ les valeurs que prend X. Alors les images réciproques $X^{-1}(\{x_1\})$, $X^{-1}(\{x_2\})$, ..., $X^{-1}(\{x_n\})$ forment un système complet d'événements de Ω : ils sont disjoints (car un élément ω ne peut avoir deux images différentes) et leur réunion est Ω (car tout ω admet comme image l'un des x_i). La somme sur Ω peut donc se calculer en sommant d'abord sur $X^{-1}(\{x_1\})$, puis sur $X^{-1}(\{x_2\})$, ..., puis sur $X^{-1}(\{x_n\})$, et en faisant la somme de ces n totaux :

$$\sum_{\omega\in\Omega} X(\omega)\, \mathbf{P}(\{\omega\}) = \sum_{i=1}^{n} \sum_{\omega\in X^{-1}(\{x_i\})} X(\omega)\, \mathbf{P}(\{\omega\})$$

D'autre part, lorsque ω appartient à $X^{-1}(\{x_i\})$, par définition, $X(\omega)$ est égal à x_i. D'où :

$$\sum_{\omega\in\Omega} X(\omega)\, \mathbf{P}(\{\omega\}) = \sum_{i=1}^{n} \sum_{\omega\in X^{-1}(\{x_i\})} x_i\ \mathbf{P}(\{\omega\}) = \sum_{i=1}^{n} x_i \sum_{\omega\in X^{-1}(\{x_i\})} \mathbf{P}(\{\omega\})$$

$$= \sum_{i=1}^{n} x_i\, \mathbf{P}(X^{-1}(\{x_i\}) = \sum_{i=1}^{n} x_i\, \mathbf{P}(X = x_i) = \mathbf{E}(X)$$

Cette formule est utile pour prouver le résultat suivant :

TEHOREME DE TRANSFERT

Soit $X : \Omega \to \{x_i,\ i \in I\} = X(\Omega)$ *une variable aléatoire définie sur l'espace probabilisé* $(\Omega, \mathbf{P})$ *et* $f : X(\Omega) \to \mathbb{R}$ *une fonction. Alors* $f \circ X$ (*notée également* $f(X)$) *est une variable aléatoire dont l'espérance se calcule de la façon suivante, au moyen de la loi de* X :

$$\mathbf{E}(f(X)) = \sum_{i\in I} f(x_i)\, \mathbf{P}(X = x_i)$$

Autrement dit, l'espérance de la variable aléatoire $f(X)$ définie sur l'espace probabilisé $(\Omega, \mathbf{P})$ est égale à l'espérance de f, vue comme variable aléatoire définie sur l'espace probabilisé $(X(\Omega), \mathbf{P}_X)$ muni de la loi de X qui vérifie $\mathbf{P}_X(\{x_i\}) = \mathbf{P}(X = x_i)$.

<u>Démonstration</u> :

❑ Elle est analogue à la démarche suivie précédemment :

$$\mathbf{E}(f(X)) = \sum_{\omega\in\Omega} f(X(\omega))\, \mathbf{P}(\{\omega\}) = \sum_{i=1}^{n} \sum_{\omega\in X^{-1}(\{x_i\})} f(x_i)\, \mathbf{P}(\{\omega\}) = \sum_{i=1}^{n} f(x_i) \sum_{\omega\in X^{-1}(\{x_i\})} \mathbf{P}(\{\omega\})$$

$$= \sum_{i=1}^{n} f(x_i)\, \mathbf{P}(X^{-1}(\{x_i\}) = \sum_{i=1}^{n} f(x_i)\, \mathbf{P}(X = x_i)$$

L'intérêt de cette formule est qu'il est inutile de rechercher la loi de $f(X)$ pour calculer son espérance. La loi de X suffit.

EXEMPLE :

❑ Soit X la somme de deux dés. La loi de X est la suivante :

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}(X = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Cherchons l'espérance de $(X - 7)^2$. La formule démontrée permet d'écrire :

$$\mathbf{E}((X - 7)^2) = 5^2 \times \frac{1}{36} + 4^2 \times \frac{2}{36} + 3^2 \times \frac{3}{36} + 1^2 \times \frac{4}{36} + \ldots + 5^2 \times \frac{1}{36} = \frac{35}{6}$$

Cette quantité mesure la dispersion de X autour de sa valeur moyenne et s'appelle variance de X.

PROPOSITION

L'espérance possède les propriétés suivantes :

i) $\forall\, a \in \mathbb{R}$, $\forall\, b \in \mathbb{R}$, $\forall$ X *et* Y *variables aléatoires sur* $(\Omega, \mathbf{P})$, *on a* :

$$\mathbf{E}(aX + bY) = a\mathbf{E}(X) + b\mathbf{E}(Y).$$

Autrement dit, l'espérance est un opérateur linéaire sur l'espace vectoriel des variables aléatoires.

ii) $X \geq 0 \Rightarrow \mathbf{E}(X) \geq 0$

ii bis) $X \geq Y \Rightarrow \mathbf{E}(X) \geq \mathbf{E}(Y)$

iii) $|\mathbf{E}(X)| \leq \mathbf{E}(|X|)$

iv) $|\mathbf{E}(XY)| \leq \sqrt{\mathbf{E}(X^2)}\,\sqrt{\mathbf{E}(Y^2)}$

v) *Si* X *est une variable certaine égale à* k, *alors* $\mathbf{E}(X) = k$

<u>Démonstration</u> :

❑ i) et ii) se déduisent facilement du fait que $\mathbf{E}(X) = \sum_{\omega\in\Omega} X(\omega)\ \mathbf{P}(\{\omega\})$, expression linéaire par rapport à X, et positive si X est positive.

❑ ii bis) se déduit de ii) en considérant X – Y:

$$X - Y \geq \Rightarrow \mathbf{E}(X - Y) \geq 0 \Rightarrow \mathbf{E}(X) - \mathbf{E}(Y) \geq 0 \Rightarrow \mathbf{E}(X) \geq \mathbf{E}(Y).$$

❑ iii) se montre en remarquant que $-X \leq |X| \leq X$, donc $-\mathbf{E}(X) \leq \mathbf{E}(|X|) \leq \mathbf{E}(X)$

❑ iv) La quantité $\langle X,Y\rangle = \mathbf{E}(XY)$ est une forme bilinéaire symétrique positive, et comme toute forme bilinéaire symétrique positive, elle vérifie l'inégalité de Schwarz :

$$\langle X,Y\rangle^2 \leq \langle X,X\rangle\ \langle Y,Y\rangle$$

(Voir L1/ESPEUCL.PDF)

❑ v) On a $\mathbf{P}(X = k) = 1$, donc $\mathbf{E}(X) = k\ \mathbf{P}(X = k) = k$.

De v), on déduit que toute variable aléatoire $X - \mathbf{E}(X)$ a une espérance nulle. Une telle variable, d'espérance nulle, est dite **centrée**.

b) Variance

On mesure la dispersion de la variable aléatoire X autour de son espérance $\mathbf{E}(X) = \mu$ en calculant la quantité $\mathbf{E}((X - \mu)^2)$. Il s'agit de la **variance** de X, notée $\mathbf{V}(X)$. Si l'on développe cette quantité, on obtient :

$$\begin{aligned}\mathbf{V}(X) &= \mathbf{E}((X - \mu)^2)\\ &= \mathbf{E}(X^2 - 2X\mu + \mu^2)\\ &= \mathbf{E}(X^2) - 2\mu\mathbf{E}(X) + \mathbf{E}(\mu^2)\\ &= \mathbf{E}(X^2) - 2\mu^2 + \mu^2 \qquad \text{car } \mathbf{E}(X) = \mu \text{ et } \mathbf{E}(\mu^2) = \mu^2.\end{aligned}$$

Donc $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

C'est cette dernière quantité que l'on calcule le plus souvent pour obtenir la variance.

On vérifiera facilement que :

$\forall\, a \in \mathbb{R}, \forall\, b \in \mathbb{R}, \mathbf{V}(aX + b) = a^2\mathbf{V}(X)$.

On appelle **écart-type** la quantité $\sigma(X) = \sqrt{V(X)}$.

La variable aléatoire $\dfrac{X - \mathbf{E}(X)}{\sigma(X)}$ a une espérance nulle, et sa variance vaut 1. On dit qu'elle est **centrée réduite**.

c) Lois usuelles

❑ *Variable certaine* :

Soit X = k une variable certaine égale à k. Le lecteur vérifiera sans difficulté que $\mathbf{E}(X) = k$, $\mathbf{E}(X^2) = k^2$, et $\mathbf{V}(X) = 0$.

Réciproquement, il sera prouvé dans la paragraphe suivant que, si X admet une variance nulle, alors, il existe k tel que $\mathbf{P}(X = k) = 1$.

❑ *Variable de Bernoulli* :

Elle admet la loi suivante :

X	0	1
$\mathbf{P}_X$	$1 - p$	p

On a donc : $\mathbf{E}(X) = p = \mathbf{E}(X^2)$ (en effet, $X^2 = X$). D'où $\mathbf{V}(X) = p - p^2 = p(1 - p) = pq$ avec $q = 1 - p$

Un exemple fréquent de variable de Bernoulli est celui des fonctions indicatrices. Soit $(\Omega, \mathbf{P})$ un espace probabilisé, et A un événement. On considère la variable aléatoire définie de la façon suivante :

$$\mathbf{1}_A : \Omega \to \{0, 1\}$$
$$\omega \to \begin{cases} 0 \text{ si } \omega \notin A \\ 1 \text{ si } \omega \in A \end{cases}$$

Cette variable aléatoire est la fonction indicatrice de la partie A. La loi de $\mathbf{1}_A$ est une loi de Bernoulli de paramètre p, avec :

$$\begin{aligned} p &= \mathbf{P}(\mathbf{1}_A = 1) \\ &= \mathbf{P}(\{\omega \in \Omega, \mathbf{1}_A(\omega) = 1\}) \\ &= \mathbf{P}(A) \end{aligned}$$

Alors, $\mathbf{E}(\mathbf{1}_A) = \mathbf{P}(A)$

Voici quelques règles de calcul des fonctions indicatrices, dont la vérification est laissée au lecteur :

$\mathbf{1}_\Omega = 1$

$\mathbf{1}_\varnothing = 0$

$\mathbf{1}_{A\cap B} = \mathbf{1}_A \times \mathbf{1}_B$

$\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$

$\mathbf{1}_{A\setminus B} = \mathbf{1}_A - \mathbf{1}_A \times \mathbf{1}_B$

$1 - \mathbf{1}_{A\cup B} = (1 - \mathbf{1}_A)(1 - \mathbf{1}_B)$

$\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A\cap B} = \mathrm{Max}(\mathbf{1}_A, \mathbf{1}_B)$

L'avant-dernière formule se généralise aisément à une réunion quelconque de n parties $A_1, ..., A_n$, de la façon suivante :

$$1 - \mathbf{1}_{\cup A_i} = (1 - \mathbf{1}_{A_1})(1 - \mathbf{1}_{A_2})...(1 - \mathbf{1}_{A_n})$$

En effet :

$$1 - \mathbf{1}_{\cup A_i} = \mathbf{1}_{(\cup Ai)^c} = \mathbf{1}_{\cap A_i^c} = \prod_{i=1}^{n} \mathbf{1}_{A_i^c} = \prod_{i=1}^{n} (1 - \mathbf{1}_{A_i})$$

Cela permet de démontrer la formule dite **de Poincaré** en développant le membre de droite, qui donne la probabilité d'une union quelconque d'événements :

$$1 - \mathbf{1}_{\cup A_i} = 1 - \sum_i \mathbf{1}_{A_i} + \sum_{i,j} \mathbf{1}_{A_i \cap A_j} - ... + (-1)^m \sum_{i_1<i_2<...<i_m} \mathbf{1}_{A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_m}} + ... + (-1)^n \mathbf{1}_{A_1 \cap A_2 \cap ... \cap A_n}$$

En prenant l'espérance des deux membres, on obtient :

$$\mathbf{P}(\bigcup_{i=1}^{n} A_i) = \sum_i \mathbf{P}(A_i) - \sum_{i<j} \mathbf{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbf{P}(A_i \cap A_j \cap A_n) - ...$$

$$+ (-1)^{m-1} \sum_{i_1<i_2<...<i_m} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_m}) + ... + (-1)^{n-1} \mathbf{P}(A_1 \cap A_2 \cap ... \cap A_n)$$

Si P est donné par $\mathbf{P}(A) = \dfrac{\text{Card}(A)}{\text{Card}(\Omega)}$, on obtient :

$$\text{Card}(\bigcup_{i=1}^{n} A_i) = \sum_i \text{Card}(A_i) - \sum_{i<j} \text{Card}(A_i \cap A_j) + \sum_{i<j<k} \text{Card}(A_i \cap A_j \cap A_n) - ...$$

$$+ (-1)^{m-1} \sum_{i_1<i_2<...<i_m} \text{Card}(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_m}) + ...$$

$$+ (-1)^{n-1} \text{Card}(A_1 \cap A_2 \cap ... \cap A_n)$$

qui généralise la formule $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$.

❑ *Loi binomiale* :

Soit $X_1 , ..., X_n$ n variables aléatoires de Bernoulli indépendantes (i.e. les événements correspondant ces variables sont indépendants), à valeur dans {0, 1}, de même paramètre p. Alors la somme $X = \sum_{i=1}^{n} X_i$ suit une loi binomiale $\mathcal{B}(n, p)$. X est égal au nombre de succès des n expériences.

Rappelons que :

$$\mathbf{E}(X_i) = p$$
$$\mathbf{V}(X_i) = p(1 - p)$$

On en déduit que :

$$\mathbf{E}(X) = \sum_{i=1}^{n} \mathbf{E}(X_i) = np$$

Par ailleurs, il est démontré plus bas au IV que la variance d'une somme de variables aléatoires indépendantes est égale à la somme des variances. D'où :

$$\mathbf{V}(X) = \sum_{i=1}^{n} \mathbf{V}(X_i) = np(1-p)$$

On peut également démontrer ce dernier résultat directement. Calculons $\mathbf{E}(X^2)$:

$$X^2 = \sum_{i=1}^{n} X_i^2 + 2 \sum_{i<j} X_i X_j$$

X_i^2 n'est autre que X_i, donc son espérance est p.
X_iX_j ne prend que les valeurs 0 ou 1. C'est une variable de Bernoulli, de paramètre p^2. En effet :

$$\begin{aligned} \mathbf{P}(X_iX_j = 1) &= \mathbf{P}(X_i = 1 \text{ et } X_j = 1) \\ &= \mathbf{P}(X_i = 1) \times \mathbf{P}(X_j = 1) \quad \text{(indépendance des variables)} \\ &= p^2 \end{aligned}$$

Donc $\mathbf{E}(X_iX_j) = p^2$. Enfin, il y a $\dfrac{n(n-1)}{2}$ couples (i, j) tels que i soit strictement inférieur à j. Ainsi :

$$\mathbf{E}(X^2) = np + n(n-1)p^2$$

$$\Rightarrow \quad \mathbf{V}(X) = np + n(n-1)p^2 - (np)^2 = np(1-p)$$

❑ *Loi hypergéométrique* :
On reprend le modèle de l'urne de N boules dont m blanches. La variable X donnant le nombre de boules blanches lors d'un tirage sans remise de n boules suit une loi hypergéométrique de paramètres N, m, n. Notons $p = \dfrac{m}{N}$ la proportion initiale de boules blanches. Lors du premier tirage, la probabilité de tirer une boule blanche est p. Curieusement, cette probabilité est la même à chaque tirage, même si celui-ci se fait sans remise. Soit en effet $i \geq 1$. Supposons que l'on ait tiré i boules, $i < N$, et que parmi celles-ci, il y ait eu k blanches (événement A_k). Quelle est la probabilité de tirer une boule blanche au $(i + 1)$-ème tirage (événement B) ? On a la probabilité conditionnelle suivante :

$$\mathbf{P}(B \mid A_k) = \frac{m-k}{N-i} \qquad \text{puisqu'il reste } m-k \text{ blanches parmi } N-i \text{ boules}$$

$$\mathbf{P}(A_k) = \frac{\binom{m}{k}\binom{N-m}{i-k}}{\binom{N}{i}} \qquad \text{nombre de tirages favorables sur nombre de tirages possibles}$$

La formule des probabilités totales donne le résultat annoncé :

$$\begin{aligned} \mathbf{P}(B) = \sum_k \mathbf{P}(B \mid A_k) \times \mathbf{P}(A_k) &= \sum_k \frac{\binom{m}{k}\binom{N-m}{i-k}}{\binom{N}{i}} \frac{m-k}{N-i} \\ &= \frac{m}{N} \sum_k \frac{\binom{m-1}{k}\binom{N-m}{i-k}}{\binom{N-1}{i}} = \frac{m}{N} = p \end{aligned}$$

car $\displaystyle\sum_k \frac{\binom{m-1}{k}\binom{N-m}{i-k}}{\binom{N-1}{i}} = 1$ (somme d'une loi de probabilité hypergéométrique de paramètres $N-1$, $m-1$, i).

Par conséquent, si, comme pour la loi binomiale, on note X_i la variable aléatoire valant 1 si le i-ème tirage est une boule blanche et 0 sinon, alors :

$$\mathbf{E}(X_i) = p$$

$$X = \sum_{i=1}^{n} X_i$$

donc $\quad \mathbf{E}(X) = np = \dfrac{nm}{N}$

L'espérance est identique à celle de la loi binomiale.

Par contre la variance de la loi hypergéométrique est différente de celle de la loi binomiale car, pour $i < j$, les événements $X_i = 1$ et $X_j = 1$ ne sont pas indépendants. On a :

$$\mathbf{P}(X_i = 1, X_j = 1) = \mathbf{P}(X_j = 1 \mid X_i = 1) \times \mathbf{P}(X_i = 1)$$

On a vu que $\mathbf{P}(X_i = 1) = p = \dfrac{m}{N}$. Quant à $\mathbf{P}(X_j = 1 \mid X_i = 1)$, elle vaut $\dfrac{m-1}{N-1}$ car, si on suppose qu'une blanche a été tirée au i-ème tirage, on peut supprimer ce tirage et supposer que les autres tirages sont effectués sans remise dans une urne de $N - 1$ boules contenant $m - 1$ blanches. La probabilité de tirer une boule blanche lors d'un tirage j donné est dans ce cas $\dfrac{m-1}{N-1}$. On a alors :

$$\mathbf{E}(X_iX_j) = \mathbf{P}(X_iX_j = 1) = \mathbf{P}(X_i = 1, X_j = 1) = \frac{m}{N}\frac{m-1}{N-1}$$

$$X^2 = \sum_{i=1}^{n} X_i^2 + 2\sum_{i<j} X_iX_j = \sum_{i=1}^{n} X_i + 2\sum_{i<j} X_iX_j \qquad \text{car } X_i^2 = X_i$$

$$\mathbf{E}(X^2) = \frac{nm}{N} + 2\binom{n}{2}\frac{m}{N}\frac{m-1}{N-1}$$

$$\mathbf{V}(X) = \frac{nm}{N} + 2\binom{n}{2}\frac{m}{N}\frac{m-1}{N-1} - \left(\frac{nm}{N}\right)^2 = \frac{mn(N-n)(N-m)}{N^2(N-1)}$$

$$= np(1-p)\frac{N-n}{N-1}$$

La variance de la loi hypergéométrique est légèrement inférieure à la variance $np(1 - p)$ de la loi binomiale.

