IV. Variables aléatoires

Aujourd'hui nous terminons la partie consacrée aux probabilités conditionnelles en parlant d'indépendance entre différents événements, puis nous étudions les variables aléatoires, ce qui nous permet de passer finalement aux statistiques dans le dernier chapitre du module.

1 Indépendance de deux événements

Nous avons vu que parfois la probabilité d'un événement est affectée par la réalisation d'un autre événement. Par exemple, le fait de tirer un as dans un paquet de cartes rend la probabilité d'en tirer un autre ensuite plus faible. Par contre, obtenir un six lors d'un lancer de dé n'affecte pas la probabilité d'en obtenir un autre au lancer suivant.

Définition 1.1. Les événements E et F sont *indépendants* si $P(E \cap F) = P(E)P(F)$.

En d'autres termes, on a

$$P(E|F) = \frac{P(E\cap F)}{P(F)} \overset{\text{déf. 1.1}}{=} \frac{P(E)\cdot P(F)}{P(F)} = P(E)$$

De même, $P(F|E) = P(F)$

L'information partielle que F s'est réalisé n'influence pas la réalisation de E et vice-versa.

Exemple 1.2. Voici les résultats simplifiés d'une enquête sur la législation de la marijuana aux USA. On interroge des habitants sur tout le territoire des Etats-Unis. On compte 30% de la population sur la côte Est. La proportion de personnes sondées pour la dépénalisation et habitant sur la côte Est est de 7,8%. Il y a 18,2% de personnes en faveur de la dépénalisation et habitant dans les autres régions.

Les événements D : "être en faveur de la dépénalisation" et E : "vivre sur la côte Est" sont-ils dépendants ?

$$P(D|E) = \frac{P(D\cap E)}{P(E)} = \frac{0,078}{0,3} = 0,26$$

et $P(D) = P(D\cap E) + P(D\cap E^c) = 0,078 + 0,182 = 0,26$

Les deux événements sont indépendants. On a autant de chance de tomber sur quelqu'un qui est pour la dépénalisation et qui sur la côte Est ou ailleurs.

	E	E^c	Total
D	7,8%	18,2%	26%
D^c	22,2%	51,8%	74%
Total	30%	70%	100%

Il y a indépendance entre les événements E et D

$\iff$

Les colonnes, respectivement les lignes du tableau sont proportionnelles.

Que se passe-t-il lorsque l'on veut parler de l'indépendance de plus de deux événements ? Il faut faire très attention avant de tirer des conclusions hâtives.
En effet, si A et B sont indépendants et que B et C le sont aussi, que dire de A et C ?

A et C ne sont en général pas indépendants !

Par exemple si $A = C$, A n'est pas indépendant de lui-même !

$$\frac{P(\emptyset)}{P(A^c)} = 0$$

$$P(A|A^c) = \frac{P(A\cap A^c)}{P(A^c)} = \frac{P(A)}{P(A)} = 1$$

Par ailleurs, si A et B sont indépendants et que A et C le sont aussi, alors A et $B\cap C$ ... ne sont en général pas indépendants !

Exemple 1.3. On dispose d'un gros dé à six faces qu'on colorie de la manière suivante.
Les faces 1, 2, 3 et 4 sont blanches, les faces 5 et 6 sont bleues ;
les nombres 1, 2, 3 et 6 sont écrits à l'encre rouge, les nombres 4 et 5 à l'encre noire.
On considère les événements A : "obtenir un nombre pair", R : "obtenir un nombre rouge" et B : "obtenir une face blanche".

- A et B sont indépendants : $P(A) = \frac{1}{2}$, $P(B) = \frac{2}{3}$ et $P(A \cap B) = \frac{1}{3}$
  $\Rightarrow P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3} = P(A \cap B)$
- A et R sont indépendants : idem
- A et $B \cap R$ ne sont pas indépendants : $P(B \cap R) = \frac{1}{2}$
  $P(A \cap (B \cap R)) = \frac{1}{6}$. Or, $P(A) \cdot P(B \cap R) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \neq \frac{1}{6}$.
  (2 seul possible)

Un autre exemple surprenant (pour toi aussi je l'espère) sera analysé dans la série d'exercices. Ceci nous pousse à définir la notion d'indépendance totale.

