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 pour les derniers chapitres 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, la probabilité de lancer un six aux dés n'affecte pas le lancer suivant.

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

En d'autres termes, on a

$$P(E|F) = \frac{P(E\cap F)}{P(F)} \overset{(*)}{=} \frac{P(E)\cdot \cancel{P(F)}}{\cancel{P(F)}} = P(E)$$

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 d'une enquête (simplifiée) 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 habitant dans les autres régions. Parmi celles qui sont contre, $22,2\%$ vivent sur la côte Est et les $51,8\%$ restant habitent dans les autres régions.

Que dire des événements D "être en faveur de la dépénalisation" et E "vivre sur la côte Est" ?

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

On a $P(D|E) = \dfrac{P(D\cap E)}{P(E)} = \dfrac{7{,}8\%}{30\%} = 0{,}26 = P(D)$

$\Rightarrow$ les deux évènements sont indépendants

Lorsque les variables sont indépendants les lignes, respectivement les colonnes 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 B et C le sont aussi, que dire de A et C ?

Il est faux de croire que A et C sont alors forcément indépendants.

Si on choisit $C = A$, A n'est pas indépendent de lui-même.

$$P(A|A) = 1 \neq P(A) \text{ en général.}$$

Plus sérieuse est la constatation suivante : Si A et B sont indépendants et A et C le sont aussi, alors A et $B \cap C$ ... ne sont en général pas indépendants !

Exemple où les variables ne sont PAS indépendantes !

	F	¬F	
G	5	45	50
¬G	25	25	50
	30	70	100

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 l'événement A "obtenir un nombre pair", R "obtenir un nombre rouge" et B "obtenir une face blanche".

- A et B sont indépendants car $P(A) = \frac{1}{2}$, $P(B) = \frac{2}{3}$ et $P(A \cap B) = \frac{2}{6} = \frac{1}{3} = \frac{1}{2} \cdot \frac{2}{3} = P(A) \cdot P(B)$.
- De même, A et R sont indépendants : $P(R) = \frac{2}{3}$
  $P(A \cap R) = \frac{2}{6} = \frac{1}{3} = \frac{1}{2} \cdot \frac{2}{3} = P(A) \cdot P(R)$
- Mais A et $B \cap R$ ne sont pas indépendants :
  $P(A \cap (B \cap R)) = \frac{1}{6} \neq P(A) \cdot P(B \cap R) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
  (seul. le chiffre 2 convient)

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

1. $P(E \cap F) = P(E) \cdot P(F)$
2. $P(E \cap G) = P(E) \cdot P(G)$
3. $P(F \cap G) = P(F) \cdot P(G)$
4. $P(E \cap F \cap G) = P(E) \cdot P(F) \cdot P(G)$

Exemple : les nombres obtenu à chaque lancé d'un dé.

Contre exemple : on lance 2 pièces de monnaie successivement.

A = "pile au 1^er lancé", B = "face au 2^e lancé"

C = "même résultat sur les deux lancés"

$P(A) = \frac{1}{2}$; $P(B) = \frac{1}{2}$; $P(C) = \frac{1}{2}$; $P(A \cap B) = \frac{1}{4}$, $P(B \cap C) = \frac{1}{4}$

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

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. S'il s'agit d'une élection au Conseil fédéral par exemple, 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.

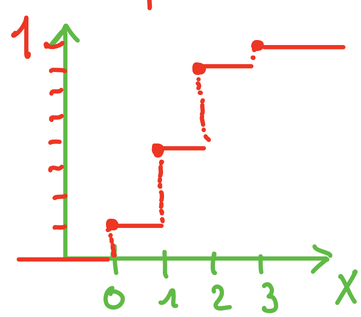
Définition 2.1. Toute fonction réelle $X : S \to \mathbb{R}$ définie sur l'ensemble fondamental S d'une expérience est appelée *variable aléatoire*.