❑ *Loi uniforme sur* $[\![1, n]\!]$:

La probabilité que $X = i$ vaut $\dfrac{1}{n}$. Donc :

$$\mathbf{E}(X) = \sum_{i=1}^{n} \frac{i}{n} = \frac{n+1}{2}$$

$$\mathbf{E}(X^2) = \sum_{i=1}^{n} \frac{i^2}{n} = \frac{(n+1)(2n+1)}{6}$$

$$\mathbf{V}(X) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

On a utilisé la formule $\displaystyle\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$, qui peut se montrer par récurrence.

EXEMPLE :

❑ On considère le groupe symétrique $\mathfrak{S}_n$ des permutations de l'ensemble $[\![1, n]\!]$. On munit $\mathfrak{S}_n$ d'une loi uniforme. On tire une permutation σ au hasard et on note X(σ) le nombre de points fixes de σ, i.e. le nombre d'éléments i de $[\![1, n]\!]$ tels que $\sigma(i) = i$. Déterminons l'espérance de X. Pour tout i élément de $[\![1, n]\!]$, posons $X_i(\sigma) = \begin{cases} 1 \text{ si } \sigma(i) = i \\ 0 \text{ sinon} \end{cases}$. On a alors $X = \sum_{i=1}^{n} X_i$. Il suffit de déterminer l'espérance des X_i. Or $\mathbf{P}(X_i) = \dfrac{(n-1)!}{n!} = \dfrac{1}{n}$ car il y a $(n-1)!$ permutations σ satisfaisant $\sigma(i) = i$, sur $n!$ permutations au total. Donc $\mathbf{E}(X_i) = \dfrac{1}{n}$ et $\mathbf{E}(X) = \sum_{i=1}^{n} \mathbf{E}(X_i) = 1$. En moyenne, le nombre de points fixes d'une permutation est égal à 1.

Cherchons maintenant la variance de X. Pour $i \neq j$, on a :

$$\mathbf{E}(X_iX_j) = \mathbf{P}(X_i = 1, X_j = 1) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

Par ailleurs, pour tout i, $\mathbf{E}(X_i^2) = \mathbf{E}(X_i) = \dfrac{1}{n}$ car $X_i^2 = X_i$. Donc :

$$\begin{aligned} \mathbf{E}(X^2) &= \mathbf{E}\Big(\sum_{i=1}^{n} X_i^2 + \sum_{i\neq j} X_iX_j\Big) \\ &= \sum_{i=1}^{n} \mathbf{E}(X_i^2) + \sum_{i\neq j} \mathbf{E}(X_iX_j) \\ &= n \times \frac{1}{n} + n(n-1) \times \frac{1}{n(n-1)} \\ &= 2 \end{aligned}$$

donc $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = 2 - 1 = 1$.

La loi de X, non évidente, peut être trouvée en utilisant le nombre a_n de permutations sans point fixe de $[\![1, n]\!]$ (ou **dérangements**). Un exercice de L1/GROUPSYM.PDF établit la formule :

$$a_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

On en déduit que :

$$\mathbf{P}(X = r) = \frac{1}{n!} \binom{n}{r} a_{n-r}$$

choisir les r points fixes, puis une permutation
sans point fixe sur les $n - r$ éléments restants

$$= \frac{1}{r!} \sum_{k=0}^{n-r} \frac{(-1)^k}{k!}$$

Si on calcule l'espérance de X en utilisant cette loi, on obtient :

$$\mathbf{E}(X) = \sum_{r=1}^{n} \sum_{k=0}^{n-r} \frac{(-1)^k}{k!(r-1)!}$$

$$= \sum_{r=0}^{n-1} \sum_{k=0}^{n-1-r} \frac{(-1)^k}{k!r!} \qquad \text{en changeant d'indice } r \leftarrow r-1$$

$$= \sum_{k=0}^{n-1} \sum_{r=0}^{n-1-k} \frac{(-1)^k}{k!r!} \qquad \text{en permutant les sommes}$$

On en déduit la formule non évidente $\sum_{k=0}^{n-1} \sum_{r=0}^{n-1-k} \frac{(-1)^k}{k!r!} = 1$. Plusieurs autres démonstrations de cette formule sont données dans les exercices de L1/DENOMBRE.PDF.

4- Inégalité de Bienaymé-Tchebychev

PROPOSITION

Soit X *une variable aléatoire d'espérance* μ *et d'écart-type* σ, *et soit* ε *un réel strictement positif. Alors* : $\mathbf{P}(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$

<u>Démonstration</u> :

❑ Soit Ω l'espace probabilisé sur lequel s'applique X. Posons $Y = |X - \mu|$. On a $\sigma^2 = \mathbf{V}(X) = \mathbf{E}(Y^2)$. Notons $\mathbf{1}_{Y\geq\varepsilon}$ la fonction indicatrice de l'ensemble $\{\omega \in \Omega, Y(\omega) \geq \varepsilon\}$. Alors, on a :

$$Y^2 \geq \varepsilon^2\, \mathbf{1}_{Y\geq\varepsilon}$$

En effet, pour tout ω de Ω :

ou bien $Y(\omega) \geq \varepsilon$ et l'inégalité ci-dessus donne bien le même résultat, puisqu'alors $\mathbf{1}_{Y\geq\varepsilon}(\omega) = 1$

ou bien $Y(\omega) < \varepsilon$, est l'inégalité ci-dessus donne $Y(\omega) \geq 0$, puisque $\mathbf{1}_{Y\geq\varepsilon}(\omega) = 0$.

Dans tous les cas, l'inégalité est vérifiée. On en déduit que :

$$\sigma^2 = \mathbf{E}(Y^2) \geq \mathbf{E}(\varepsilon^2\, \mathbf{1}_{Y\geq\varepsilon}) = \varepsilon^2\, \mathbf{E}(\mathbf{1}_{Y\geq\varepsilon}) = \varepsilon^2\, \mathbf{P}(Y \geq \varepsilon)$$

D'où la proposition.

Cette inégalité permet d'évaluer la probabilité de s'écarter de sa moyenne. Son avantage certain est qu'elle ne dépend pas de la loi de X. Cette loi peut être ignorée. Son inconvénient est que la majoration est grossière. Pour une loi binomiale $\mathcal{B}(20, \frac{1}{2})$ et $\varepsilon = 3$, on a $\sigma^2 = 5$, $\mu = 10$ et l'inégalité de Bienaymé-Tchebychev donne :

$$\mathbf{P}(|X - 10| \geq 3) \leq \frac{5}{9} \approx 0{,}56$$

alors que le calcul donne $\mathbf{P}(|X - 10| \geq 3) \approx 0{,}26$

Notons que l'argument utilisé sur la variable aléatoire Y^2 se généralise à n'importe quelle variable aléatoire Z positive ou nulle sous la forme :

$$\forall\, a > 0,\ \mathbf{E}(Z) \geq a\, \mathbf{P}(Z \geq a) \qquad (\textbf{inégalité de Markov})$$

en remarquant que $Z \geq a\, \mathbf{1}_{Z\geq a}$ et en prenant l'espérance des deux membres.

APPLICATIONS :

❑ Si $X \geq 0$ et $\mathbf{E}(X) = 0$, alors $\mathbf{P}(X = 0) = 1$.

En effet, pour tout $a > 0$, l'inégalité de Markov donne $0 \geq a\, \mathbf{P}(X \geq a)$, donc $\mathbf{P}(X \geq a) = 0$. Donc en prenant pour a la plus petite valeur strictement positive susceptible d'être prise par X, on obtient $\mathbf{P}(X > 0) = 0$ et par passage au complémentaire, $\mathbf{P}(X = 0) = 1$.

❑ Si $\mathbf{V}(X) = 0$, alors il existe μ tel que $\mathbf{P}(X = \mu) = 1$.
En effet, l'inégalité de Bienaymé-Tchebychev donne $\mathbf{P}(|X - \mu| \geq \varepsilon) = 0$ pour tout $\varepsilon > 0$, ce qui permet de conclure que $\mathbf{P}(|X - \mu| > 0) = 0$ comme ci-dessus, et par passage au complémentaire, $\mathbf{P}(X = \mu) = 1$.

Dans le premier cas, on dit que X est **presque sûrement nulle** (ou presque partout nul), dans le second, que X est **presque sûrement constante**.

❑ Soit une suite de variables aléatoires indépendantes X_i, $1 \leq i \leq n$ de même loi, d'espérance μ et d'écart-type σ. Appliquons l'inégalité de Bienaymé-Tchebychev à $Z_n = \dfrac{X_1 + ... + X_n}{n}$. On a $\mathbf{E}(Z_n) = \mu$ et :

$$\mathbf{V}(Z_n) = \frac{\mathbf{V}(X_1 + ... + X_n)}{n^2}$$

$$= \frac{\mathbf{V}(X_1) + ... + \mathbf{V}(X_n)}{n^2} \qquad \text{voir IV, avec les } X_i \text{ variables aléatoires indépendantes}$$

$$= \frac{\sigma^2}{n}$$

Donc :

$$\mathbf{P}(|Z_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

quantité qui tend vers 0 quand n tend vers l'infini. Cela exprime que la probabilité que Z_n appartienne à l'intervalle $]\mu - \varepsilon, \mu + \varepsilon[$ tend vers 1 quand n tend vers l'infini, et ceci, quel que soit le nombre ε choisi. Cette propriété s'appelle la **loi faible des grands nombres**.

5- Espérance conditionnelle

Soit $(\Omega, \mathbf{P})$ un espace probabilisé et X une variable aléatoire réelle définie sur Ω. Soit B un événement de probabilité non nulle. On peut alors définir sur Ω une probabilité conditionnelle $\mathbf{P}_B$. On appelle **espérance conditionnelle** de X selon l'événement B l'espérance de X en utilisant la loi de probabilité $\mathbf{P}_B$. On la note $\mathbf{E}(X \mid B)$.

EXEMPLE :
❑ On considère un dé pipé ayant les probabilités de sortie suivantes.

1	2	3	4	5	6
$\frac{2}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{3}{15}$

Soit X la variable aléatoire égale au nombre sorti par le dé en un lancer. On a :

$$\mathbf{E}(X) = 1 \times \frac{2}{15} + 2 \times \frac{3}{15} + 3 \times \frac{2}{15} + 4 \times \frac{2}{15} + 5 \times \frac{3}{15} + 6 \times \frac{3}{15} = \frac{55}{15} = \frac{11}{3}$$

Soit maintenant B l'événement "tirer un nombre pair". On a :

$\mathbf{P}(B) = \frac{8}{15}$

La loi conditionnelle de X est :

$\mathbf{P}_B(X = 2) = \mathbf{P}_B(X = 6) = \frac{3}{8}$

$\mathbf{P}_B(X = 4) = \frac{2}{8}$

$\mathbf{P}_B(X \in \{1, 3, 5\}) = 0$

donc $\mathbf{E}(X \mid B) = 2 \times \frac{3}{8} + 4 \times \frac{2}{8} + 6 \times \frac{3}{8} = \frac{32}{8} = 4$

De même, si on conditionne par l'événement contraire B^c = "tirer un nombre impair", on a :

$\mathbf{P}(B^c) = \frac{7}{15}$

$\mathbf{P}_{B^c}(X = 1) = \mathbf{P}_{B^c}(X = 3) = \frac{2}{7}$

$\mathbf{P}_{B^c}(X = 5) = \frac{3}{7}$

donc $\mathbf{E}(X \mid B^c) = 1 \times \frac{2}{7} + 3 \times \frac{2}{7} + 5 \times \frac{3}{7} = \frac{23}{7}$

THEOREME DE L'ESPERANCE TOTALE

Soit $(A_i)_{1 \leq i \leq n}$ *un système complet d'événements d'un espace probabilisé fini* $(\Omega, \mathbf{P})$, *et* X *une variable aléatoire réelle définie sur* Ω. *Alors* :

$$\mathbf{E}(X) = \sum_{i=1}^{n} \mathbf{P}(A_i) \times \mathbf{E}(X \mid A_i)$$

<u>Démonstration</u> :

❑ On a :

$$\mathbf{E}(X) = \sum_{\omega \in \Omega} X(\omega)\, \mathbf{P}(\{\omega\}) = \sum_{i=1}^{n} \sum_{\omega \in A_i} X(\omega)\, \mathbf{P}(\{\omega\})$$

$$= \sum_{i=1}^{n} \sum_{\omega \in A_i} X(\omega)\, \mathbf{P}_{A_i}(\{\omega\})\, \mathbf{P}(A_i)$$

$$= \sum_{i=1}^{n} \mathbf{P}(A_i) \sum_{\omega \in A_i} X(\omega)\, \mathbf{P}_{A_i}(\{\omega\})$$

$$= \sum_{i=1}^{n} \mathbf{P}(A_i) \sum_{\omega \in \Omega} X(\omega)\, \mathbf{P}_{A_i}(\{\omega\}) \qquad \text{car si } \omega \notin A_i,\ \mathbf{P}_{A_i}(\{\omega\}) = 0$$

$$= \sum_{i=1}^{n} \mathbf{P}(A_i)\, \mathbf{E}(X \mid A_i)$$

EXEMPLES :

❑ Si on reprend l'exemple précédent du dé pipé, avec $A_1 = B$ = "tirer un nombre pair", $A_2 = B^c$ = "tirer un nombre impair", et X la variable aléatoire donnant le numéro tiré lors d'un lancer, on a :

$$\mathbf{P}(A_1) = \frac{8}{15}$$

$$\mathbf{P}(A_2) = \frac{7}{15}$$

$$\mathbf{E}(X \mid A_1) = 4$$

$$\mathbf{E}(X \mid A_2) = \frac{23}{7}$$

Donc :

$$\mathbf{P}(A_1) \times \mathbf{E}(X \mid A_1) + \mathbf{P}(A_2) \times \mathbf{E}(X \mid A_2) = \frac{8}{15} \times 4 + \frac{7}{15} \times \frac{23}{7} = \frac{32 + 23}{15} = \frac{55}{15} = \frac{11}{3} = \mathbf{E}(X)$$

❑ On considère une urne contenant n boules numérotées de 1 à n. On en tire sans remise un nombre Y, variable aléatoire. On calcule la somme X des boules tirées. Exprimons $\mathbf{E}(X)$ en fonction de $\mathbf{E}(Y)$. Posons $X_i = \left[\begin{array}{l} i \text{ si on a tiré la boule n°}i \\ 0 \text{ sinon} \end{array}\right.$. On a donc $X = \sum_{i=1}^{n} X_i$. Si on tire m boules, la probabilité qu'une boule donnée soit tirée est $\frac{m}{n}$. Par conséquent :

$$\mathbf{P}(X_i = i \mid Y = m) = \frac{m}{n}$$

donc $$\mathbf{E}(X_i \mid Y = m) = \frac{im}{n}$$

$$\mathbf{E}(X \mid Y = m) = \sum_{i=1}^{n} \mathbf{E}(X_i \mid Y = m) = \sum_{i=1}^{n} \frac{im}{n} = \frac{n+1}{2}\, m$$

donc $$\mathbf{E}(X) = \sum_{m=0}^{n} \mathbf{E}(X \mid Y = m)\, \mathbf{P}(Y = m) = \sum_{m=0}^{n} \frac{n+1}{2}\, m\mathbf{P}(Y = m) = \frac{n+1}{2}\, \mathbf{E}(Y)$$

On notera que, si Y est la variable aléatoire certaine m, alors $\mathbf{E}(X) = \frac{n+1}{2}\, m$, et que, dans le cas général, tout se passe comme si Y était la variable aléatoire certaine égale à son espérance.

IV : Couple de variables aléatoires

1- Définition

On considère un espace probabilisé $(\Omega, \mathbf{P})$ et deux variables aléatoires réelles X et Y définies sur Ω. (X, Y) est un couple de variables aléatoires. C'est l'application de Ω à valeurs dans $\mathbb{R}^2$, définie naturellement par :

$$\omega \in \Omega \to (X(\omega), Y(\omega)) \in \mathbb{R}^2$$

Si X est à valeurs dans $\{x_i, i \in I\}$ et Y à valeurs dans $\{y_j, j \in J\}$ où I et J sont deux ensembles finis, alors (X, Y) est à valeurs dans $\{(x_i, y_j), (i, j) \in I \times J\}$. Pour déterminer la loi du couple, dite loi conjointe, il suffit donc de donner la probabilité des événements élémentaires $\mathbf{P}(X = x_i \cap Y = y_j)$

noté également $\mathbf{P}(X = x_i, Y = y_j)$. Quant à la probabilité d'une partie A quelconque de $\mathbb{R}^2$, on aura, comme pour toute variable aléatoire sur un espace fini :

$$\mathbf{P}((X, Y) \in A) = \mathbf{P}(\{\omega \in \Omega \mid (X(\omega), Y\,\omega)) \in A\}$$

$$= \sum_{(i,j) \text{ tels que } (x_i, y_j) \in A} \mathbf{P}(X = x_i, Y = y_j)$$

EXEMPLE :

❑ On lance deux dés, X est le plus grand numéro tiré, Y le plus petit. Quelle est la loi conjointe ? On la présentera sous la forme d'un tableau à double entrée.

X↓ Y→	1	2	3	4	5	6	loi de X
1	1/36	0	0	0	0	0	1/36
2	2/36	1/36	0	0	0	0	3/36
3	2/36	2/36	**1/36**	0	0	0	5/36
4	2/36	**2/36**	**2/36**	**1/36**	0	0	7/36
5	**2/36**	**2/36**	**2/36**	**2/36**	**1/36**	0	9/36
6	**2/36**	**2/36**	**2/36**	**2/36**	**2/36**	**1/36**	11/36
loi de Y	11/36	9/36	7/36	5/36	3/36	1/36	total 1

La convention est ici de mettre les valeurs prises par X dans la première colonne et celles prises par Y dans la première ligne. On prendra garde qu'on aurait pu prendre la convention inverse, ce qui aurait conduit à inverser le rôle des lignes et des colonnes dans toute la suite de ce paragraphe.