Définition 1.4. Les événements E, F et G sont *totalement indépendants* s'ils sont indépendants deux à deux et de plus $P(E \cap F \cap G) = P(E)P(F)P(G)$.

Explicitement on demande donc que

$P(E \cap F) =$ $P(E) \cdot P(F)$

$P(E \cap G) =$ $P(E) \cdot P(G)$

$P(F \cap G) =$ $P(F) \cdot P(G)$

et $P(E \cap F \cap G) =$ $P(E) \cdot P(F) \cdot P(G)$

Exemple 1.5. *ou plutôt un contre-exemple :*
On lance deux pièces de monnaie et on considère les événements
A = "pile au 1er lancé", B = "face au 2ème lancé", C = "même résultat au deux lancés".

$P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$, $P(C) = \frac{1}{2}$; $P(A \cap B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$, $P(A \cap C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
$P(B \cap C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. mais $P(A \cap B \cap C) = 0 \neq P(A) \cdot P(B) \cdot P(C)$.

2 Variables aléatoires

Après avoir réalisé une expérience, il arrive qu'on s'intéresse plus à une fonction du résultat qu'au résultat lui-même. Par exemple, lors d'une élection au Conseil fédéral, peut-être que certains voudront seulement savoir combien de femmes ont été élues. S'il s'agit des élèves sélectionnés au cours Euler, ce sera le nombre d'élèves Genevois, ou alors le nombre d'élèves en 11ème année, etc.

Définition 2.1. Toute fonction réelle $X : S \to \mathbb{R}$ définie sur l'ensemble fondamental S d'une expérience est appelée *variable aléatoire.*

Exemple 2.2. On jette trois pièces de monnaie équilibrées et on définit X par le nombre de piles obtenus. Ainsi X est une variable aléatoire pouvant prendre les valeurs 0, 1, 2 ou 3. La probabilité de chacun de ces événements est :

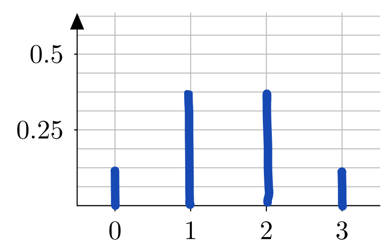
$$P\{X = 0\} = \frac{1}{8} = \left(\frac{1}{2}\right)^3 \quad (FFF)$$

$$P\{X = 1\} = \frac{3}{8} \quad (PFF, FPF, FFP)$$

$$P\{X = 2\} = \frac{3}{8}$$

$$P\{X = 3\} = \frac{1}{8} \quad (PPP)$$

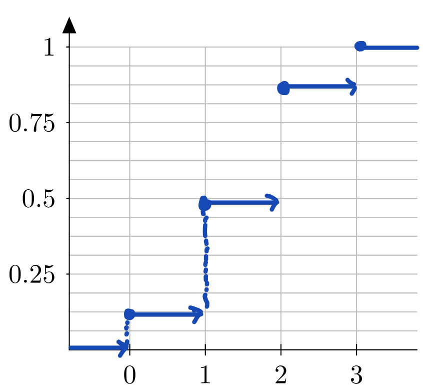
Représentation graphique de la distribution de X



Définition 2.3. La *fonction de répartition* d'une variable aléatoire X est la fonction réelle

$$F(b) = P\{X \leq b\}.$$

La fonction de répartition de l'exemple précédent se représente graphiquement comme suit :