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 :

$X =$ nb de piles sur 3 lancés

1. $P\{X = 0\} = \frac{1}{8}$ $\longrightarrow$ FFF
2. $P\{X = 1\} = \frac{3}{8}$ $\longrightarrow$ PFF FPF FFP
3. $P\{X = 2\} = \frac{3}{8}$ $\longrightarrow$ FPP PFP PPF
4. $P\{X = 3\} = \frac{1}{8}$ $\longrightarrow$ PPP

Graphe de F



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

Dans l'exemple 2.2

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

$F(2) = P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$

Nous étudierons principalement des variables aléatoires discrètes (qui ne peut prendre au plus qu'une quantité dénombrable de valeurs), mais il en existe d'autres, dites continues. Cela pourrait être par exemple le poids d'une personne où il serait plus agréable de considérer le poids comme un nombre réel plutôt que rationnel.

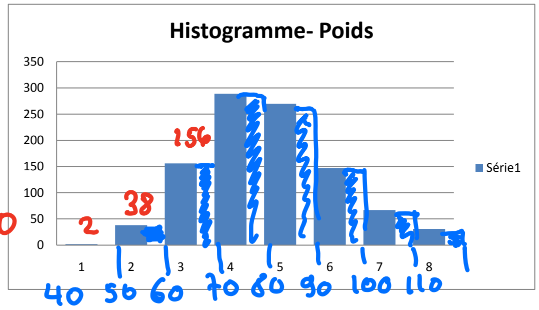
Exemple 2.4. Une étude de 1995 mesure le poids d'un adulte en Australie. On considère ici le poids comme une variable discrète et ne s'intéresse qu'aux dizaines de kilos. Sur 1000 hommes pesés, on constate qu'il y en a 2 de moins de 50 kilos, 38 entre 50 et 60, 156 entre 60 et 70, 289 entre 70 et 80, 270 entre 80 et 90, 147 entre 90 et 100, 67 entre 100 et 110 et enfin 31 de plus de 110 kilos. On peut représenter cela par un *histogramme*, où l'on indique la valeur de la variable sur l'axe horizontal et le nombre de personnes concernées est reporté sur l'axe vertical :

En fait, on a regroupé les poids par classe $[40;50[$, $[50;60[$ etc.

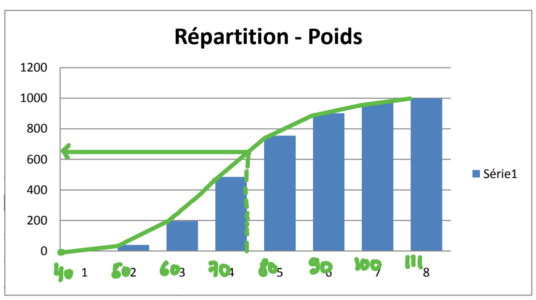
$F(70) = P(X \leq 70)$

$= \dfrac{2+38+156}{1000} = \dfrac{196}{1000}$

$= 0{,}196 = 19{,}6\,\%$



Une mise à l'échelle donne la probabilité. Par exemple $P\{70 < X \leq 80\} = 0,289$. Si l'on somme le nombre de personnes ayant un poids inférieur à x kilos, on obtient la fonction de répartition de cette variable



$F(75) \cong \dfrac{630}{1000} = 63\,\%$

Attention ! Il ne s'agit pas vraiment d'une variable aléatoire ici, mais nous établissons le comportement probable de cette variable grâce à une mesure statistique. Ce sera notre prochain sujet et cet exemple tente de montrer le lien étroit entre les probabilités et les statistiques.

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

$$F(b) \overset{\text{déf}}{=} P\{X \leq b\} = P\{X \leq a\} + P\{a < X \leq b\}$$
$$\overset{\text{déf}}{=} F(a) + P\{a < X \leq b\}$$
$$\Rightarrow P\{a < X \leq b\} = F(b) - F(a)$$

□

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. Il y a donc $6^3 = 216$ issues possibles. 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 honnête ?