Quelle est la probabilité de l'événement X + Y > 5 ? Cette événement consiste en l'appartenance de (X, Y) à la partie $A = \{(x, y) \mid 1 \leq x \leq 6,\ 1 \leq y \leq 6,\ x + y > 5\}$. Les probabilités des événements élémentaires correspondants sont indiquées en gras dans le tableau ci-dessus. On trouve donc pour $\mathbf{P}(X + Y > 5)$, égal à $\mathbf{P}((X, Y) \in A)$, la valeur $\dfrac{26}{36}$ ou $\dfrac{13}{18}$.

2- Lois marginales

Un couple étant donné, on appelle **lois marginales** les lois de X et de Y. Les lois marginales ont été indiquées dans l'exemple précédent. On remarque qu'elles se trouvent facilement à partir de la loi conjointe par addition des lignes ou des colonnes. En effet :

$$X = x_i \Leftrightarrow \exists\, j,\ (X, Y) = (x_i, y_j)$$

donc $\{\omega \in \Omega \mid X(\omega) = x_i\} = \bigcup_{j \in J} \{\omega \in \Omega \mid X(\omega) = x_i,\ Y(\omega) = y_j\}$. Donc :

$$\mathbf{P}(X = x_i) = \sum_{j \in J} \mathbf{P}(X = x_i, Y = y_j\} \qquad \text{(somme des termes de la ligne } i\text{).}$$

De même,

$$\mathbf{P}(Y = y_j) = \sum_{i \in I} \mathbf{P}(X = x_i, Y = y_j\} \qquad \text{(somme des termes de la colonne } j\text{).}$$

On remarquera que la connaissance des lois marginales ne permet pas de reconstituer la loi du couple. On possède moins d'informations.

EXEMPLE :

❑ Considérons un couple (X, Y) dont la loi conjointe est donnée par le tableau suivant :

X↓ Y→	0	1	loi X
0	p	$1/2 - p$	1/2
1	$1/2 - p$	p	1/2
loi de Y	1/2	1/2	total 1

Quelle que soit la valeur de p, entre 0 et $\frac{1}{2}$, la loi conjointe admet les mêmes lois marginales.

3- Lois conditionnelles

Soit (X, Y) un couple de variables aléatoires sur Ω. On suppose l'événement $Y = y_j$ réalisé. On peut alors munir Ω de la loi conditionnelle $\mathbf{P}_{Y=y_j}$. La loi de X sur cet espace s'appelle loi de X sachant $Y = y_j$. On a donc :

$$\mathbf{P}_{Y=y_j}(X = x_i) = \mathbf{P}(X = x_i \mid Y = y_j) = \frac{\mathbf{P}(X = x_i, Y = y_j)}{\mathbf{P}(Y = y_j)}$$

La loi de X sachant $Y = y_j$ se trouve donc en divisant la j-ème colonne par $\mathbf{P}(Y = y_j)$, qui n'est autre que la somme de cette j-ème colonne. Le calcul peut être fait pour toutes les valeurs de y_j, et le résultat présenté sous la forme d'un tableau à double entrée. On pourrait également chercher la loi de Y sachant que $X = x_i$.

Dans le cas du tirage de deux dés, où X représente le plus grand tirage et Y le plus petit, on obtient :

Loi de X↓ sachant que →	Y = 1	Y = 2	Y = 3	Y = 4	Y = 5	Y = 6
1	1/11	0	0	0	0	0
2	2/11	1/9	0	0	0	0
3	2/11	2/9	1/7	0	0	0
4	2/11	2/9	2/7	1/5	0	0
5	2/11	2/9	2/7	2/5	1/3	0
6	2/11	2/9	2/7	2/5	2/3	1

La loi de X sachant que $Y = j$ se lit dans la j-ème colonne

Loi de Y → sachant que ↓	1	2	3	4	5	6
X = 1	1	0	0	0	0	0
X = 2	2/3	1/3	0	0	0	0
X = 3	2/5	2/5	1/5	0	0	0
X = 4	2/7	2/7	2/7	1/7	0	0
X = 5	2/9	2/9	2/9	2/9	1/9	0
X = 6	2/11	2/11	2/11	2/11	2/11	1/11

La loi de Y sachant que $X = i$ se lit dans la i-ème ligne.

Les lois conditionnelles présentent l'intérêt suivant. Si l'on connaît la loi de Y et les lois conditionnelles de X sachant Y, alors on peut reconstituer la loi du couple. En effet :

$$\mathbf{P}(X = x_i, Y = y_j) = \mathbf{P}(Y = y_j) \times \mathbf{P}_{Y=y_j}(X = x_i)$$

4- Variables indépendantes

a) Cas de deux variables

Deux variables aléatoires X et Y à valeurs respectivement dans $\{x_i,\ i \in I\}$ et $\{y_j,\ j \in J\}$, sont dites **indépendantes** si, pour tout événement A et B :

$$\mathbf{P}(X \in A,\ Y \in B) = \mathbf{P}(X \in A) \times \mathbf{P}(Y \in B)$$

En particulier, pour tout i et tout j, les événements $X = x_i$ et $Y = y_j$ sont indépendants. On a donc :

$$\forall\ i \in I,\ \forall\ j \in J,\ \mathbf{P}(X = x_i,\ Y = y_j) = \mathbf{P}(X = x_i) \times \mathbf{P}(Y = y_j)$$

Connaissant les lois marginales, on est alors capable de reconstituer la loi du couple. Les lois conditionnelles de X sachant $Y = y_j$ sont identiques à la loi de X.

Réciproquement, il suffit de vérifier les indépendances des événements $X = x_i$ et $Y = y_j$ pour conclure à l'indépendance de X et Y, car :

$$\begin{aligned}
\mathbf{P}(X \in A,\ Y \in B) &= \sum_{x_i \in A, y_j \in B} \mathbf{P}(X = x_i,\ Y = y_j) \\
&= \sum_{x_i \in A, y_j \in B} \mathbf{P}(X = x_i)\, \mathbf{P}(Y = y_j) \qquad \text{d'après l'hypothèse} \\
&= \sum_{x_i \in A} \mathbf{P}(X = x_i) \sum_{y_j \in B} \mathbf{P}(Y = y_j) \\
&= \mathbf{P}(X \in A)\, \mathbf{P}(Y \in B)
\end{aligned}$$

donc, X et Y sont bien indépendantes.

EXEMPLES :

❑ Soit (X, Y) un couple à valeurs dans $[\![1, n]\!] \times [\![1, n]\!]$. On pose :

$$\mathbf{P}(X = i, Y = j) = C \times \frac{i}{j}$$

où C est une constante telle que la somme des probabilités soit égale à 1. Les variables sont-elles indépendantes ?

La loi de X est :

$$\mathbf{P}(X = i) = \sum_{j=1}^{n} \mathbf{P}(X = i,\ Y = j) = \sum_{j=1}^{n} C \times \frac{i}{j} = CK \times i$$

où $K = \sum_{j=1}^{n} \frac{1}{j}$. C et K sont reliés par la relation suivante :

$$1 = \sum_{i=1}^{n} \mathbf{P}(X = i) = CK \sum_{i=1}^{n} i = CK\, \frac{n(n+1)}{2}$$

La loi de Y est :

$$\mathbf{P}(Y = j) = \sum_{i=1}^{n} \mathbf{P}(X = i,\ Y = j) = \sum_{i=1}^{n} C \times \frac{i}{j} = \frac{C}{j} \sum_{i=1}^{n} i = \frac{C}{j}\, \frac{n(n+1)}{2} = \frac{1}{Kj}$$

On a donc bien :

$$\mathbf{P}(X = i) \times \mathbf{P}(Y = j) = \frac{Ci}{j} = \mathbf{P}(X = i, Y = j).$$

En fait, on peut prévoir l'indépendance des variables si la loi conjointe est de la forme suivante :

$$\mathbf{P}(X = i, Y = j) = f(i)g(j)$$

où f et g sont deux fonctions.

En effet, soit $G = \sum_j g(j)$ et $F = \sum_i f(i)$. Alors F et G sont reliés par la relation :

$$FG = \sum_{i,j} f(i)g(j) = \sum_{i,j} \mathbf{P}(X = i, Y = j) = 1$$

$$\mathbf{P}(X = i) = \sum_j f(i)g(j) = f(i)G$$

$$\mathbf{P}(Y = j) = \sum_i f(i)g(j) = g(j)F$$

D'où $\mathbf{P}(X = i) \times \mathbf{P}(Y = j) = FG \times f(i)g(j) = f(i)g(j) = \mathbf{P}(X = i, Y = j)$ et les variables sont bien indépendantes. Cette propriété apparaît dans le tableau donnant la loi du couple quand les lignes (ou les colonnes) sont proportionnelles entre elles.

❑ Soit (X, Y) un couple à valeurs dans $[\![1, n]\!] \times [\![1, n]\!]$. On pose :

$$\mathbf{P}(X = i, Y = j) = C(i + j) \text{ où C est une constante.}$$

Les variables sont-elles indépendantes ?

La loi de X est :

$$\mathbf{P}(X = i) = \sum_{j=1}^{n} \mathbf{P}(X = i, Y = j) = \sum_{j=1}^{n} C(i + j) = C(ni + \frac{n(n+1)}{2})$$

C vérifie la relation suivante : $1 = \sum_{i=1}^{n} \mathbf{P}(X = i) = Cn^2(n + 1) = 1$

La loi de Y est, par symétrie :

$$\mathbf{P}(Y = j) = C(nj + \frac{n(n+1)}{2})$$

On a donc :

$$\mathbf{P}(X = i) \times \mathbf{P}(Y = j) \neq \mathbf{P}(X = i, Y = j).$$

Les variables ne sont pas indépendantes.

PROPOSITION

Soit $X : \Omega \to \mathbb{R}$ *et* $Y : \Omega \to \mathbb{R}$ *deux variables aléatoires indépendantes. Alors pour toute fonction f et g de* $\mathbb{R}$ *dans* $\mathbb{R}$, $f(X)$ *et* $g(Y)$ *sont indépendantes.*

<u>Démonstration</u> :

❑ Pour toute valeur u_i et v_j que peuvent prendre $f(X)$ et $g(Y)$, on a :

$$\begin{aligned}\mathbf{P}(f(X) = u_i, g(Y) = v_j) &= \mathbf{P}(X \in f^{-1}(u_i), Y \in g^{-1}(v_j)) \\ &= \mathbf{P}(X \in f^{-1}(u_i))\, \mathbf{P}(Y \in g^{-1}(v_j)) \\ &= \mathbf{P}(f(X) = u_i)\, \mathbf{P}(g(Y) = v_j)\end{aligned}$$

b) Cas de *n* variables

n variables aléatoires X_1, X_2, ... X_n à valeurs respectivement dans $E_1 = \{x_{i_1}, i_1 \in I_1\}$, ..., $E_n = \{x_{i_n}, i_n \in I_n\}$, sont dites indépendantes si, pour tout événement A_1 de E_1, ..., A_n de E_n les événements $X_1 \in A_1$, ..., $X_n \in A_n$ sont indépendants, ou, de manière équivalente, si pour tout i_1, ..., tout i_n, les événements $X_1 = x_{i_1}$, ..., $X_n = x_{i_n}$ sont indépendants. Il suffit de vérifier que :

$$\forall\ i_1 \in I_1, ..., \forall\ i_n \in I_n, \mathbf{P}(X_1 = x_{i_1}, ..., X_n = x_{i_n}) = \mathbf{P}(X_1 = x_{i_1}) ...\mathbf{P}(X_n = x_{i_n})$$

car une relation analogue sera vérifiée pour toute sous-famille d'événements en sommant de manière adéquate les indices.

Si X_1, X_2, ... X_n sont indépendantes, on peut vérifier les propriétés suivantes :
i) Toute sous-famille extraite de $(X_1, ..., X_n)$ est indépendante
ii) si f est une fonction de $E_1 \times ... \times E_k$ et g une fonction de $E_{k+1} \times ... \times E_n$ à valeurs dans $\mathbb{R}$, alors les deux variables aléatoires $Y = f(X_1, ..., X_k)$ et $Z = g(X_{k+1}, ..., X_n)$ sont indépendantes.

5- Espérance d'une fonction d'un couple

a) Cas du produit XY :
On souhaite calculer $\mathbf{E}(XY)$, connaissant la loi du couple (X, Y). On a :

$$\mathbf{E}(XY) = \sum_{\omega\in\Omega} X(\omega)Y(\omega)\ \mathbf{P}(\{\omega\})$$

Nous reprenons ci-dessous un raisonnement comparable à celui que nous avons tenu pour une seule variable aléatoire dans III.3). Soient x_1, x_2 , ..., x_n les valeurs que prend X et y_1, y_2, ..., y_m les valeurs prises par Y. Alors les images réciproques $X^{-1}(\{x_1\})$, $X^{-1}(\{x_2\})$, ..., $X^{-1}(\{x_n\})$ forment un système complet d'événements de Ω : ils sont disjoints (car un élément ω ne peut avoir deux images différentes) et leur réunion est Ω (car tout ω admet comme image l'un des x_i) . Il en est de même de $Y^{-1}(\{y_1\})$, $Y^{-1}(\{y_2\})$, ..., $Y^{-1}(\{y_m\})$. Et il en est encore de même de la famille $X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})$, $1 \le i \le n$, $1 \le j \le m$. On peut calculer la somme sur Ω en sommant d'abord sur chaque $X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})$, puis en faisant la somme de ces $n \times m$ totaux :

$$\sum_{\omega\in\Omega} X(\omega)Y(\omega)\ \mathbf{P}(\{\omega\}) = \sum_{i,j}\ \sum_{\omega\in X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})} X(\omega)Y(\omega)\ \mathbf{P}(\{\omega\})$$

D'autre part, lorsque ω appartient à $X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})$, par définition, $X(\omega)$ est égal à x_i et $Y(\omega)$ à y_j. D'où :

$$\sum_{i,j}\ \sum_{\omega\in X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})} X(\omega)Y(\omega)\ \mathbf{P}(\{\omega\}) = \sum_{i,j}\ \sum_{\omega\in X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})} x_i\, y_j\ \mathbf{P}(\{\omega\})$$

$$= \sum_{i,j} x_i\, y_j \sum_{\omega\in X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})} \mathbf{P}(\{\omega\})$$

Enfin $\displaystyle\sum_{\omega\in X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})} \mathbf{P}(\{\omega\})$ n'est autre que :

$$\mathbf{P}(X^{-1}(\{x_i\}) \cap Y^{-1}(\{y_j\})) = \mathbf{P}(\{\omega \in \Omega \mid X(\omega) = x_i \cap Y(\omega) = y_j\}) = \mathbf{P}(X = x_i, Y = y_j)\ .$$

Donc :

$$\sum_{\omega\in\Omega} X(\omega)Y(\omega)\,\mathbf{P}(\{\omega\}) = \sum_{i,j} x_i\, y_j\, \mathbf{P}(X = x_i, Y = y_j)$$

Ainsi :

$$\mathbf{E}(XY) = \sum_{i,j} x_i\, y_j\, \mathbf{P}(X = x_i, Y = y_j)$$

Examinons maintenant le cas où X et Y sont indépendantes. Alors, la formule obtenue devient :

$$\mathbf{E}(XY) = \sum_{i,j} x_i\, y_j\, \mathbf{P}(X = x_i)\,\mathbf{P}(Y = y_j)$$

$$= \sum_{i} x_i\, \mathbf{P}(X = x_i) \sum_{j} y_j\, \mathbf{P}(Y = y_j)$$

$$= \mathbf{E}(X)\mathbf{E}(Y)$$

On peut donc énoncer la proposition suivante :

PROPOSITION

Soit X *et* Y *deux variables indépendantes, alors* $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$

Cette propriété se généralise par récurrence à un produit quelconque de variables aléatoires indépendants.

Si X et Y ne sont pas indépendantes, en général, il n'y a pas égalité.

EXEMPLE :

❑ Soit $0 < p < \frac{1}{2}$. On considère un couple (X, Y) de variables aléatoires dont la loi est donnée par le tableau ci-dessous :

X↓ Y→	0	1	loi de X
0	p	$1/2 - p$	1/2
1	$1/2 - p$	p	1/2
loi de Y	1/2	1/2	total 1

XY suit une loi de Bernoulli de paramètre p. Donc $\mathbf{E}(XY) = p$. Mais $\mathbf{E}(X) = \mathbf{E}(Y) = \frac{1}{2}$, donc $\mathbf{E}(X)\mathbf{E}(Y) = \frac{1}{4}$, différent de p en général.

On prendra garde que la proposition énonce une implication et non une équivalence. Il est tout à fait possible que $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$ sans que X et Y soient indépendantes.

EXEMPLE :

❑ On dispose de trois urnes. On y place trois boules au hasard. X est le nombre de boules dans l'urne 1, et Y est le nombre d'urnes vides. On vérifiera que la loi conjointe est :

X↓ Y→	0	1	2	loi de X
0	0	6/27	2/27	8/27
1	6/27	6/27	0	12/27
2	0	6/27	0	6/27
3	0	0	1/27	1/27
loi de Y	2/9	6/9	1/9	total 1

Par exemple, X = 0 et Y = 1 signifie que l'urne 1 seule est vide, et il y a 6 façons sur 27 possibles de répartir les trois boules dans les urnes 2 et 3, avec au moins une boule dans ces deux urnes (deux façons de choisir l'urne qui ne possèdera qu'une boule, et trois façons de choisir la dite boule).
On a :

$$\mathbf{E}(\mathrm{X}) = 1 \times \frac{12}{27} + 2 \times \frac{6}{27} + 3 \times \frac{1}{27} = 1$$

$$\mathbf{E}(\mathrm{Y}) = 1 \times \frac{6}{9} + 2 \times \frac{1}{9} = \frac{8}{9}$$

$$\mathbf{E}(\mathrm{XY}) = 1 \times \frac{6}{27} + 2 \times \frac{6}{27} + 6 \times \frac{1}{27} = \frac{24}{27} = \frac{8}{9} = \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y})$$

Pourtant, X et Y ne sont pas indépendantes.

b) Cas d'une fonction *h*(X,Y) :

Le raisonnement appliqué au produit se généralise de la façon suivante

$$\mathbf{E}(h(\mathrm{X},\mathrm{Y})) = \sum_{i,j} h(x_i, y_j)\, \mathbf{P}(\mathrm{X} = x_i, \mathrm{Y} = y_j)$$

(Reprendre la démonstration pour le produit XY, en remplaçant X(ω)Y(ω) par h(X(ω), Y(ω)) et x_iy_j par $h(x_i, y_j)$.
Dans le cas de deux variables aléatoires indépendantes, on a :

$$\mathbf{E}(h(\mathrm{X},\mathrm{Y})) = \sum_{i,j} h(x_i, y_j)\, \mathbf{P}(\mathrm{X} = x_i)\, \mathbf{P}(\mathrm{Y} = y_j)$$

De plus, si $h(\mathrm{X}, \mathrm{Y}) = f(\mathrm{X})\, g(\mathrm{Y})$, on obtient :

$$\mathbf{E}(f(\mathrm{X})\, g(\mathrm{Y})) = \sum_{i,j} f(x_i)\, g(y_j)\, \mathbf{P}(\mathrm{X} = x_i)\, \mathbf{P}(\mathrm{Y} = y_j) = \mathbf{E}(f(\mathrm{X}))\, \mathbf{E}(g(\mathrm{Y}))$$

ce qui découle également directement du fait que f(X) et g(Y) sont deux variables aléatoires indépendantes, et que l'espérance d'un produit de variables aléatoires indépendantes est le produit des espérances.