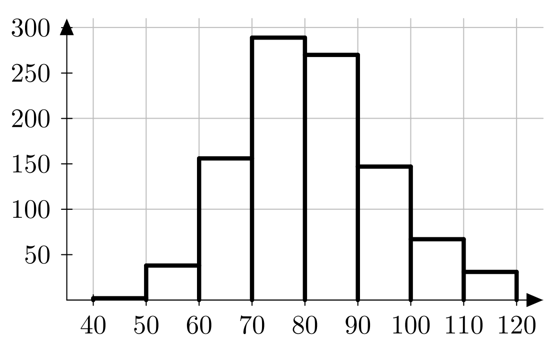
Nous étudierons principalement des variables aléatoires discrètes, qui ne prennent au plus qu'une quantité dénombrable de valeurs, mais il en existe d'autres, dites continues. Par exemple, les mesure, taille ou durée sont vues comme des grandeurs réelles plutôt qu'entières ou rationnelles.

Les concepts de fonction de distribution et de répartition se retrouvent aussi dans le domaine de la statistique comme le montre l'exemple ci-dessous.

Exemple 2.4. Le tableau suivant donne les résultats d'une étude du poids d'Australiens.
Le poids étant une mesure de type continu, les données ont été regroupés en classes

Poids X en kg	*Effectif*	*Pourcentage*	*Pourcentage cumulé*
$X < 50$	2	0.2%	0,2%
$50 \leqslant X < 60$	38	3.8%	4%
$60 \leqslant X < 70$	156	15.6%	19,6%
$70 \leqslant X < 80$	289	28.9%	48,5%
$80 \leqslant X < 90$	270	27.0%	75,5%
$90 \leqslant X < 100$	147	14.7%	90,2%
$100 \leqslant X < 110$	67	6.7%	96,9%
$110 \leqslant X$	31	3.1%	100%
Total	1000	100%	

Histogramme de distribution des poids.

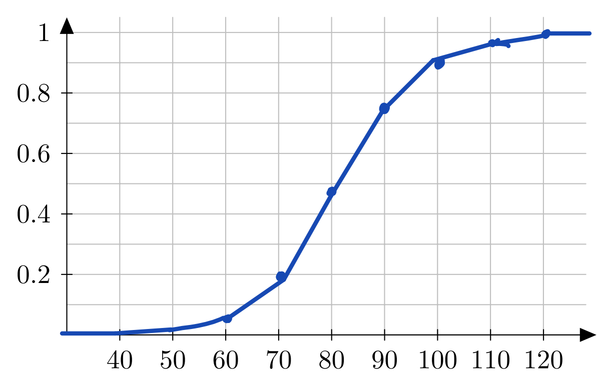


On peut alors considérer X comme une variable aléatoire continue.

Si on tire une personne au hasard dans l'échantillon, quelle est la probabilité que son poids soit compris entre 60 et 70 kg ?

En calculant les pourcentages cumulés des personnes ayant un poids inférieur à x kilos, on peut tracer la fonction de répartition de X.

Si on tire une personne au hasard dans l'échantillon, quelle est la probabilité son poids soit inférieur à 80 kg ?



Proposition 2.5. *Soit X une variable aléatoire et F sa fonction de répartition. Alors*

$$P\{a < X \leq b\} = F(b) - F(a) \text{ si } a < b.$$

Démonstration. Décrivons l'événement $\{X \leq b\}$ comme réunion disjointe des deux événements $\{X \leq a\}$ et $\{a < X \leq b\}$. Ainsi

$$F(b) = P\{X \leq b\} = P\{X \leq a\} + P\{a < X \leq b\}$$
$$= F(a) + P\{a < X \leq b\}$$

d'où $F(b) - F(a) = P\{a < X \leq b\}$

□

Exemple 2.6. La roue de la fortune
On fait tourner une triple roue qui s'immobilise en laissant apparaître trois nombres compris entre 1 et 6. Le joueur mise un franc et choisit un nombre. Il gagne k fois sa mise si le nombre apparaît k fois. Il perd sa mise si le nombre n'apparaît pas.
Ce jeu est-il équitable ? Etudions la variable aléatoire X = "gain net du joueur".