Nous travaillons avec la variable aléatoire X qui mesure les gains nets du joueur.

X prends les valeurs -1 ; 0 ; 1 ; 2

- $P\{X = 2\} = \frac{1}{6^3} = \frac{1}{216}$
- $P\{X = 1\} = \frac{3 \cdot 5}{6^3} = \frac{15}{216}$ (2 roues sur 3 affichent le nb choisi)
- $P\{X = 0\} = \frac{3 \cdot 5 \cdot 5}{6^3} = \frac{75}{216}$

On perd sa mise dans tous les autres cas, soit avec une probabilité $\frac{216 - 1 - 15 - 75}{216} = \frac{125}{216}$

3 Espérance

L'espérance d'une variable aléatoire est la moyenne pondérée des valeurs possibles de cette variable. Elle mesure donc ce qu'on peut espérer obtenir en "moyenne" lors d'une expérience, si on la répète un grand nombre de fois.
Par exemple, si l'on lance une pièce mal équilibrée qui a deux chances sur trois de donner pile ($X = 0$) et une seule chance sur trois de donner face ($X = 1$), alors l'espérance de X vaut $1/3$.

Exemple 2.6 : $E[X] = 2 \cdot \frac{1}{216} + 1 \cdot \frac{15}{216} + 0 \cdot \frac{75}{216} - 1 \cdot \frac{125}{216}$
$= \frac{-108}{216} = -0,5$

A ce jeu, on perd en moyenne 50 cts par partie...

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 la probabilité d'apparition de $\frac{1}{6}$.

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6\,\frac{1}{6}$$
$$= \frac{7}{2} = 3,5$$

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

$$E[X^2] = \frac{1}{3}(t-2)^2 + \frac{1}{3}(t-3)^2 + \frac{1}{3}(t-5)^2$$
$$= t^2 - \frac{20}{3}t + \frac{38}{3} = \left(t - \frac{10}{3}\right)^2 - \left(\frac{10}{3}\right)^2 + \frac{38}{3}$$
$$= \left(t - \frac{10}{3}\right)^2 + \frac{14}{9}$$

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

$$\frac{14}{9} = 1,56$$

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

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 $\sum_c \sum_a aP\{X = a, Y = c - a\}$ comme suit :

$$\sum_a a \cdot P\left\{\coprod_c \{X=a\} \cap \{Y=c-a\}\right\} = \sum_a a P\{X=a\} = E[X]$$

+ réunion disjointe

c prend toutes les valeurs possibles telles que $P\{Y=c-a\} > 0$ $\Rightarrow \left\{\coprod_c \{X=a\} \cap \{Y=c-a\}\right\} = \{X=a\}$

De même, la deuxième partie $\sum_c \sum_a (c - a) P\{X = a, Y = c - a\}$ devient, en posant $b = c-a$

$$\sum_b \sum_a b P\{X=a ; Y=b\} = \sum_b b P\left\{\coprod_a \{X=a\} \cap \{Y=b\}\right\}$$

$$= \sum_b b P\{Y=b\} = E[Y]$$

idem ci-dessus

□

Ainsi, $E[X+Y] = E[X] + E[Y]$

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 illsustres 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 ? Ce modèle fut proposé pour déterminer combien il reste de couples mariés après la mort de m personnes...

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 ?

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

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

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

$$E\left[\sum_{i=1}^{N} X_i\right] = \sum_{i=1}^{N} E[X_i] = N \cdot E[X_i] = \frac{(2N-m)(2N-m-1)}{2\,(2N-1)}$$

espérance est linéaire

toutes les $E[x_i]$ sont égales.

Ainsi, s'il y a 100 couples au départ et que 50 personnes meurent, on peut espérer avoir encore $\dfrac{150 \cdot 149}{2(199)} \cong 56$, c'est-à-dire environ 56 couples mariés...