6- Somme de variables aléatoires

On s'intéresse à la loi de la somme Z = X + Y de deux variables aléatoires. L'événement $\{\mathrm{Z} = z\}$ est égal à la réunion disjointe des événements $\{\mathrm{X} = x_i, \mathrm{Y} = y_j \mid x_i + y_j = z\}$. Donc :

$$\mathbf{P}(\mathrm{Z} = z) = \sum_{(i,j) \text{ tq } x_i + y_j = z} \mathbf{P}(\mathrm{X} = x_i, \mathrm{Y} = y_j)$$

Dans le cas de deux variables à valeurs dans $\mathbb{N}$, on obtient :

$$\mathbf{P}(Z = n) = \sum_{i=0}^{n} \mathbf{P}(X = i, Y = n - i)$$

Lorsque les variables sont indépendantes, cette formule devient :

$$\mathbf{P}(Z = n) = \sum_{i=0}^{n} \mathbf{P}(X = i)\, \mathbf{P}(Y = n - i)$$

EXEMPLES :

❑ X et Y deux variables aléatoires indépendantes de même loi de Bernouilli de paramètre p :

$$\mathbf{P}(X + Y = 0) = \mathbf{P}(X = 0)\, \mathbf{P}(Y = 0) = (1 - p)^2$$
$$\mathbf{P}(X + Y = 1) = \mathbf{P}(X = 0)\, \mathbf{P}(Y = 1) + \mathbf{P}(X = 1)\, \mathbf{P}(Y = 0) = 2(1 - p)p$$
$$\mathbf{P}(X + Y = 2) = \mathbf{P}(X = 1)\, \mathbf{P}(Y = 1) = p^2.$$

On reconnaît une loi binomiale $\mathcal{B}(2, p)$, ce qui n'est pas surprenant puisque la loi $\mathcal{B}(2, p)$ est précisément définie comme la loi de la somme de deux variables de Bernoulli indépendantes de même paramètre p.

❑ On généralise ce résultat à la somme de deux variables aléatoires indépendantes suivant une loi binomiale de même paramètre. Soit X de loi $\mathcal{B}(n, p)$ et Y de loi $\mathcal{B}(m, p)$ deux variables aléatoires indépendantes. Alors X + Y admet pour loi une loi $\mathcal{B}(n + m, p)$. En effet :

$$\mathbf{P}(X + Y = k) = \sum_{r} \mathbf{P}(X = r)\, \mathbf{P}(Y = k - r)$$

$$= \sum_{r} \binom{n}{r} p^r (1 - p)^{n-r} \binom{m}{k-r} p^{k-r} (1 - p)^{m-k+r}$$

$$= \sum_{r} \binom{n}{r} \binom{m}{k-r} p^k (1 - p)^{n+m-k}$$

Or $\sum_{r} \binom{n}{r}\binom{m}{k-r} = \binom{n+m}{k}$. En effet, $\binom{n+m}{k}$ est le nombre de parties de k éléments parmi $n + m$ et on peut aussi dénombrer ce nombre de parties de la façon suivante : choisit r éléments parmi n, et $k - r$ parmi m. On obtient finalement :

$$\mathbf{P}(X + Y = k) = \binom{n+m}{k} p^k (1 - p)^{n+m-k}$$

qui est bien la loi binomiale annoncée. Ce résultat est parfaitement compréhensible si on imagine que X est le nombre de Pile en n lancers, et Y le nombre de P en m autres lancers. Alors, X + Y est le nombre de Pile en $n + m$ lancers.

7- Covariance

On s'intéresse à la variance d'un somme :

$$\mathbf{V}(X + Y) = \mathbf{E}(\, (X + Y - \mathbf{E}(X) - \mathbf{E}(Y))^2\,)$$
$$= \mathbf{E}((X - \mathbf{E}(X))^2 + (Y - \mathbf{E}(Y))^2 + 2(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))\,)$$
$$= \mathbf{V}(X) + \mathbf{V}(Y) + 2\, \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$$

On pose alors la **covariance** de X et Y comme égale à :

$$\mathrm{cov}(X,Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$$

D'où :

$$\mathbf{V}(\mathrm{X} + \mathrm{Y}) = \mathbf{V}(\mathrm{X}) + \mathbf{V}(\mathrm{Y}) + 2\ \mathrm{cov}(\mathrm{X},\mathrm{Y})$$

Cette formule est à rapprocher de ce qui ce passe sur $\mathbb{R}$ avec : $(x + y)^2 = x^2 + y^2 + 2xy$. Le rôle du double produit est joué ici par la covariance.

La formule ci-dessus se généralise à n variables de la façon suivante :

$$\mathbf{V}(\mathrm{X}_1 + ... + \mathrm{X}_n) = \mathbf{V}(\mathrm{X}_1) + ... + \mathbf{V}(\mathrm{X}_n) + 2 \sum_{i<j} \mathrm{cov}(\mathrm{X}_i, \mathrm{X}_j)$$

PROPOSITION

La covariance vérifie les propriétés suivantes :

i) $\mathrm{cov}(\mathrm{X}, \mathrm{X}) = \mathbf{V}(\mathrm{X})$

ii) $\mathrm{cov}(\mathrm{X},\mathrm{Y}) = \mathbf{E}(\mathrm{XY}) - \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y})$

iii) $(\mathrm{X},\mathrm{Y}) \to \mathrm{cov}(\mathrm{X},\mathrm{Y})$ *est bilinéaire symétrique.*

iv) $|\mathrm{cov}(\mathrm{X},\mathrm{Y})| \leq \sqrt{\mathrm{V}(\mathrm{X})}\,\sqrt{\mathrm{V}(\mathrm{Y})}$

v) X *et* Y *indépendantes* $\Rightarrow \mathrm{cov}(\mathrm{X}, \mathrm{Y}) = 0$

vi) X *et* Y *indépendantes* $\Rightarrow \mathbf{V}(\mathrm{X} + \mathrm{Y}) = \mathbf{V}(\mathrm{X}) + \mathbf{V}(\mathrm{Y})$

Démonstration :

❑ i) est évident, d'après les définitions de la variance et de la covariance.

❑ ii) se montre en développant cov(X,Y) :

$$\begin{aligned}\mathrm{cov}(\mathrm{X}, \mathrm{Y}) &= \mathbf{E}((\mathrm{X} - \mathbf{E}(\mathrm{X}))(\mathrm{Y} - \mathbf{E}(\mathrm{Y}))) \\ &= \mathbf{E}(\mathrm{XY} - \mathrm{X}\mathbf{E}(\mathrm{Y}) - \mathrm{Y}\mathbf{E}(\mathrm{X}) + \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y})) \\ &= \mathbf{E}(\mathrm{XY}) - \mathbf{E}(\mathrm{X}\mathbf{E}(\mathrm{Y})) - \mathbf{E}(\mathrm{Y}\mathbf{E}(\mathrm{X})) + \mathbf{E}(\mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y})) \\ &= \mathbf{E}(\mathrm{XY}) - \mathbf{E}(\mathrm{Y})\mathbf{E}(\mathrm{X}) - \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y}) + \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y}) \\ &= \mathbf{E}(\mathrm{XY}) - \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y})\end{aligned}$$

❑ iii) La vérification est laissée au lecteur.

❑ iv) résulte du fait qu'une forme bilinéaire symétrique positive vérifie l'inégalité de Schwarz.

❑ v) Si X et Y sont indépendantes, alors $\mathbf{E}(\mathrm{XY}) = \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y})$. La propriété v) en résulte.

❑ vi) est clairement équivalente à v), car en général, $\mathbf{V}(\mathrm{X} + \mathrm{Y}) = \mathbf{V}(\mathrm{X}) + \mathbf{V}(\mathrm{Y}) + 2\ \mathrm{cov}(\mathrm{X},\mathrm{Y})$.

On définit le **coefficient de corrélation** de X et Y comme étant égal à $\rho(\mathrm{X}, \mathrm{Y}) = \dfrac{\mathrm{cov}(\mathrm{X},\mathrm{Y})}{\sqrt{\mathbf{V}(\mathrm{X})}\sqrt{\mathbf{V}(\mathrm{Y})}}$. Ce nombre est compris entre –1 et 1, d'après la relation iv). Il joue un rôle analogue à celui du cosinus pour deux vecteurs.

Si $\mathrm{Y} = a\mathrm{X} + b$, alors $\mathrm{cov}(\mathrm{X}, \mathrm{Y}) = a\mathbf{V}(\mathrm{X})$ et $\mathbf{V}(\mathrm{Y}) = a^2\mathbf{V}(\mathrm{X})$, donc le coefficient de corrélation vaut 1 si $a > 0$ et – 1 si $a < 0$.

Si X et Y sont indépendantes, alors le coefficient de corrélation est nul. On dit que les variables sont **non corrélées**. On prendra garde que cette condition n'est pas équivalente à l'indépendance de X et Y (pas plus que le fait que $\mathbf{E}(\mathrm{XY}) = \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y})$ n'entraîne l'indépendance de X et de Y).

EXEMPLE :

❑ Soit un couple de variables aléatoires ayant la loi suivante :

$\mathbf{P}(X = -1, Y = 0) = \frac{1}{2}$,

$\mathbf{P}(X = 1, Y = 1) = \frac{1}{4}$,

$\mathbf{P}(X = 1, Y = -1) = \frac{1}{4}$

On a :

$\mathbf{P}(X = -1) = \mathbf{P}(X = 1) = \frac{1}{2}$ donc $\mathbf{E}(X) = 0$,

$\mathbf{P}(XY = 0) = \frac{1}{2}$

$\mathbf{P}(XY = 1) = \frac{1}{4}$

$\mathbf{P}(XY = -1) = \frac{1}{4}$ donc $\mathbf{E}(XY) = 0$

donc cov(X, Y) = 0.

Cependant les variables aléatoires ne sont pas indépendantes puisque $\mathbf{P}(X = -1, Y = 0) = \frac{1}{2}$ alors que $\mathbf{P}(X = -1)\mathbf{P}(Y = 0) = \frac{1}{4}$.

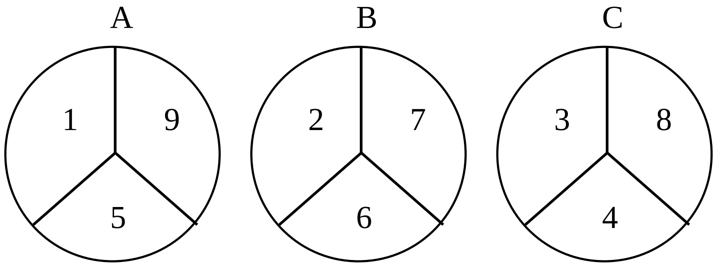
Exercices

1- Enoncés

Exo.1) Titus et Bérénice jouent aux dés. Le dé de Titus est équilibré. Le dé de Bérénice a une probabilité p de tomber sur le chiffre 6 et une probabilité $\frac{1-p}{5}$ de tomber sur chacun des autres chiffres. Le jeu se joue en un lancer de dé pour chaque joueur, le gagnant étant celui qui sort le plus grand chiffre. En cas d'égalité la partie est nulle. Quelle est la probabilité que la partie soit nulle ? Que Titus gagne ? Que Bérénice gagne ? Vérifier que, si $p > \frac{1}{6}$, Bérénice a plus de chance de gagner que Titus.

Exo.2) Dans un groupe de n personnes, quelle est la probabilité que deux personnes au moins aient leur anniversaire le même jour ? (on ne tient pas compte des années bissextiles). Pour quelle valeur de n cette probabilité dépasse-t-elle $\frac{1}{2}$?

Exo.3) On dispose de trois roulettes A, B, C disposant chacune de trois secteurs numérotés comme suit :



Quand on lance une de ces roulettes, chaque secteur a la même probabilité d'être sélectionné. Titus choisit une des roulettes à sa guise, puis Bérénice prend la roulette suivante dans l'ordre $A \to B \to C \to A$. Chacun fait tourner sa propre roulette. Le gagnant est celui qui tire le plus grand nombre. Quelle est la roulette que doit choisir de préférence Titus ?

Exo.4) Dans une loterie, il y a 1000 billets dont 2 sont gagnants. Combien en prendre pour avoir une probabilité de posséder au moins un billet gagnant supérieure à $\frac{5}{100}$?

Exo.5) Soit n un entier strictement positif. On considère un jeu de dominos, chaque domino étant constitué d'une paire de nombres entiers (éventuellement égaux) de 1 à n. Si $i \neq j$, le domino (i, j) est identique au domino (j, i).

a) Combien y a-t-il de tels dominos ?

b) On en prend deux. Quelle est la probabilité qu'ils portent un nombre commun ?

Exo.6) a) Dans une urne, on dispose $2n$ boules numérotées de 1 à $2n$. On en tire une première boule que l'on remet dans l'urne, puis on en tire une deuxième. Quelle est la probabilité que la somme des deux numéros tirés soit pair ? Même question si le tirage est effectué sans remise.

b) Dans une urne, on dispose $3n$ boules numérotées de 1 à $3n$. On en tire trois avec remise de chaque boule tirée. Quelle est la probabilité que la somme des deux numéros tirés soit un multiple de 3 ? Même question si le tirage est effectué sans remise.

Exo.7) On donne ci-dessous plusieurs problèmes datant des débuts du calcul des probabilités, et proposés par Christian Huygens, *De ratiociniis in ludo aleae* (1657)[1] :

a) Proposition XI : "Trouver en combien de fois l'on peut accepter de jeter deux 6 avec deux dés". En version moderne : on lance deux dés n fois de suite. Quelle est la probabilité que l'un des lancers au moins donne un double 6 ? Quelle valeur donner à n pour que cette probabilité dépasse $\frac{1}{2}$?

b) Proposition XII. "Trouver le nombre de dés avec lequel on peut accepter de jeter deux 6 du premier coup". En version moderne : On lance n dés. Quelle est la probabilité d'obtenir au moins deux 6 ? Quelle valeur donner à n pour que cette probabilité dépasse $\frac{1}{2}$?

c) Deuxième problème. "Trois joueurs, A, B et C prennent 12 jetons dont 4 blancs et 8 noirs ; ils jouent à cette condition que celui gagnera qui aura le premier, en choisissant à l'aveuglette, tiré un jeton blanc, et que A choisira le premier, B ensuite, puis C, puis de nouveau A et, ainsi de suite, à tour de rôle. On demande le rapport de leurs chances".

[1] Voir le texte original en ligne : https://gallica.bnf.fr/ark:/12148/bpt6k77862v/f85.item et pages suivantes.

d) Troisième problème. "A parie contre B, que de 40 cartes, dont 10 de chaque couleur, il en tirera 4 de manière à en avoir une de chaque couleur. On trouve dans ce cas que la chance de A est à celle de B comme 1000 est à 8139".

e) Quatrième problème. "On prend 12 jetons, dont 4 blancs et 8 noirs. A parie contre B que parmi 7 jetons qu'il en tirera à l'aveuglette, il se trouvera 3 blancs. On demande le rapport de la chance de A à celle de B".

Exo.8) En 2010, en France, la proportion de naissances issues d'une fécondation in vitro (FIV) était de 1,5%. Parmi les naissances issues d'une FIV, 18% sont des naissances de jumeaux, contre 2% pour l'ensemble de toutes les naissances. Quelle est la probabilité que deux jumeaux nés en 2010, pris au hasard, soient issus d'une FIV ?

Exo.9) On met en œuvre un test détectant une maladie qui touche 1 personne sur 145 dans la population visée. Lors d'un test de dépistage, la probabilité qu'un individu atteint ait un test positif est de 0,96 (les 4% restants sont appelés faux négatifs). La probabilité qu'un individu non atteint ait un test négatif est 0,94 (les 6% restants sont appelés faux positifs). On teste un individu quelconque de la population visée. On observe que le test est positif. Quelle est la probabilité qu'il soit atteint par la maladie ?

Exo.10) Soient A, B et C trois événements dans un espace probabilisé $(\Omega, \mathbf{P})$. Montrer que :

a) A et B indépendants $\Rightarrow$ A et B^c indépendants

b) $\begin{cases} \mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\,\mathbf{P}(B \cap C) \\ \mathbf{P}(A \cap B \cap C^c) = \mathbf{P}(A)\,\mathbf{P}(B \cap C^c) \end{cases} \Rightarrow$ A et B sont indépendants

c) $\mathbf{P}(A \cap B)\,\mathbf{P}(A^c \cap B^c) = \mathbf{P}(A \cap B^c)\,\mathbf{P}(A^c \cap B) \Leftrightarrow$ A et B sont des événements indépendants.

Exo.11) Au casino, un joueur décide de jouer à la roulette tant qu'il n'a pas gagné, tout en se limitant à un maximum de N parties. On note p la probabilité qu'il gagne à chaque partie et X le nombre de parties qu'il joue.

a) Donner $\mathbf{P}(X = n)$ pour $n < N$, puis pour $n = N$. Vérifier que $\sum_{n=1}^{N} \mathbf{P}(X = n) = 1$.

b) Calculer et simplifier l'espérance $\mathbf{E}(X)$ et vérifier le résultat obtenu dans les cas $N = 1$, $N = 2$, ou p tendant vers 0.

Exo.12) On considère une urne de N boules dont n sont blanches. On tire les boules une par une sans remise jusqu'à avoir tiré toutes les boules blanches. Soit X le nombre de boules ainsi tirées. Donner la loi de X et son espérance.

Exo.13) On considère une urne contenant n boules. On tire n boules avec remise. On peut donc tirer plusieurs fois la même boule. Soit X le nombre de boules distinctes tirées au bout des n tirages.

a) Donner l'expression de $\mathbf{E}(X)$ ainsi qu'un équivalent de cette quantité quand n tend vers $+\infty$.

b) Donner la loi de X.

Exo.14) On considère un groupe de n personnes. Chaque personne envoie une lettre à une des $n - 1$ autres personnes choisie au hasard suivant une loi uniforme.

a) Combien existe-t-il de répartitions possibles d'envois ?

b) On considère une personne donnée. Soit X le nombre de lettres qu'elle reçoit. Quelle est la loi de X ? Son espérance ? Sa variance ?

c) Pour tout entier naturel k, déterminer $\alpha_k = \lim\limits_{n\to\infty} \mathbf{P}(X = k)$. Vérifier que $\sum\limits_{k=0}^{\infty} \alpha_k = 1$.