$X \in \{-1; 0, 1; 2\}$

$$P\{X = -1\} = \left(\frac{5}{6}\right)^3 = \frac{125}{216}$$
$$P\{X = 0\} = C_1^3 \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6} = \frac{75}{216}$$
$$P\{X = 1\} = C_2^3 \left(\frac{5}{6}\right) \cdot \left(\frac{1}{6}\right)^2 = \frac{15}{216}$$
$$P\{X = 2\} = \left(\frac{1}{6}\right)^3 = \frac{1}{216}$$
$$P(0 \leq X \leq 2) = \frac{91}{216}$$

3 Espérance

L'espérance d'une variable aléatoire est la moyenne pondérée des valeurs possibles de cette variable. Elle mesure la valeur qu'on peut espérer obtenir en "moyenne" si on la répète un grand nombre de fois l'expérience aléatoire.

Définition 3.1. Soit X une variable aléatoire discrète. L'*espérance* de X est le nombre réel

$$E[X] = \sum_{a \in \mathbb{R}} a\, P\{X = a\}.$$

Ceci est bien défini puisque la variable aléatoire est discrète. Il n'y a qu'un nombre fini de nombres réels a pour lesquels $P\{X = a\} \neq 0$.

Exemple 3.2. Quelle est l'espérance du résultat du lancer d'un dé bien équilibré ?

Chaque valeur a probabilité $\frac{1}{6}$ d'apparaître, $S = \{1; 2; 3; 4; 5; 6\}$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = 3{,}5$$

Dans l'exemple 2.6 $E[X] = -1 \cdot \frac{125}{216} + 0 \cdot \frac{75}{216} + 1 \cdot \frac{15}{216} + 2 \cdot \frac{1}{216} = \frac{-108}{216} = -\frac{1}{2}$

En moyenne, on perd 50 centimes par partie à ce jeu.

En "moyenne" on obtient donc $3,5$, ce qui veut dire qu'on peut espérer obtenir 350 points si l'on additionne les résultats de 100 lancers.

Il arrive aussi qu'on cherche à connaître l'espérance non pas d'une variable aléatoire X, mais d'une fonction de celle-ci.

Exemple 3.3. On choisit au hasard un mot dans la phrase "TU VAS PAYER". Tu dois deviner en combien de lettres s'écrit le mot choisi et tu devras payer le carré de la différence entre la longueur devinée et la vraie longueur. Ainsi si tu devines 5, tu ne devra rien payer si le mot "PAYER" est choisi, mais $(5-2)^2 = 9$ francs si le mot "TU" est choisi. Que faut-il choisir pour payer le moins possible ?

Soit X la variable aléatoire donnant la différence entre la longueur devinée et la longueur choisie. Si t est la longueur choisie, tu as donc une chance sur trois de devoir payer $(t-2)^2$, une sur trois de payer $(t-3)^2$ et enfin une sur trois de devoir payer $(t-5)^2$. L'espérance est donc

$$E[X^2] = \frac{1}{3}(t-2)^2 + \frac{1}{3}(t-3)^2 + \frac{1}{3}(t-5)^2$$

$$= \frac{1}{3}\left(t^2 - 4t + 4 + t^2 - 6t + 9 + t^2 - 10t + 25\right)$$

$$= t^2 - \frac{20}{3}t + \frac{38}{3} = \left(t - \frac{10}{3}\right)^2 - \left(\frac{10}{3}\right)^2 + \frac{38}{3} = \left(t - \frac{10}{3}\right)^2 + \frac{14}{9}$$

complétion du carré

Il faut donc choisir $\frac{10}{3}$ pour espérer payer seulement $1,56$ francs en moyenne.