Exo.15) Soit λ un réel élément de]0, 1[et n un entier strictement positif. On désigne par $\lfloor n(1 + \lambda)\rfloor$ le plus grand entier inférieur ou égal à $n(1 + \lambda)$, et par $\lceil n(1 - \lambda)\rceil$ le plus petit entier supérieur ou égal à $n(1 - \lambda)$. Montrer, en utilisant une inégalité de Bienaymé-Tchebychev, que :

$$\sum_{k=\lceil n(1-\lambda)\rceil}^{\lfloor n(1+\lambda)\rfloor} \binom{2n}{k} \geq \left(1 - \frac{1}{2\lambda^2 n}\right) 4^n$$

Exo.16) Un appareil, prévu pour fonctionner pendant une durée T, est formé de n composants. Ces composants sont disposés en parallèle de façon à réduire le risque de panne globale de l'appareil. On cherche à déterminer n pour minimiser le coût de fonctionnement de l'appareil pendant la durée T donnée. Plus précisément, chaque composant coûte un prix égal à a. La probabilité qu'un composant tombe en panne pendant la durée T est p, le risque de panne étant indépendant d'un composant à l'autre. Le système complet tombe en panne quand tous ses composants sont défectueux. Dans ce cas, la panne du système au cours de la durée T entraîne un préjudice dont le coût est égal à b. Déterminer la valeur de n en fonction de a, b, p pour que l'espérance du coût pendant la durée T soit le plus petit possible.

Applications numériques : $a = 1$, $b = 1000$, $p = \frac{1}{100}$, puis $a = 1$, $b = 100\,000$, $p = \frac{1}{100}$

Exo.17) Soit $\Omega = [\![1, 6]\!] \times [\![1, 6]\!]$ l'univers correspondant au lancé de deux dés équiprobables. On considère les variables aléatoires suivantes :

$X(i, j) = \mathrm{Max}(i, j)$ maximum des deux dés

$Y(i, j) = i + j$ somme des deux dés.

Pour tout $x \in [\![1, 6]\!]$, calculer $\mathbf{E}(Y \mid X = x)$ et $\mathbf{P}(X = x)$.

Vérifier que $\mathbf{E}(Y) = \sum\limits_{x=1}^{6} \mathbf{E}(Y \mid X = x) \times \mathbf{P}(X = x)$.

Exo.18) Soit $\mathfrak{S}_n$ le groupe symétrique, dont les éléments sont les permutations de $[\![1, n]\!]$. On définit sur $\mathfrak{S}_n$ une loi de probabilité uniforme et on note X_n la variable aléatoire telle que, pour tout σ élément de $\mathfrak{S}_n$, $X_n(\sigma)$ est égale au nombre de cycles de la permutation σ (voir L1/GROUPSYM.PDF). Par exemple, pour $n = 5$ et σ définie par $\sigma(1) = 4$, $\sigma(2) = 3$, $\sigma(3) = 2$, $\sigma(4) = 1$, $\sigma(5) = 5$, il y a trois cycles : {1, 4}, {2, 3} et {5}. Donc $X(\sigma) = 3$. Si σ est définie par $\sigma(1) = 4$, $\sigma(2) = 1$, $\sigma(3) = 5$, $\sigma(4) = 2$, $\sigma(5) = 3$, il y a deux cycles : {1, 4, 2} et {3, 5}. Donc $X(\sigma) = 2$.

a) Donner une expression de l'espérance de X_n. Pour cela, on pourra déterminer une relation entre X_{n+1} et X_n, en définissant les permutations φ de $[\![1, n + 1]\!]$ à partir des permutations σ de $[\![1, n]\!]$, soit en ajoutant le cycle constitué du seul élément $\{n + 1\}$, soit en insérant l'élément $n + 1$ parmi un cycle de σ. On en déduira ensuite une relation entre $\mathbf{E}(X_{n+1})$ et $\mathbf{E}(X_n)$.

b) Donner une expression de la variance de X_n.

Exo.19) Dans un jeu, on dispose de n cartes portant chacune un nombre entier positif différent. Les nombres en question sont inconnus du joueur. Un animateur bat les cartes (les permutations des cartes ainsi obtenues sont équiprobables), puis les retourne une par une devant le joueur de façon à ce que celui-ci voit le nombre indiqué. Celui-ci arrête l'animateur à son gré et gagne un prix dont le montant est égal au nombre figurant sur la dernière carte retournée.

a) Un joueur décide d'arrêter l'animateur à un rang quelconque, le rang étant choisi avec une loi uniforme sur $[\![1, n]\!]$. Quelle est la probabilité qu'il tombe sur la carte portant le plus grand nombre ?

b) Un autre joueur se donne un entier $m < n$, laisse passer les m premières cartes, note le maximum de ces m cartes puis arrête l'animateur à la première carte qui suit dont le nombre est supérieur à ce maximum (si elle existe, sinon il désigne la dernière carte). On note :

E_k l'événément "La carte portant le plus grand nombre se trouve au rang k",

F_k l'événement "La carte ayant le nombre maximal parmi les cartes de rang 1 à k est l'une des m premières",

G l'événement "Le joueur choisit la carte portant le plus grand nombre".

Donner $\mathbf{P}(E_k)$, $\mathbf{P}(F_k)$, $\mathbf{P}(E_k \cap F_{k-1})$, $\mathbf{P}(G \mid E_k)$, et enfin $\mathbf{P}(G)$.

Donner un équivalent de $\mathbf{P}(G)$ quand n et m tendent vers l'infini de façon que $\dfrac{m}{n}$ tende vers un réel α élément de $]0, 1[$. Comment choisir α pour que ce résultat soit le plus grand possible ?

Exo.20) Une société démarche n clients potentiels par téléphone. Soit $p \in]0,1[$ la probabilité, supposée identique pour tous, qu'un client contacté réponde. Le fait qu'un client réponde est indépendant de ce que font les autres clients. Soit X le nombre de clients qui ont répondu à la société.

a) Quelle est la loi de X et son espérance $\mathbf{E}(X)$?

b) Peu de temps après, la société tente de nouveau de contacter les clients qui n'ont pas répondu la première fois. Si $X = k$, il y a donc $n - k$ clients à démarcher. Soit q (éventuellement différente de p) la probabilité qu'un client contacté pour la deuxième fois réponde. Soit Y le nombre de clients qui répondent. Quelle est la loi de Y, conditionnelle à l'événement $X = k$, et son espérance, notée $\mathbf{E}(Y \mid X = k)$?

c) Quelle est la loi du couple (X, Y) ?

d) Quelle est la loi marginale de Y ? Quelle est son espérance $\mathbf{E}(Y)$? Vérifier que :

$$\mathbf{E}(Y) = \sum_{k=0}^{n} \mathbf{E}(Y \mid X = k)\mathbf{P}(X = k)$$

e) Quelle est la loi de la somme X + Y et son espérance ?

f) Si la société lance un troisième cycle d'appels pour les clients n'ayant pas répondu les deux premières fois, avec une probabilité r qu'un client donné réponde lors de ce troisième cycle, et si Z est le nombre de clients répondant à ce dernier appel, que vaut à votre avis $\mathbf{E}(X + Y + Z)$?

Exo.21) Soit n un entier strictement positif. On considère l'ensemble Ω constitué des points (i, j) de $\mathbb{Z}^2$ tels que $0 \leq i \leq n - |j|$. On définit sur Ω une probabilité P uniforme et on considère sur Ω les deux variables aléatoires X et Y suivantes : $X(i, j) = i$ et $Y(i, j) = j$.

a) Déterminer la loi de X, la loi de Y, $\mathbf{E}(X)$, $\mathbf{E}(Y)$, $\mathbf{E}(|Y|)$. Donner un équivalent de $\mathbf{E}(X)$ et de $\mathbf{E}(|Y|)$ quand n tend vers l'infini.

b) Pour tout j, déterminer la loi conditionnelle de X relativement à l'événement $Y = j$, ainsi que son espérance, que l'on notera $\mathbf{E}(X \mid Y = j)$. Vérifier que $\mathbf{E}(X) = \sum_{j=-n}^{n} \mathbf{E}(X \mid Y = j)\mathbf{P}(Y = j)$.

c) On note f la fonction $j \to \mathbf{E}(X \mid Y = j)$. Calculer $\mathbf{E}(f(Y))$ (où $f(Y) = f \circ Y$) et la comparer à $\mathbf{E}(X)$. Montrer que, pour toute fonction g, $\mathbf{E}((X - f(Y))g(Y)) = 0$. En déduire que $f(Y)$ est la fonction de Y vérifiant : pour toute fonction g, $\mathbf{V}(X - f(Y)) \leq \mathbf{V}(X - g(Y))$. ($f(Y)$ est la fonction de Y approchant le mieux X au sens de la moindre variance).

Exo.22) On considère n variables aléatoires indépendantes X_i de même loi, d'espérance μ et de variance σ^2. Soit $Z = \frac{1}{n}\sum_{i=1}^{n} X_i$. Montrer que $\frac{1}{n-1}\sum_{i=1}^{n} \mathbf{E}((X_i - Z)^2) = \sigma^2$.

Exo.23) a) n personnes se répartissent au hasard dans trois pièces. Soit X_i le nombre de personnes dans la pièce i. Donner la loi des X_i, celle de $X_1 + X_2$, et calculer le coefficient de corrélation $\rho(X_1, X_2)$.

b) On lance un dé n fois. Soit X_1 (respectivement X_2) le nombre de fois où le chiffre 1 (respectivement le nombre de 2) sort. Calculer le coefficient de corrélation $\rho(X_1, X_2)$.

c) Soient $X_1, ..., X_m$ des variables aléatoires ayant même variance non nulle $\mathbf{V}(X_i) = \sigma^2$, et même covariance $\mathrm{cov}(X_i, X_j) = \rho\sigma^2$ ($i \neq j$). Montrer que $\rho \geq -\frac{1}{m-1}$.

d) Justifier que le a) et le b) relèvent du c).

Exo.24) Soit X de loi binomiale $\mathcal{B}(N, p)$, n un entier inférieur ou égal à N, et Y une variable aléatoire telle que, pour tout m élément de $[\![0, N]\!]$, et tout k tel que $0 \leq k \leq m$, et $0 \leq n - k \leq N - m$:

$$\mathbf{P}(Y = k \mid X = m) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$$

Autrement dit, la loi de Y conditionnellement à l'événement $X = m$ et la loi hypergéométrique de paramètres N, m, n.

a) Déterminer la loi de Y.

b) Déterminer $\mathbf{P}(X = m + k \mid Y = k)$

c) Considérant N objets, on les prend un par un et on dispose une marque sur chacun d'eux avec une probabilité p. Puis on en prend n au hasard sans remise et on compte combien sont marqués. Utiliser cette situation pour interpréter le a) et le b).

2- Solutions

Sol.1) $\mathbf{P}(\text{nulle}) = 5 \times \frac{1-p}{5} \times \frac{1}{6} + p \times \frac{1}{6} = \frac{1}{6}$

$\mathbf{P}$(Titus gagne) = $\sum_{k=2}^{6} \frac{1}{6} \times (k-1) \times \frac{1-p}{5} = \frac{1}{6} \times \frac{5 \times 6}{2} \times \frac{1-p}{5} = \frac{1-p}{2}$. Le calcul est fait en considérant que Titus sort un nombre k variant de 2 à 6, et Bérénice un nombre strictement inférieur.

$\mathbf{P}$(Bérénice gagne) = 1 – $\mathbf{P}$(nulle) – $\mathbf{P}$(Titus gagne) = $\frac{1}{3} + \frac{p}{2}$. On peut aussi faire un calcul direct en considérant que Titus sort un nombre k de 1 à 5, et que Bérénice sort un nombre strictement supérieur :

$\mathbf{P}$(Bérénice gagne) = $\sum_{k=1}^{5} \frac{1}{6} \times (p + (5-k)\frac{1-p}{5}) = \frac{5p}{6} + \frac{1}{6} \times \frac{4 \times 5}{2} \times \frac{1-p}{5} = \frac{1}{3} + \frac{p}{2}$

$\mathbf{P}$(Bérénice gagne) > $\mathbf{P}$(Titus gagne) $\Leftrightarrow \frac{1}{3} + \frac{p}{2} > \frac{1-p}{2} \Leftrightarrow p > \frac{1}{6}$

Sol.2) Cherchons plutôt la probabilité que les anniversaires soient tous différents :

$$\frac{365 \times 364 \times ... \times (365-n+1)}{365^n}$$

Donc la probabilité cherchée est $1 - \frac{365 \times 364 \times ... \times (365-n+1)}{365^n}$. Elle dépasse $\frac{1}{2}$ pour $n = 23$. Autrement dit, dans une classe de 23 élèves ou plus, il y a plus d'une chance sur deux que deux élèves aient leur anniversaire le même jour.

Sol.3) Si Titus choisit A, la probabilité qu'il gagne est $\frac{1}{9} + \frac{3}{9} = \frac{4}{9}$

S'il choisit B, la probabilité qu'il gagne est $\frac{2}{9} + \frac{2}{9} = \frac{4}{9}$

S'il choisit C, la probabilité qu'il gagne est $\frac{1}{9} + \frac{1}{9} + \frac{2}{9} = \frac{4}{9}$

Le choix de la roulette n'a aucune importance, et il n'y a pas de roulette plus favorable qu'une autre. Quelle que soit la roulette choisie par Titus, Bérénice a plus de chance de gagner que lui.

Sol.4) Si on achète k billets, la probabilité d'avoir au moins un gagnant (événement complémentaire de n'avoir aucun billet gagnant) est :

$$1 - \frac{\binom{998}{k}}{\binom{1000}{k}} = 1 - \frac{(1000-k)!\ 998!}{(998-k)!\ 1000!} = 1 - \frac{(1000-k)(999-k)}{1000 \times 999}$$

On veut que :

$$1 - \frac{(1000-k)(999-k)}{1000 \times 999} > \frac{5}{100}$$

$$\Leftrightarrow \quad \frac{(1000-k)(999-k)}{1000 \times 999} < \frac{95}{100}$$

$$\Leftrightarrow \quad k^2 - 1999k + 999000 < 950 \times 999$$

$$\Leftrightarrow \quad k^2 - 1999k + 999 \times 50 < 0$$

Les zéros du trinôme sont (avec deux décimales) 25.31 et 1973.69, donc on prend au moins $k = 26$ billets.

Sol.5) a) $\dfrac{n(n+1)}{2}$

b) ❑ Pour tout couple (i, j), supposons que le premier domino soit (i, j). Notons A_{ij} l'événement consistant à tirer un deuxième domino ayant un numéro en commun avec le domino (i, j). La probabilité de l'événement A_{ij} sachant que le premier domino tiré est (i, j) est :

si $i = j$: $\mathbf{P}_{ij}(A_{ij}) = \dfrac{n-1}{\dfrac{n(n+1)}{2} - 1}$ (il y a $n - 1$ dominos possibles)

si $i \neq j$: $\mathbf{P}_{ij}(A_{ij}) = \dfrac{2(n-1)}{\dfrac{n(n+1)}{2} - 1}$ (il y a $2(n - 1)$ dominos possibles)

Pour tout couple (i, j), la probabilité de tirer un premier domino (i, j) puis un second domino dont un des numéros est identique à ceux du premier est alors :

si $i = j$: $\mathbf{P}(\text{tirer } (i, j) \text{ et } A_{ij}) = \mathbf{P}(\text{tirer } (i, j)) \times \mathbf{P}_{ij}(A_{ij}) = \dfrac{2}{n(n+1)} \dfrac{n-1}{\dfrac{n(n+1)}{2} - 1}$

si $i \neq j$: $\mathbf{P}(\text{tirer } (i, j) \text{ et } A_{ij}) = \dfrac{2}{n(n+1)} \dfrac{2(n-1)}{\dfrac{n(n+1)}{2} - 1}$

La probabilité de tirer un deuxième domino ayant un numéro commun avec le premier est :

$$\sum_{i,j} \mathrm{P}((i, j) \text{ et } A_{ij}) = n \times \frac{2}{n(n+1)} \frac{n-1}{\dfrac{n(n+1)}{2} - 1} + \left(\frac{n(n+1)}{2} - n\right) \frac{2}{n(n+1)} \frac{2(n-1)}{\dfrac{n(n+1)}{2} - 1}$$

$$= \frac{4n}{(n+1)(n+2)}$$

❑ On peut aussi procéder comme suit. On considère l'ensemble Ω des couples de dominos. Notons un tel couple de domino sous la forme $((a, b), (c, d))$ où (a, b) représente le premier domino, $1 \leq a \leq b \leq n$, et (c, d) le deuxième domino, $1 \leq c \leq d \leq n$, $(a, b) \neq (c, d)$.

$$\mathrm{Card}(\Omega) = \frac{n(n+1)}{2} \left(\frac{n(n+1)}{2} - 1\right) = \frac{n(n+1)}{2} \frac{n^2 + n - 2}{2} = \frac{(n-1)n(n+1)(n+2)}{4}$$

La probabilité sur Ω est implicitement supposée uniforme. Soit A l'ensemble des couples $((a, b), (c, d))$ avec $a = c$ ou $a = d$ ou $b = c$ ou $b = d$. Partitionnons A en trois sous-ensembles :

A_1 : Il n'y a pas de dominos double. $\mathrm{Card}(A_1) = n(n-1)(n-2)$

A_2 : Le premier domino est un double ($a = b$). $\mathrm{Card}(A_2) = n(n-1)$

A_3 : Le deuxième domino est un double ($c = d$). $\mathrm{Card}(A_2) = n(n-1)$

Donc $\mathrm{Card}(A) = n^2(n-1)$ et la probabilité est donc de $\dfrac{n^2(n-1)}{\dfrac{(n-1)n(n+1)(n+2)}{4}} = \dfrac{4n}{(n+1)(n+2)}$

Sol.6) a) Avec remise, la probabilité est celle de tirer une deuxième boule de parité donnée (celle de la première boule) parmi $2n$. C'est $\dfrac{1}{2}$.

Sans remise, le tirage peut s'opérer comme suit. On tire une première boule portant le numéro a_1. Il reste $2n - 1$ boules dont $n - 1$ ont la même parité que a_1 et n une parité différente. La probabilité est $\dfrac{n-1}{2n-1}$.

b) Avec remise, la probabilité est celle de tirer la troisième boule parmi $3n$ de façon que la valeur de cette troisième boule modulo 3 soit donnée (opposée à la somme des valeurs des deux premières boules). C'est $\frac{1}{3}$.

Sans remise, le tirage peut s'opérer comme suit. On tire une première boule portant le numéro a_1. Il reste $3n - 1$ boules dont $n - 1$ ont le même reste modulo 3 que a_1 et $2n$ un reste différent. On tire ensuite une boule portant le numéro a_2. Il y a deux cas :

ou bien a_2 a même reste que a_1 (cas qui se produit avec une probabilité $\frac{n-1}{3n-1}$) et il reste $3n - 2$ boules dont $n - 2$ ont le même reste. On souhaite que la troisième boule a_3 ait ce reste, ce qui arrive avec une probabilité $\frac{n-2}{3n-2}$.

ou bien a_2 a un reste différent de celui de a_1 (cas qui se produit avec une probabilité $\frac{2n}{3n-1}$). Il reste $3n - 2$ boules, dont $n - 1$ ont même reste que a_1, $n - 1$ même reste que a_2, n un reste différent. C'est parmi ces dernières qu'on doit tirer a_3, ce qui arrive avec une probabilité $\frac{n}{3n-2}$.

La formule des probabilités totales donne :

$$\frac{n-1}{3n-1} \times \frac{n-2}{3n-2} + \frac{2n}{3n-1} \times \frac{n}{3n-2} = \frac{3n^2 - 3n + 2}{(3n-1)(3n-2)}$$

Sol.7) a) La probabilité de sortir un double 6 en lançant deux dés est $\frac{1}{36}$. La probabilité de ne jamais obtenir un double 6 en n lancers est $(\frac{35}{36})^n$. La probabilité d'obtenir au moins un double 6 est $1 - (\frac{35}{36})^n$. On dépasse $\frac{1}{2}$ pour $n \geq \frac{\ln(2)}{\ln(\frac{36}{35})} \approx 24.605$ donc $n = 25$ convient.

b) La probabilité de ne sortir aucun 6 est $(\frac{5}{6})^n$ et la probabilité d'en sortir un seul est $n\,\frac{5^{n-1}}{6^n}$. Donc la probabilité demandée est $1 - (\frac{5}{6})^n - n\,\frac{5^{n-1}}{6^n}$. Pour $n = 9$, on obtient 0.457..., et pour $n = 10$, on obtient 0.515.... Donc $n = 10$ convient.

c) Les chances de gain successifs sont :

A	tirage : B	probabilité : $\frac{4}{12}$
B	tirage : NB	probabilité : $\frac{8}{12} \times \frac{4}{11}$
C	tirage : NNB	probabilité : $\frac{8}{12} \times \frac{7}{11} \times \frac{4}{10}$
A	tirage : NNNB	probabilité : $\frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{4}{9}$
B	tirage : NNNNB	probabilité : $\frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8}$
C	tirage : NNNNNB	probabilité : $\frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{4}{7}$

A tirage : NNNNNNB probabilité : $\frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7} \times \frac{4}{6}$

B tirage : NNNNNNNB probabilité : $\frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7} \times \frac{2}{6} \times \frac{4}{5}$

C tirage : NNNNNNNNB probabilité : $\frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7} \times \frac{2}{6} \times \frac{1}{5} \times 1$

Les probabilités de gain respectives sont :

A : $\frac{4}{12} + \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{4}{9} + \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7} \times \frac{4}{6} = \frac{7}{15} = \frac{77}{165}$

B : $\frac{8}{12} \times \frac{4}{11} + \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} + \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7} \times \frac{2}{6} \times \frac{4}{5} = \frac{53}{165}$

C : $\frac{8}{12} \times \frac{7}{11} \times \frac{4}{10} + \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{4}{7} + \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7} \times \frac{2}{6} \times \frac{1}{5} \times 1 = \frac{7}{33} = \frac{35}{165}$

d) La probabilité de gain de A est $1 \times \frac{30}{39} \times \frac{20}{38} \times \frac{10}{37} = \frac{1000}{9139}$ et celle de B est $1 - \frac{1000}{9139} = \frac{8139}{9139}$.

e) Le nombre de cas possibles est $\binom{12}{7} = 792$. Le nombre de cas favorable est $\binom{4}{3}\binom{8}{4} = 280$. La probabilité de gain de A est donc $\frac{280}{792} = \frac{35}{99}$. Celle de B est $\frac{64}{99}$.

Sol.8) On applique la formule de Bayes.

$\mathbf{P}(\text{FIV}) = 0{,}015$

$\mathbf{P}(\text{J}) = 0{,}02$

$\mathbf{P}_{\text{FIV}}(\text{J}) = 0{,}18$

donc $\mathbf{P}(\text{FIV} \cap \text{J}) = \mathbf{P}_{\text{FIV}}(\text{J}) \times \mathbf{P}(\text{FIV}) = 0{,}18 \times 0{,}015 = 0{,}0027$

et $\mathbf{P}_{\text{J}}(\text{FIV}) = \frac{\mathbf{P}(\text{FIV} \cap \text{J})}{\mathbf{P}(\text{J})} = \frac{0{,}0027}{0{,}02} = 0{,}135$

Sol.9) On applique la formule de Bayes. Soit A l'événement "être atteint par la maladie", et B l'évènement "réagir positivement au test".

On a, selon les hypothèses données :

$\mathbf{P}(\text{B} \mid \text{A}) = 0{,}96$

$\mathbf{P}(\text{B}^c \mid \text{A}^c) = 0{,}94$ donc $\mathbf{P}(\text{B} \mid \text{A}^c) = 0{,}06$ (faux positifs)

$\mathbf{P}(\text{A}) = \frac{1}{145}$ donc $\mathbf{P}(\text{A}^c) = \frac{144}{145}$

On demande $\mathbf{P}(\text{A} \mid \text{B})$:

$$\mathbf{P}(\text{A} \mid \text{B}) = \frac{\mathbf{P}(\text{A} \cap \text{B})}{\mathbf{P}(\text{B})} = \frac{\mathbf{P}(\text{B} \mid \text{A})\, \mathbf{P}(\text{A})}{\mathbf{P}(\text{B} \mid \text{A})\, \mathbf{P}(\text{A}) + \mathbf{P}(\text{B} \mid \text{A}^c)\, \mathbf{P}(\text{A}^c)} = \frac{0{,}96}{0{,}96 + 0{,}06 \times 144} = \frac{1}{10}$$

Cet exercice illustre la difficulté d'application des tests de dépistage. Appliqués sur une population trop grande dans laquelle la proportion de malades est faible, la plupart des tests positifs proviennent de faux positifs.

Sol.10) a) Comme A est réunion disjointe de $\text{A} \cap \text{B}$ et de $\text{A} \cap \text{B}^c$, on a :

$\mathbf{P}(\text{A}) = \mathbf{P}(\text{A} \cap \text{B}) + \mathbf{P}(\text{A} \cap \text{B}^c)$

$= \mathbf{P}(\text{A})\mathbf{P}(\text{B}) + \mathbf{P}(\text{A} \cap \text{B}^c)$ car A et B sont indépendants

donc $\mathbf{P}(\text{A} \cap \text{B}^c) = \mathbf{P}(\text{A}) - \mathbf{P}(\text{A})\mathbf{P}(\text{B}) = \mathbf{P}(\text{A})(1 - \mathbf{P}(\text{B})) = \mathbf{P}(\text{A})\mathbf{P}(\text{B}^c)$

donc A et B^c sont indépendants. On montre de même que A^c et B^c sont indépendants.
b) Il suffit d'ajouter membre à membre les deux égalités
c) $\Leftarrow$ est facile en utilisant le fait que l'on a également A et B^c indépendants, A^c et B^c indépendants, et A^c et B^c indépendants. Les deux membres sont alors égaux à $\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(A^c)\mathbf{P}(B^c)$.
Réciproquement, si $\mathbf{P}(A \cap B)\ \mathbf{P}(A^c \cap B^c) = \mathbf{P}(A \cap B^c)\ \mathbf{P}(A^c \cap B)$, alors :

$$\mathbf{P}(A\cap B)\ \mathbf{P}(A^c\cap B^c) + \mathbf{P}(A\cap B)\ \mathbf{P}(A^c\cap B) = \mathbf{P}(A\cap B^c)\ \mathbf{P}(A^c\cap B) + \mathbf{P}(A\cap B)\ \mathbf{P}(A^c\cap B)$$

donc $\quad \mathbf{P}(A \cap B)\ \mathbf{P}(A^c) = \mathbf{P}(A)\ \mathbf{P}(A^c \cap B)$

$\Leftrightarrow \quad \mathbf{P}(A \cap B)\ (1 - \mathbf{P}(A)) = \mathbf{P}(A)\ \mathbf{P}(A^c \cap B)$

$\Leftrightarrow \quad \mathbf{P}(A \cap B) = \mathbf{P}(A)\ \mathbf{P}(A^c \cap B) + \mathbf{P}(A \cap B)\mathbf{P}(A)$

$\Leftrightarrow \quad \mathbf{P}(A \cap B) = \mathbf{P}(A)\ \mathbf{P}(B)$

et A et B sont bien indépendants.

Sol.11) a) $\mathbf{P}(X = n) = (1-p)^{n-1}p$ pour $n < N$

$\mathbf{P}(X = N) = (1-p)^{N-1}p + (1-p)^N$ (soit le joueur gagne pour la première fois à la N-ème partie, soit il perd sur ses N parties).

b) $\mathbf{E}(X) = \sum\limits_{n=1}^{N} n(1-p)^{n-1}p + N(1-p)^N$

On calcule la première somme en partant de $\sum\limits_{n=0}^{N} x^n = \dfrac{1-x^{N+1}}{1-x}$ qu'on dérive :

$$\sum_{n=1}^{N} nx^{n-1} = \frac{-(N+1)x^N(1-x) + 1 - x^{N+1}}{(1-x)^2} = \frac{1 + Nx^{N+1} - (N+1)x^N}{(1-x)^2}$$

Donc :

$$\begin{aligned}\mathbf{E}(X) &= \frac{1 + N(1-p)^{N+1} - (N+1)(1-p)^N}{p} + N(1-p)^N \\ &= \frac{1 + N(1-p)^{N+1} - (N+1)(1-p)^N + Np(1-p)^N}{p} \\ &= \frac{1 + N(1-p)^{N+1} - (N+1)(1-p)^N - N(1-p)^{N+1} + N(1-p)^N}{p} \\ &= \frac{1-(1-p)^N}{p}\end{aligned}$$

Pour $N = 1$, on a $X = 1$ et $\mathbf{E}(X) = 1$.
Pour $N = 2$, on a $\mathbf{P}(X = 1) = p$ et $\mathbf{P}(X = 2) = 1 - p$ donc $\mathbf{E}(X) = 2 - p$.
Si p tend vers 0, $\mathbf{E}(X)$ tend vers N.

Sol.12) Tirons toutes les boules une par une en notant pour chacune d'elle son rang de tirage. On obtient ainsi une numérotation des N boules de 1 à N et X et le rang le plus élevé correspondant à une boule blanche. Le problème est alors équivalent au suivant : parmi les N entiers de 1 à N, on choisit une partie de n éléments et on note X le plus grand élément choisi (au lieu de fixer les boules blanches et de choisir un tirage aléatoire, on fixe le tirage et ce sont n boules qu'on colorie aléatoirement en blanc). Dans cette configuration, le nombre de parties possibles est $\dbinom{N}{n}$ et le nombre de parties satisfaisant l'événement $X = k$ est $\dbinom{k-1}{n-1}$ (la partie comprend l'élément k et $n-1$ éléments plus petits). Donc :

$$\mathbf{P}(X = k) = \frac{\binom{k-1}{n-1}}{\binom{N}{n}}$$

$$\mathbf{E}(X) = \sum_{k=n}^{N} k \frac{\binom{k-1}{n-1}}{\binom{N}{n}} = n \sum_{k=n}^{N} \frac{\binom{k}{n}}{\binom{N}{n}} = n \frac{\binom{N+1}{n+1}}{\binom{N}{n}} = \frac{n(N+1)}{n+1}$$

On a utilisé la formule $\sum_{k=n}^{N} \binom{k}{n} = \binom{N+1}{n+1}$ vue dans L1/DENOMBRE.PDF.

Sol.13) a) Supposons les boules numérotées de 1 à n et notons $X_i = 1$ si la boule i a été tirée au moins une fois. X_i suit une loi de Bernoulli avec :

$$\mathbf{P}(X_i = 0) = (1 - \frac{1}{n})^n \text{ et } \mathbf{P}(X_i = 1) = 1 - (1 - \frac{1}{n})^n = \mathbf{E}(X_i)$$

Par ailleurs, $X = \sum_{i=1}^{n} X_i$, donc :

$$\mathbf{E}(X) = n(1 - (1 - \frac{1}{n})^n) = n(1 - \exp(n\ln(1 - 1/n)) = n(1 - \exp(-1 + o(1)) \sim n(1 - \frac{1}{e})$$

b) On prendra garde que les X_i ne sont pas indépendantes et donc que X ne suit pas une loi binomiale. L'univers Ω constitué de tous les tirages possibles est l'ensemble des fonctions f de $[\![1, n]\!]$ dans $[\![1, n]\!]$, où $f(i)$ est le numéro de la boule tirée au i-ème tirage, et son cardinal est n^n. Dire que $X = k$, c'est dire que f prend k valeurs distinctes, donc est une application surjective de $[\![1, n]\!]$ sur une partie constituée de k éléments. Si on note $\sigma(n, k)$ le nombre de surjections d'un ensemble à n éléments sur un ensemble à k éléments, il y a $\binom{n}{k} \sigma(n, k)$ telles applications (choix de k éléments parmi $n \times$ choix d'une surjection sur ces k éléments). Donc :

$$\mathbf{P}(X = k) = \binom{n}{k} \frac{\sigma(n, k)}{n^n}$$

Une expression explicite de $\sigma(n, k)$ est donnée en annexe de L1/DENOMBRE.PDF. On peut alors retrouver l'expression de $\mathbf{E}(X)$ à partir de la loi de X, mais à l'issue de calculs assez longs (en particulier, il conviendra de permuter de signes de sommation).

Sol.14) a) $(n - 1)^n$.

b) Il s'agit d'une loi binomiale $\mathcal{B}(n - 1, \frac{1}{n-1})$.

$$\mathbf{E}(X) = 1$$

$$\mathbf{V}(X) = (n - 1) \frac{1}{n-1} (1 - \frac{1}{n-1}) = \frac{n-2}{n-1}$$

c) $\mathbf{P}(X = k) = \binom{n}{k} \frac{1}{(n-1)^k} (1 - \frac{1}{n-1})^{n-k} = \frac{n(n-1)...(n-k+1)}{k!} \frac{1}{(n-1)^k} (1 - \frac{1}{n-1})^{n-k}$

$$\sim \frac{n^k}{k!} \frac{1}{n^k} \ (1 - \frac{1}{n-1})^{n-1} \to \frac{e^{-1}}{k!} \text{ quand } n \text{ tend vers l'infini}$$

Nous verrons dans L2/PROBA2.PDF que les valeurs $\alpha_k = \dfrac{e^{-1}}{k!}$ définissent une loi de probabilité sur l'ensemble infini $\mathbb{N}$ appelée loi de Poisson de paramètre 1. On a $\sum_{k=0}^{\infty} \alpha_k = \sum_{k=0}^{\infty} \dfrac{e^{-1}}{k!} = 1$ car $\sum_{k=0}^{\infty} \dfrac{1}{k!} = e$. Voir L1/DLTAYLOR.PDF.

Sol.15) Soit X une variable aléatoire suivant une loi binomiale $\mathcal{B}(2n, \frac{1}{2})$. On a :

$$\sum_{k=\lceil n(1-\lambda)\rceil}^{\lfloor n(1+\lambda)\rfloor} \binom{2n}{k} \frac{1}{2^{2n}} = \mathbf{P}(n(1-\lambda) \leq X \leq n(1+\lambda)) = \mathbf{P}(|X-n| \leq \lambda n) = 1 - \mathbf{P}(|X-n| > \lambda n)$$

Or l'inégalité de Bienaymé-Tchebychev donne :

$$\mathbf{P}(|X-n| > \lambda n) \leq \mathbf{P}(|X-n| \geq \lambda n) \leq \frac{\mathbf{V}(X)}{\lambda^2 n^2} = 2n \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{\lambda^2 n^2} = \frac{1}{2\lambda^2 n}$$

donc $\displaystyle\sum_{k=\lceil n(1-\lambda)\rceil}^{\lfloor n(1+\lambda)\rfloor} \binom{2n}{k} \geq (1 - \frac{1}{2\lambda^2 n})\, 4^n$.

On rappelle que, d'après la formule du binôme de Newton, $\displaystyle\sum_{k=0}^{2n} \binom{2n}{k} = 2^{2n} = 4^n$. La formule précédente donne une évaluation grossière de la concentration des valeurs de $\binom{2n}{k}$ dans l'intervalle $[n(1-\lambda), n(1+\lambda)]$. Elle ne présente d'intérêt que si $\lambda > \dfrac{1}{\sqrt{2n}}$. Par exemple, pour $n = 20$ et $\lambda = 0.6$, $\displaystyle\sum_{k=\lceil n(1-\lambda)\rceil}^{\lfloor n(1+\lambda)\rfloor} \binom{2n}{k}$ vaut 1.099×10^{12}, alors que $(1 - \frac{1}{2\lambda^2 n})\, 4^n = 1.023 \times 10^{12}$.

Sol.16) Le coût vaut na s'il y a au moins un composant en marche à l'issue de la durée garantie, et $na + b$ si tous les composants tombent en panne. Donc l'espérance du coût est :

$$E_n = na + bp^n$$

On étudie le sens de variation de la suite (E_n) :

$$E_{n+1} - E_n = a + bp^n(p-1) \leq 0 \Leftrightarrow bp^n \geq \frac{a}{1-p} \Leftrightarrow \ln(b) + n\ln(p) \geq \ln(\frac{a}{1-p})$$

$$\Leftrightarrow n \leq \frac{\ln(\frac{a}{b(1-p)})}{\ln(p)}$$

Donc le minimum de l'espérance est obtenu pour $\left\lfloor \dfrac{\ln(\frac{a}{b(1-p)})}{\ln(p)} \right\rfloor + 1$.

Applications numériques : On trouve $n = 2$, puis $n = 3$.