Théorème 3.4. *Soit X une variable aléatoire discrète et $g(x)$ une fonction réelle. Alors*

$$E[g(X)] = \sum_{a \in \mathbb{R}} g(a) P\{X = a\}.$$

Démonstration. Il s'agit en fait simplement de la définition de l'espérance appliquée à la variable aléatoire $g(X)$. Il faut remarquer néanmoins que $P\{g(X) = b\} = \sum_{a | g(a) = b} P\{X = a\}$. □

D'autre part l'espérance est linéaire :

Théorème 3.5. *Soient X, Y deux variables aléatoires discrètes. Alors $E[X + Y] = E[X] + E[Y]$.*

Démonstration. On vérifie cela en calculant

$$\begin{aligned}
E[X+Y] &= \sum_c cP\{X+Y=c\} \\
&= \sum_c c\left(\sum_a P\{X=a, Y=c-a\}\right) \\
&= \sum_c \sum_a (a+(c-a))P\{X=a, Y=c-a\} \\
&= \sum_c \sum_a aP\{X=a, Y=c-a\} + \sum_c \sum_a (c-a)P\{X=a, Y=c-a\}
\end{aligned}$$

On calcule la première partie comme suit :

$$\begin{aligned}
\sum_c \sum_a a \cdot P\{X=a, Y=c-a\} &= \sum_a a \cdot P\left\{\coprod_c (\{X=a\} \cap \{Y=c-a\})\right\} \\
&= \sum_a a \cdot P\{X=a\} = E[X]
\end{aligned}$$

De même, la deuxième partie devient, avec le changement de variable $b = c - a$:

$$\begin{aligned}
\sum_b \sum_a b \cdot P\{X=a, Y=b\} &= \sum_b b \cdot P\left\{\coprod_a (\{X=a\} \cap \{Y=b\})\right\} \\
&= \sum_b b \cdot P\{Y=b\} = E[Y]
\end{aligned}$$

Ainsi, on obtient bien que $E[X+Y] = E[X] + E[Y]$. □

Exemple 3.6. Un problème de Daniel Bernoulli.
Daniel Bernoulli (1700–1782) est le fils de Jean et le neveu de Jacques Bernoulli, l'un des mathématiciens les plus illustres de cette famille bâloise. Ami d'Euler, il travailla avec lui à Saint-Petersbourg pendant plusieurs années, avant de s'installer à Bâle.

Le problème est le suivant. Une urne contient $2N$ cartes numérotées de 1 à N par paire. On tire m cartes au hasard. Quel est le nombre moyen de paires de cartes encore présentes dans l'urne ?

On définit X_i la variable aléatoire qui vaut 1 si la i-ème paire est intacte, et zéro sinon. Quelle est l'espérance de X_i ?

$$E[X_i] = 1 \cdot P\{X_i = 1\} + 0 \cdot P\{X_i = 0\} = 1 \cdot \frac{\binom{2N-2}{m}}{\binom{2N}{m}}$$

$$= \frac{\dfrac{\cancel{(2N-2)!}}{\cancel{m!}\,\cancel{(2N-2-m)!}}}{\dfrac{\cancel{(2N)!}\ 2N\cdot(2N-1)}{\cancel{m!}\,\cancel{(2N-m)!}(2N-m)(2N-m-1)}} = \frac{(2N-m)(2N-m-1)}{2N\cdot(2N-1)}$$

Le but est de trouver l'espérance de la somme de toutes les variables X_i.
Le théorème précédent affirme que cette espérance vaut

$$E\left[\sum_{i=1}^{N} X_i\right] \underset{\text{linéarité de l'espérance}}{=} \sum_{i=1}^{N} E[X_i] = N \cdot E[X_i] = \frac{(2N-m)(2N-m-1)}{2\,(2N-1)}$$

Ce modèle fut proposé pour déterminer combien il reste de couples après la mort de m personnes... Par exemple, s'il y a 100 couples au départ et que 50 personnes meurent, on peut espérer trouver encore $\frac{150 \cdot 149}{2 \cdot 199} \cong 56$ couples "indemmes" après ce coup du sort.