Sol.17) On vérifiera que :

$\mathbf{E}(Y \mid X = 1\} = 2$ $\qquad\qquad$ $\mathbf{P}(X = 1) = \dfrac{1}{36}$

$\mathbf{E}(Y \mid X = 2\} = \dfrac{10}{3}$ $\qquad$ $\mathbf{P}(X = 2) = \dfrac{3}{36}$

$\mathbf{E}(Y \mid X = 3\} = \dfrac{24}{5}$ $\qquad$ $\mathbf{P}(X = 3) = \dfrac{5}{36}$

$\mathbf{E}(Y \mid X = 4\} = \dfrac{44}{7}$ $\qquad$ $\mathbf{P}(X = 4) = \dfrac{7}{36}$

$\mathbf{E}(Y \mid X = 5\} = \dfrac{70}{9}$ $\qquad$ $\mathbf{P}(X = 5) = \dfrac{9}{36}$

$\mathbf{E}(Y \mid X = 6\} = \dfrac{102}{11}$ $\qquad$ $\mathbf{P}(X = 6) = \dfrac{11}{36}$

On retrouve :

$$\mathbf{E}(Y) = \sum_{x=1}^{6} \mathbf{E}(Y \mid X = x) \times \mathbf{P}(X = x) = \frac{2 + 10 + 24 + 44 + 70 + 102}{36} = 7$$

Sol.18) a) Soit σ une permutation de $[\![1, n]\!]$. On peut lui associer $n + 1$ permutations φ de $[\![1, n + 1]\!]$ de la façon suivante :

❑ La permutation φ est telle que $\varphi(n + 1) = n + 1$ et $\forall\ k \leq n$, $\varphi(k) = \sigma(k)$. On a dans ce cas $X_{n+1}(\varphi) = X_n(\sigma) + 1$ car φ possède le cycle $\{n + 1\}$ en plus.

❑ Pour chaque m élément de $[\![1, n]\!]$, on définit une permutation φ par $\varphi(\sigma^{-1}(m)) = n + 1$, $\varphi(n + 1) = m$, $\varphi(k) = \sigma(k)$ pour k différent de $\sigma^{-1}(m)$ et de $n + 1$. Dans le cycle de σ contenant les éléments $\sigma^{-1}(m)$ et son image m par σ, on intercale le nombre $n + 1$ entre $\sigma^{-1}(m)$ et m. On a dans ce cas $X_{n+1}(\varphi) = X_n(\sigma)$ car on a juste allongé un cycle de σ sans modifier le nombre de cycles.

Notons A l'événement $\{\varphi \in \mathfrak{S}_{n+1} \mid \varphi(n + 1) = n + 1\}$ et B son complémentaire.
Inversement, on retrouve σ à partir de φ en prenant $\sigma(k) = \varphi(k)$ pour tout $k \leq n$ si A est réalisé, et sinon, $\sigma(k) = \varphi(k)$ pour k différent de $\varphi^{-1}(n + 1)$, $\sigma(\varphi^{-1}(n + 1)) = \varphi(n + 1)$. On oublie le terme $n + 1$ dans le cycle de φ le contenant.
Notons π l'application ainsi définie $\varphi \to \sigma$.
Dans l'espace probabilisé $\mathfrak{S}_{n+1}$, on a :

$$\mathbf{P}(A) = \frac{n!}{(n + 1)!} = \frac{1}{n + 1} \qquad \text{et} \qquad \mathbf{P}(B) = 1 - \mathbf{P}(A) = \frac{n}{n + 1}$$

La relation $X_{n+1}(\varphi) = \begin{cases} X_n(\sigma) + 1 \text{ si A est réalisé} \\ X_n(\sigma) \text{ sinon} \end{cases}$ peut se traduire, à l'aide de la fonction indicatrice $\mathbf{1}_A$ de l'événement A, sous la forme :

$$X_{n+1} = X_n \circ \pi + \mathbf{1}_A$$

Donc :

$$\mathbf{E}(X_{n+1}) = \mathbf{E}(X_n \circ \pi) + \mathbf{E}(\mathbf{1}_A)$$

Mais $\mathbf{E}(X_n \circ \pi) = \mathbf{E}(X_n)$ car :

$$\mathbf{E}(X_n \circ \pi) = \frac{1}{(n + 1)!} \sum_{\varphi} (X_n \circ \pi)(\varphi)$$ où la somme porte sur tous les éléments $\varphi \in \mathfrak{S}_{n+1}$

$$= \frac{1}{(n + 1)!} \sum_{\sigma} (n + 1) X_n(\sigma)$$ où la somme porte sur tous les éléments $\sigma \in \mathfrak{S}_n$ et qu'il y a $n + 1$ permutations φ de $\mathfrak{S}_{n+1}$ donnant la même permutation $\sigma = \pi(\varphi)$ de $\mathfrak{S}_n$.

$$= \frac{1}{n!} \sum_{\sigma} X_n(\sigma) = \mathbf{E}(X_n)$$

Donc :

$$\mathbf{E}(X_{n+1}) = \mathbf{E}(X_n) + \mathbf{P}(A) = \mathbf{E}(X_n) + \frac{1}{n+1} \qquad \text{avec } \mathbf{E}(X_1) = 1$$

d'où $\quad \mathbf{E}(X_n) = \sum_{k=1}^{n} \frac{1}{k}$

b) En élevant au carré la relation $X_{n+1} = X_n \circ \pi + \mathbf{1}_A$, on obtient :

$$X_{n+1}^2 = (X_n \circ \pi)^2 + 2(X_n \circ \pi)\mathbf{1}_A + \mathbf{1}_A$$

donc $\quad \mathbf{E}(X_{n+1}^2 \mid A) = \mathbf{E}((X_n \circ \pi)^2 \mid A) + \mathbf{E}(2(X_n \circ \pi)\mathbf{1}_A \mid A) + \mathbf{E}(\mathbf{1}_A \mid A)$

$\quad = \mathbf{E}((X_n \circ \pi)^2 \mid A) + 2\mathbf{E}(X_n \circ \pi \mid A) + 1$

et $\quad \mathbf{E}(X_{n+1}^2 \mid A^c) = \mathbf{E}((X_n \circ \pi)^2 \mid A^c) + \mathbf{E}(2(X_n \circ \pi)\mathbf{1}_A \mid A^c) + \mathbf{E}(\mathbf{1}_A \mid A^c)$

$\quad = \mathbf{E}((X_n \circ \pi)^2 \mid A^c)$

donc $\quad \mathbf{E}(X_{n+1}^2) = \mathbf{P}(A)(\mathbf{E}((X_n \circ \pi)^2 \mid A) + 2\mathbf{E}(X_n \circ \pi \mid A) + 1) + \mathbf{P}(A^c)\mathbf{E}((X_n \circ \pi)^2 \mid A^c)$

$\quad = \mathbf{P}(A)\mathbf{E}((X_n \circ \pi)^2 \mid A) + \mathbf{P}(A^c)\mathbf{E}((X_n \circ \pi)^2 \mid A^c) + 2\mathbf{P}(A)\mathbf{E}(X_n \circ \pi \mid A) + \mathbf{P}(A)$

$\quad = \mathbf{E}((X_n \circ \pi)^2) + 2\mathbf{P}(A)\mathbf{E}(X_n \circ \pi \mid A) + \mathbf{P}(A)$

Comme dans le a), on vérifiera que $\mathbf{E}((X_n \circ \pi)^2) = \mathbf{E}(X_n^2)$ et que $\mathbf{E}(X_n \circ \pi \mid A) = \mathbf{E}(X_n)$. Donc :

$$\mathbf{E}(X_{n+1}^2) = \mathbf{E}(X_n^2) + 2\mathbf{P}(A)\mathbf{E}(X_n) + \mathbf{P}(A)$$

donc $\quad \mathbf{V}(X_{n+1}) = \mathbf{E}(X_{n+1}^2) - \mathbf{E}(X_{n+1})^2$

$\quad = \mathbf{E}(X_n^2) + 2\mathbf{P}(A)\mathbf{E}(X_n) + \mathbf{P}(A) - (\mathbf{E}(X_n) + \mathbf{P}(A))^2$

$\quad = \mathbf{E}(X_n^2) + 2\mathbf{P}(A)\mathbf{E}(X_n) + \mathbf{P}(A) - \mathbf{E}(X_n)^2 - 2\mathbf{P}(A)\mathbf{E}(X_n) - \mathbf{P}(A)^2$

$\quad = \mathbf{V}(X_n) + \mathbf{P}(A)(1 - \mathbf{P}(A))$

$\quad = \mathbf{V}(X_n) + \frac{n}{(n+1)^2} \qquad \text{avec } \mathbf{V}(X_1) = 0$

d'où $\quad \mathbf{V}(X_n) = \sum_{k=1}^{n-1} \frac{k}{(k+1)^2}$

❑ On peut aussi vérifier que, pour $1 \le k \le n+1$:

$\mathbf{P}(X_{n+1} = k) = \mathbf{P}(X_n \circ \pi + \mathbf{1}_A = k)$

$\quad = \mathbf{P}(X_n \circ \pi + \mathbf{1}_A = k \mid A)\mathbf{P}(A) + \mathbf{P}(X_n \circ \pi + \mathbf{1}_A = k \mid A^c)\mathbf{P}(A^c)$

$\quad = \mathbf{P}(X_n \circ \pi = k - 1 \mid A)\mathbf{P}(A) + \mathbf{P}(X_n \circ \pi = k \mid A^c)\mathbf{P}(A^c)$

$\quad = \mathbf{P}(X_n = k - 1)\mathbf{P}(A) + \mathbf{P}(X_n = k)\mathbf{P}(A^c)$

$\quad = \frac{1}{n+1}\mathbf{P}(X_n = k - 1) + \frac{n}{n+1}\mathbf{P}(X_n = k)$

ce qui permet de calculer récursivement la loi de X_n.

Sol.19) a) $\frac{1}{n}$

b) $\quad \mathbf{P}(E_k) = \frac{1}{n}$

$\mathbf{P}(F_k) = \frac{m}{k}$

$\mathbf{P}(E_k \cap F_{k-1}) = \mathbf{P}(F_{k-1} \mid E_k)\,\mathbf{P}(E_k) = \mathbf{P}(F_{k-1})\,\mathbf{P}(E_k) = \frac{m}{n(k-1)}$

car le fait de savoir que la carte portant le plus grand nombre se trouve en k ne modifie en rien la probabilité que la carte ayant le nombre maximal parmi celles de rang 1 à $k - 1$ se trouve en telle ou telle position. Ainsi, les événement E_k et F_{k-1} sont indépendants.
On peut aussi dénombrer le nombre de cas favorables réalisant l'événement $E_k \cap F_{k-1}$:

placer la carte portant le plus grand nombre (1 choix) au rang k,

choisir $k - 1$ cartes parmi les $n - 1$ qui restent ($\binom{n-1}{k-1}$ choix),

choisir l'emplacement entre 1 et m qui accueillera la carte ayant le nombre maximal parmi celles de rang 1 à $k - 1$ (m choix possibles),

placer la carte ayant le nombre maximal parmi le paquet des $k - 1$ cartes sélectionnées à l'emplacement précédemment choisi, et placer les $k - 2$ autres cartes au hasard aux places restantes ($(k - 2)!$ choix),

placer les $n - k$ cartes non utilisées au hasard après la k-ème ($(n - k)!$ choix).

Cela donne $\binom{n-1}{k-1} \times m \times (k - 2)! \times (n - k)! = \dfrac{(n - 1)!\, m}{(k - 1)}$ cas favorables. D'où la probabilité cherchée en divisant par $n!$.

$$\mathbf{P}(G \mid E_k) = \begin{cases} \mathbf{P}(F_{k-1} \mid E_k) = \mathbf{P}(F_{k-1}) = \dfrac{m}{k-1} \text{ si } k > m \\ 0 \text{ sinon} \end{cases}$$

$$\mathbf{P}(G) = \sum_{k=m+1}^{n} \mathbf{P}(G \mid E_k)\, \mathbf{P}(E_k) = \sum_{k=m+1}^{n} \frac{m}{n(k-1)} = \frac{m}{n} \sum_{k=m}^{n-1} \frac{1}{k}$$

On peut donner un équivalent de $\sum_{k=m}^{n-1} \frac{1}{k}$ au moyen d'encadrement par des intégrales :

$$\sum_{k=m}^{n-1} \int_k^{k+1} \frac{1}{t}\, dt \leq \sum_{k=m}^{n-1} \frac{1}{k} \leq \sum_{k=m}^{n-1} \int_{k-1}^{k} \frac{1}{t}\, dt$$

$$\Leftrightarrow \quad \int_m^n \frac{1}{t}\, dt \leq \sum_{k=m}^{n-1} \frac{1}{k} \leq \int_{m-1}^{n-1} \frac{1}{t}\, dt$$

$$\Leftrightarrow \quad \ln\left(\frac{n}{m}\right) \leq \sum_{k=m}^{n-1} \frac{1}{k} \leq \ln\left(\frac{n-1}{m-1}\right)$$

donc $\sum_{k=m}^{n-1} \frac{1}{k} \to -\ln(\alpha)$ et $\mathbf{P}(G) \sim -\alpha\ln(\alpha)$. Cette quantité est maximale pour $\alpha = \frac{1}{e}$ et $\mathbf{P}(G) \sim \frac{1}{e}$

Cet exercice est parfois présenté de façon assez cynique sous le vocable *problème du recrutement*. On a n candidats pour un poste, de valeurs différentes. On passe en revue m d'entre eux, avec $m \approx \frac{n}{e}$. On prend le premier candidat qui suit les m premiers et qui est meilleur que ceux-ci, ou le dernier si le meilleur de tous était dans les m premiers. La probabilité de recruter ainsi le meilleur de tous est de l'ordre de $\frac{1}{e}$.

Sol.20) a) La loi de X est binomiale $\mathcal{B}(n, p)$. $\mathbf{E}(X) = np$

b) La loi de Y, conditionnée par l'événement X = k, est binomiale $\mathcal{B}(n-k, q)$.

$$\mathbf{E}(Y \mid X = k) = (n-k)q$$

c) Pour $0 \le k \le n$ et $0 \le r \le n-k$:

$$\mathbf{P}(X = k, Y = r) = \mathbf{P}(X = k)\,\mathbf{P}(Y = r \mid X = k) = \binom{n}{k} p^k(1-p)^{n-k} \binom{n-k}{r} q^r(1-q)^{n-k-r}$$

$$= \frac{n!}{k!r!(n-k-r)!} p^k(1-p)^{n-k}\, q^r(1-q)^{n-k-r}$$

$$= \binom{n}{r}\binom{n-r}{k} p^k(1-p)^{n-k}\, q^r(1-q)^{n-k-r}$$

d) $$\mathbf{P}(Y = r) = \sum_{k=0}^{n-r} \binom{n}{r}\binom{n-r}{k} p^k(1-p)^{n-k}\, q^r(1-q)^{n-k-r}$$

$$= \binom{n}{r} (1-p)^r\, q^r \sum_{k=0}^{n-r} \binom{n-r}{k} p^k(1-p)^{n-r-k}(1-q)^{n-k-r}$$

$$= \binom{n}{r} (1-p)^r\, q^r\, (p + (1-p)(1-q))^{n-r}$$

$$= \binom{n}{r} (q-pq)^r\, (1-q+qp)^{n-r}$$

Il s'agit d'une loi binomiale $\mathcal{B}(n, q-qp)$, ce qui peut s'expliquer directement comme suit. Pour être comptabilisé dans Y, le client doit subir un échec lors du premier appel (probabilité $1-p$) et un succès lors du second appel (probabilité q), soit une probabilité totale de $(1-p)q$. On effectue alors n expériences de Bernoulli sur chaque client, avec une probabilité de succès $(1-p)q$.

$$\mathbf{E}(Y) = nq(1-p)$$

et $$\sum_{k=0}^{n} \mathbf{E}(Y \mid X = k)\mathbf{P}(X = k) = \sum_{k=0}^{n} (n-k)q \binom{n}{k} p^k (1-p)^{n-k}$$

$$= nq(1-p) \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

$$= nq(1-p)\,(p+1-p)^{n-1} = nq(1-p) = \mathbf{E}(Y)$$

e) $$\mathbf{P}(X + Y = s) = \sum_{k=0}^{s} \mathbf{P}(X + Y = s \mid X = k)\,\mathbf{P}(X = k)$$

$$= \sum_{k=0}^{s} \mathbf{P}(Y = s-k \mid X = k)\,\mathbf{P}(X = k)$$

$$= \sum_{k=0}^{s} \binom{n-k}{s-k} q^{s-k}\,(1-q)^{n-s} \binom{n}{k} p^k\,(1-p)^{n-k}$$

$$= \sum_{k=0}^{s} \frac{n!}{k!(s-k)!(n-s)!} q^{s-k}\,(1-q)^{n-s}\, p^k\,(1-p)^{n-k}$$

$$= \binom{n}{s} \sum_{k=0}^{s} \binom{s}{k} \; q^{s-k}\,(1-q)^{n-s}\,p^k\,(1-p)^{n-k}$$

$$= \binom{n}{s} (1-q)^{n-s}\,(1-p)^{n-s} \sum_{k=0}^{s} \binom{s}{k} \; q^{s-k}\,p^k\,(1-p)^{s-k}$$

$$= \binom{n}{s} (1-q)^{n-s}\,(1-p)^{n-s}\,(p + q(1-p))^s$$

$$= \binom{n}{s} (1-q)^{n-s}\,(1-p)^{n-s}\,(p + q - qp)^s$$

Il s'agit d'une loi binomiale $\mathcal{B}(n, 1-(1-p)(1-q))$. On peut la trouver directement en observant que la probabilité que chaque client contacté réponde à l'un des deux appels est $1-(1-p)(1-q)$.

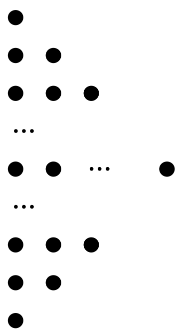
$$\mathbf{E}(X+Y) = n(1-(1-p)(1-q)) = n(p+q-pq)$$

$$= \mathbf{E}(X) + \mathbf{E}(Y) \qquad \text{en cohérence avec les questions précédentes.}$$

f) $X+Y+Z$ suit une loi $\mathcal{B}(n, 1-(1-p)(1-q)(1-r))$. Son espérance est :

$$n(1-(1-p)(1-q)(1-r))$$

Sol.21) a) Ω est constitué de $(n+1)^2$ points.



Pour $0 \leq i \leq n$, $\mathbf{P}(X = i) = \dfrac{2n+1-2i}{(n+1)^2}$. $2n+1-2i$ est le nombre de points sur la colonne i.

Pour $-n \leq j \leq n$, $\mathbf{P}(Y = j) = \dfrac{n+1-|j|}{(n+1)^2}$. $n+1-|j|$ est le nombre de points sur la ligne j.

$$\mathbf{E}(X) = \sum_{i=0}^{n} i\,\frac{2n+1-2i}{(n+1)^2} = \frac{(2n+1)n}{2(n+1)} - \frac{n(2n+1)}{3(n+1)} = \frac{n(2n+1)}{6(n+1)} \sim \frac{n}{3}$$

$$\mathbf{E}(Y) = \sum_{j=-n}^{n} j\,\frac{n+1-|j|}{(n+1)^2} = 0$$

$$\mathbf{E}(|Y|) = \sum_{j=-n}^{n} |j|\,\frac{n+1-|j|}{(n+1)^2} = 2\sum_{j=1}^{n} j\,\frac{n+1-j}{(n+1)^2} = n - \frac{n(2n+1)}{3(n+1)} = \frac{n^2+2n}{3(n+1)} \sim \frac{n}{3}$$

b) $\quad \mathbf{P}(X = i \mid Y = j) = \dfrac{\mathbf{P}(X = i, Y = j)}{\mathbf{P}(Y = j)} = \dfrac{1}{n+1-|j|}$, loi uniforme sur $[\![0, n-|j|]\!]$

$$\mathbf{E}(\mathrm{X} \mid \mathrm{Y} = j) = \sum_{i=0}^{n-|j|} \frac{i}{n+1-|j|} = \frac{n-|j|}{2}$$

$$\sum_{j=-n}^{n} \mathbf{E}(\mathrm{X} \mid \mathrm{Y} = j)\mathbf{P}(\mathrm{Y} = j) = \sum_{j=-n}^{n} \frac{n-|j|}{2} \frac{n+1-|j|}{(n+1)^2}$$

$$= \frac{n}{2}\frac{1}{n+1} + \frac{2}{(n+1)^2}\sum_{j=1}^{n}\binom{n+1-j}{2} = \frac{n}{2}\frac{1}{n+1} + \frac{2}{(n+1)^2}\binom{n+1}{3}$$

$$= \frac{n}{2}\frac{1}{n+1} + \frac{n(n-1)}{3(n+1)} = \frac{n(2n+1)}{6(n+1)} = \mathbf{E}(\mathrm{X})$$

c) $f(\mathrm{Y})(i, j) = f(\mathrm{Y}(i, j)) = f(j) = \mathbf{E}(\mathrm{X} \mid \mathrm{Y} = j) = \dfrac{n-|j|}{2}$ donc $f(\mathrm{Y}) = \dfrac{n-|\mathrm{Y}|}{2}$ et son espérance est :

$$\mathbf{E}(f(\mathrm{Y})) = \frac{n}{2} - \frac{1}{2}\mathbf{E}(|\mathrm{Y}|) = \frac{n}{2} - \frac{n^2+2n}{6(n+1)} = \frac{2n^2+n}{6(n+1)} = \mathbf{E}(\mathrm{X})$$

Pour montrer que $\mathbf{E}((\mathrm{X} - f(\mathrm{Y}))g(\mathrm{Y})) = 0$, linéarité, puisque $g(\mathrm{Y}) = \sum\limits_{j} g(j)\, \mathbf{1}_{\mathrm{Y}=j}$, il suffit de le montrer pour toute fonction $g(\mathrm{Y})$ de la forme $\mathbf{1}_{\mathrm{Y}=j}$. Pour une telle fonction :

$$\begin{aligned}
\mathbf{E}((\mathrm{X} - f(\mathrm{Y}))g(\mathrm{Y})) &= \mathbf{E}((\mathrm{X} - f(\mathrm{Y}))\, \mathbf{1}_{\mathrm{Y}=j}) \\
&= \mathbf{E}((\mathrm{X} - f(j))\, \mathbf{1}_{\mathrm{Y}=j}) \\
&= \mathbf{E}(\mathrm{X}\, \mathbf{1}_{\mathrm{Y}=j}) - f(j)\, \mathbf{P}(\mathrm{Y} = j) \\
&= \sum_{k=-n}^{n} \mathbf{E}(\mathrm{X}\, \mathbf{1}_{\mathrm{Y}=j} \mid \mathrm{Y} = k)\, \mathbf{P}(\mathrm{Y} = k) - f(j)\, \mathbf{P}(\mathrm{Y} = j) \\
&= \mathbf{E}(\mathrm{X}\, \mathbf{1}_{\mathrm{Y}=j} \mid \mathrm{Y} = j)\, \mathbf{P}(\mathrm{Y} = j) - f(j)\, \mathbf{P}(\mathrm{Y} = j) \\
&= \mathbf{E}(\mathrm{X} \mid \mathrm{Y} = j)\, \mathbf{P}(\mathrm{Y} = j) - f(j)\, \mathbf{P}(\mathrm{Y} = j) \\
&= 0
\end{aligned}$$

Enfin :

$\quad \mathbf{V}(\mathrm{X} - f(\mathrm{Y})) \le \mathbf{V}(\mathrm{X} - g(\mathrm{Y}))$

$\Leftrightarrow \quad \mathbf{E}((\mathrm{X} - f(\mathrm{Y}))^2) \le \mathbf{E}((\mathrm{X} - g(\mathrm{Y}))^2) - (\mathbf{E}(\mathrm{X} - g(\mathrm{Y})))^2 \quad$ car $(\mathbf{E}(\mathrm{X} - f(\mathrm{Y})))^2 = 0$

$\Leftrightarrow \quad \mathbf{E}(f(\mathrm{Y})^2 - g(\mathrm{Y})^2 - 2\mathrm{X}f(\mathrm{Y}) + 2\mathrm{X}g(\mathrm{Y})) \le -(\mathbf{E}(\mathrm{X} - g(\mathrm{Y})))^2$

$\Leftrightarrow \quad \mathbf{E}(f(\mathrm{Y})^2 - g(\mathrm{Y})^2) + \mathbf{E}(\mathrm{X}(-2f(\mathrm{Y}) + 2g(\mathrm{Y}))) \le -(\mathbf{E}(\mathrm{X} - g(\mathrm{Y})))^2$

or $\mathbf{E}(\mathrm{X}(-2f(\mathrm{Y}) + 2g(\mathrm{Y}))) = \mathbf{E}(f(\mathrm{Y})(-2f(\mathrm{Y}) + 2g(\mathrm{Y})))$ d'après la question précédente, appliquée avec $-2f(\mathrm{Y}) + 2g(\mathrm{Y})$ au lieu de $g(\mathrm{Y})$. On obtient :

$\Leftrightarrow \quad \mathbf{E}(f(\mathrm{Y})^2 - g(\mathrm{Y})^2) + \mathbf{E}(-2f(\mathrm{Y})^2 + 2g(\mathrm{Y})f(\mathrm{Y}))) \le -(\mathbf{E}(\mathrm{X} - g(\mathrm{Y})))^2$

$\Leftrightarrow \quad \mathbf{E}(-f(\mathrm{Y})^2 - g(\mathrm{Y})^2 + 2g(\mathrm{Y})f(\mathrm{Y}))) \le -(\mathbf{E}(\mathrm{X} - g(\mathrm{Y})))^2$

$\Leftrightarrow \quad \mathbf{E}(-(f(\mathrm{Y}) - g(\mathrm{Y}))^2) \le -(\mathbf{E}(\mathrm{X} - g(\mathrm{Y})))^2$

$\Leftrightarrow \quad \mathbf{E}((f(\mathrm{Y}) - g(\mathrm{Y}))^2) \ge (\mathbf{E}(f(\mathrm{Y}) - g(\mathrm{Y})))^2 \qquad$ car $\mathbf{E}(\mathrm{X}) = \mathbf{E}(f(\mathrm{Y}))$

$\Leftrightarrow \quad \mathbf{V}(f(\mathrm{Y}) - g(\mathrm{Y})) \ge 0$ ce qui est vrai.

La fonction $f(\mathrm{Y})$ est notée usuellement $\mathbf{E}(\mathrm{X} \mid \mathrm{Y})$.

Sol.22) En raisonnant sur les variables $\mathrm{X}_i - \mu$ et $\mathrm{Z} - \mu$, on se ramène au cas où toutes les variables sont centrées. On a alors, pour tout i :

$$\mathbf{E}((X_i - Z)^2) = \mathbf{E}((-\frac{X_1}{n} - \ldots - \frac{X_{i-1}}{n} + \frac{n-1}{n} X_i - \frac{X_{i+1}}{n} - \ldots - \frac{X_n}{n})^2)$$
$$= \mathbf{V}(-\frac{X_1}{n} - \ldots - \frac{X_{i-1}}{n} + \frac{n-1}{n} X_i - \frac{X_{i+1}}{n} - \ldots - \frac{X_n}{n})$$
$$= \mathbf{V}(\frac{X_1}{n}) + \ldots \mathbf{V}(\frac{X_{i-1}}{n}) + \mathbf{V}(\frac{n-1}{n} X_i) + \mathbf{V}(\frac{X_{i+1}}{n}) + \ldots + \mathbf{V}(\frac{X_n}{n})$$

car les variables sont indépendantes

$$= \frac{n-1}{n^2} \sigma^2 + \frac{(n-1)^2}{n^2} \sigma^2 = \frac{n-1}{n} \sigma^2 \qquad \text{indépendant de } i$$

donc $\frac{1}{n-1} \sum_{i=1}^{n} \mathbf{E}((X_i - Z)^2) = \sigma^2$.

Certaines calculatrices, dotées de fonctions statistiques, proposent deux manières de calculer l'écart-type σ d'une liste $(x_1, \ldots, x_n)$ de valeurs. Ayant calculé $m = \frac{1}{n} \sum_{i=1}^{n} x_i$, elles proposent de prendre :

ou bien $\sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - m)^2}$

ou bien $\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2}$

Le premier cas s'applique si on dispose d'une liste complète des données. Si une variable aléatoire X prend les valeurs $x_1, \ldots, x_n$ de façon équiprobable, m n'est autre que $\mathbf{E}(X)$, $\frac{1}{n} \sum_{i=1}^{n} (x_i - m)^2$ vaut $\mathbf{V}(X)$ et sa racine carrée est bien l'écart-type σ.
Mais si on ne dispose que d'un échantillon d'une liste de données dont l'étendue est inconnue, chaque valeur x_i est considérée comme un tirage aléatoire parmi cette liste, correspondant à la valeur prise par une variable aléatoire X_i. m n'est autre que la valeur prise par la variable aléatoire Z, et $\frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2$ est la valeur prise par la variable aléatoire $\frac{1}{n-1} \sum_{i=1}^{n} (X_i - Z)^2$. C'est seulement une estimation de σ^2, au sens un tirage de n autres valeurs donnerait un résultat différent, avec cependant en moyenne la valeur de σ^2 cherchée. L'utilisation de la formule $\frac{1}{n} \sum_{i=1}^{n} (X_i - Z)^2$ donnerait un résultat légèrement sous-estimé.

Sol.23) a) La loi de X_i est une loi binomiale $\mathcal{B}(n, \frac{1}{3})$. Comme $X_1 + X_2 = n - X_3$, on a :

$$\mathbf{P}(X_1 + X_2 = k) = \mathbf{P}(X_3 = n - k) = \binom{n}{n-k} \frac{1}{3^{n-k}} (\frac{2}{3})^k = \binom{n}{k} \frac{1}{3^{n-k}} (\frac{2}{3})^k$$

qui est une loi binomiale $\mathcal{B}(n, \frac{2}{3})$. On peut aussi dire que $X_1 + X_2$ est le résultat de n expériences indépendantes de Bernoulli où le succès consiste à entrer dans les pièces 1 ou 2, avec une probabilité $\frac{2}{3}$. En particulier :

$$\mathbf{V}(X_i) = n \times \frac{1}{3} \times \frac{2}{3} = \frac{2n}{9} = \mathbf{V}(X_1 + X_2)$$

Comme $\mathbf{V}(X_1 + X_2) = \mathbf{V}(X_1) + \mathbf{V}(X_2) + 2\text{cov}(X_1, X_2)$, on en déduit :

$$\text{cov}(X_1, X_2) = -\frac{n}{9}$$

et $$\rho(X_1, X_2) = -\frac{\frac{n}{9}}{\frac{2n}{9}} = -\frac{1}{2}$$

b) X_1 et X_2 suivent des lois binomiales $\mathcal{B}(n, \frac{1}{6})$, donc $\mathbf{E}(X_1) = \mathbf{E}(X_2) = \frac{n}{6}$, et $\mathbf{V}(X_1) = \mathbf{V}(X_2) = \frac{5n}{36}$. Par ailleurs, la loi de X_1 sachant que $X_2 = k$ est une loi binomiale $\mathcal{B}(n - k, \frac{1}{5})$, donc :

$$\mathbf{E}(X_1X_2 \mid X_2 = k) = \mathbf{E}(X_1k \mid X_2 = k) = k\,\mathbf{E}(X_1 \mid X_2 = k) = \frac{k(n-k)}{5}$$

et
$$\begin{aligned}
\mathbf{E}(X_1X_2) &= \sum_{k=0}^{n} \mathbf{E}(X_1X_2 \mid X_2 = k)\,\mathbf{P}(X_2 = k) \\
&= \sum_{k=0}^{n} \frac{k(n-k)}{5}\,\mathbf{P}(X_2 = k) \\
&= \mathbf{E}\left(\frac{X_2(n - X_2)}{5}\right) \\
&= \frac{1}{5}\,(n\mathbf{E}(X_2) - \mathbf{E}(X_2^2)) \\
&= \frac{1}{5}\,(n\mathbf{E}(X_2) - \mathbf{V}(X_2) - \mathbf{E}(X_2)^2) \\
&= \frac{1}{5}\left(\frac{n^2}{6} - \frac{5n}{36} - \frac{n^2}{36}\right) \\
&= \frac{n^2 - n}{36}
\end{aligned}$$

donc $$\text{cov}(X_1, X_2) = \mathbf{E}(X_1X_2) - \mathbf{E}(X_1)\mathbf{E}(X_2) = -\frac{n}{36}$$

On peut aussi calculer $\mathbf{E}(X_1X_2)$ à partir des formules :

$$\mathbf{P}(X_1 = k, X_2 = l) = \binom{n}{k}\binom{n-k}{l}\frac{1}{6^{k+l}}\left(\frac{4}{6}\right)^{n-k-l}$$

et $$\mathbf{E}(X_1X_2) = \sum_{k,l} \binom{n}{k}\binom{n-k}{l}\,kl\,\frac{1}{6^{k+l}}\left(\frac{4}{6}\right)^{n-k-l}$$

mais c'est assez fastidieux.

On peut aussi dire que $X_1 + X_2$ suit une loi binomiale $\mathcal{B}(n, \frac{2}{6})$ (on répète n expériences indépendantes de Bernoulli où le succès consiste à tirer 1 ou 2, avec une probabilité $\frac{2}{6}$), donc :

$$\mathrm{cov}(X_1, X_2) = \frac{1}{2}(\mathbf{V}(X_1 + X_2) - \mathbf{V}(X_1) - \mathbf{V}(X_2)) = \frac{1}{2}(\frac{2n}{9} - \frac{5n}{18}) = -\frac{n}{36}$$

Finalement, on en déduit que :

$$\rho(X_1, X_2) = -\frac{\frac{n}{36}}{\frac{5n}{36}} = -\frac{1}{5}$$

c) On a :

$$0 \leq \mathbf{V}(X_1 + ... + X_m) = \sum_{i=1}^{m} \mathbf{V}(X_i) + \sum_{i\neq j} \mathrm{cov}(X_i, X_j) = m\sigma^2 + m(m-1)\rho\sigma^2$$

d'où le résultat.

d) Le a) est une application directe du c) avec $m = 3$. On a précisément $\rho = -\frac{1}{2} = -\frac{1}{m-1}$ dans ce cas, car $X_1 + X_2 + X_3$ est la variable certaine n dont la variance est nulle, et le calcul de c) donne alors une égalité.

Le b) relève aussi du c) en notant X_i le nombre de fois où le chiffre i est sorti, $1 \leq i \leq 6$. On a donc ici $m = 6$. On retrouve l'égalité $\rho = -\frac{1}{5}$ en remarquant que, là aussi, $\sum_{i=1}^{6} X_i = n$ de variance nulle, donc on a également une égalité dans le calcul du c).

Sol.24) a) pour k élément de $[\![0, n]\!]$:

$$\mathbf{P}(Y = k) = \sum_{m=k}^{N-n+k} \mathbf{P}(Y = k \mid X = m)\, \mathbf{P}(X = m)$$

$$= \sum_{m=k}^{N-n+k} \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} \binom{N}{m} p^m (1-p)^{N-m}$$

$$= \sum_{m=k}^{N-n+k} \frac{m!\,(N-m)!\,N!\,n!\,(N-n)!}{k!\,(m-k)!\,(n-k)!\,(N-m-n+k)!\,m!\,(N-m)!\,N!}\; p^m (1-p)^{N-m}$$

$$= \sum_{m=k}^{N-n+k} \frac{n!\,(N-n)!}{k!\,(m-k)!\,(n-k)!\,(N-m-n+k)!}\; p^m (1-p)^{N-m}$$

$$= \binom{n}{k} \sum_{m=k}^{N-n+k} \binom{N-n}{m-k} p^m (1-p)^{N-m}$$

$$= \binom{n}{k} \sum_{m=0}^{N-n} \binom{N-n}{m} p^{m+k} (1-p)^{N-m-k}$$

en changeant d'indice $m - k \to m$

$$= \binom{n}{k} p^k (1-p)^{n-k} \sum_{m=0}^{N-n} \binom{N-n}{m} p^m (1-p)^{N-n-m}$$

on reconnaît un binôme de Newton

$$= \binom{n}{k} p^k (1-p)^{n-k} (p + 1 - p)^{N-n} = \binom{n}{k} p^k (1-p)^{n-k}$$

Y suit une loi $\mathcal{B}(n, p)$.

b) On utilise la formule Bayes :

$$\mathbf{P}(X = m + k \mid Y = k) = \frac{\mathbf{P}(Y = k \mid X = m + k)\ \mathbf{P}(X = m + k)}{\mathbf{P}(Y = k)}$$

$$= \frac{\binom{m+k}{k}\binom{N-m-k}{n-k}}{\binom{N}{n}} \frac{\binom{N}{m+k} p^{m+k} (1-p)^{N-m-k}}{\binom{n}{k} p^k (1-p)^{n-k}}$$

$$= \frac{(m+k)!\ (N-m-k)!\ N!\ n!\ (N-n)!\ k!\ (n-k)!}{k!\ m!\ (n-k)!\ (N-m-n)!\ (m+k)!\ (N-m-k)!\ N!\ n!} p^m (1-p)^{N-m-n}$$

$$= \frac{(N-n)!}{m!\ (N-m-n)!} p^m (1-p)^{N-m-n}$$

$$= \binom{N-n}{m} p^m (1-p)^{N-m-n}$$

c) X est le nombre d'objets marqués. Y est le nombre d'objets marqués parmi ceux choisis. Si on choisit d'abord les objets, puis qu'on les marque ensuite, on voit qu'on obtient bien comme loi de Y une loi $\mathcal{B}(n, p)$.

Sachant qu'il y a k objets marqués parmi les n choisis, la probabilité qu'il y en ait en tout $m + k$ marqués est la probabilité qu'il y ait m marqués parmi les $N - n$ non choisis, d'où la loi $\mathcal{B}(N - n, p)$.

◆