Étude des événements et variables aléatoires

Axel Péneau

19 mai 2022

1 Événements aléatoire et Tribut

En probabilités on étudie des événements que l'on appelle observables. On se fixe un univers Ω et un ensemble $\mathcal{A}$ d'observables, c'est-à-dire des parties de Ω. On demande à ce que l'ensemble $\mathcal{A}$ vérifie les propriétés suivantes. L'idée derrière la notion d'observable peut être illustrée par une manche de poker Texas hold'em par exemple. On commence à distribuer deux cartes à chaque joueur puis on pose cinq cartes sur la table en en mettant d'abord trois cartes face visible puis une autre puis une autre. L'univers Ω est l'ensemble des arrangements du paquet au début de la partie (il y en a $52! = 52 \times 51 \times \cdots \times 3 \times 2$). En effet un joueur qui connaîtrait l'ordre des cartes dans le paquet saurait exactement qui va gagner la manche. Sauf que en début de partie le joueur ne voit que ses deux cartes, puis après le flop il ne vois que ses deux cartes et les trois sur la table, ensuite la 4è carte puis la 5è carte et à la fin de la manche il voit le jeu de ses adversaires et peut enfin déterminer si il a gagné ou non.

Proposition 1.1 (Négation). *Si $A \subset \Omega$ est observable (c'est-à-dire $A \in \mathcal{A}$) alors son complémentaire $\Omega \backslash A$ (qu'on note $\bar{A}$ au lycée et parfois A^c) est aussi observable.*

Proposition 1.2 (Combinaison finie). *Si $A \subset \Omega$ et $B \subset \Omega$ sont observables alors leur union $A \cup B$ et leur intersection $A \cap B$ sont aussi observables.*

L'intuition derrière ces deux propriétés est assez simple, si je peux observer un événement comme "j'ai un as dans ma main" alors je peux observer le contraire "je n'ai pas d'as dans ma main". Si je peux observer l'événement "j'ai un roi" et "j'ai un as" alors je peux observer leur intersection : "j'ai un roi et un as" et leur union "j'ai un roi ou un as". En Théorie des jeux on aime aussi étudier des jeux potentiellement infinis, par exemple un tournoi de poker n'a pas de nombre de manches fixé donc on a une infinité de tournois de poker possible (alors qu'on avait que 52! manches possibles). admettons que 8 joueurs misent un jeton à chaque manche quelque soit leur jeu (ce qui est une très mauvaise stratégie). On aimerait par exemple pouvoir décrire l'événement "un des joueurs finit ruiné", c'est en fait l'union des événements $A_n =$"un des joueurs est ruiné à la n-ème manche" pour n allant de 1 à l'infini.

Proposition 1.3 (σ-intersection et σ-union). *Si $A_1 \subset A_2 \subset \ldots$ est une suite croissante d'événements observables alors leur limite :*

$$A = \bigcup_{n=1}^{\infty} A_n = \{\omega \in \Omega | \omega \in A_n \text{ à partir d'un certain indice } n\}$$

est aussi un événement observable.
De manière équivalente, si $A_1 \supset A_2 \supset \ldots$ est une suite décroissante d'événements observables alors leur intersection est elle aussi observable.

Remarquez que si on a une suite pas forcément décroissante $A_1, A_2, \dots$ alors on peut toujours voir son intersection comme une intersection décroissante car $A_1 \supset A_1 \cap A_2 \supset A_1 \cap A_2 \cap A_3 \supset \dots$ et pareil pour les unions. Intuitivement l'inclusion $A \subset B$ se traduit comme l'implication "si A se réalise alors B se réalise".

Maintenant on veut associer une probabilité $\mathbb{P}$ qui est un nombre entre 0 et 1 à chacun des événements observables. On demande à $\mathbb{P}$ de vérifier les propriétés suivantes.

Proposition 1.4 (Croissance). *Si A et B sont des événements observables et que $A \subset B$ alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.*

Proposition 1.5 (Additivité). *Si A et B sont des événements observables alors $\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$.*

Proposition 1.6 (σ-additivité). *Si $A_1 \supset A_2 \supset \dots$ est une suite décroissante d'événements observables, $\mathbb{P}(\bigcap A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$. De manière équivalente si $A_1 \subset A_2 \subset \dots$ est une suite croissante d'événements observables alors $\mathbb{P}(\bigcup A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$*

La question qu'il est légitime de se poser est bien sûr celle de l'existence des limites. C'est pour ce problème d'existence qu'on impose la condition de monotonie (croissante ou décroissante) sur la suite (A_n). En effet si $A_1 \supset A_2 \supset \dots$ alors d'après la propriété 1.4 de croissance, on a $\mathbb{P}(A_1) \geqslant \mathbb{P}(A_2) \geqslant \dots$. La suite $(\mathbb{P}(A_n))$ est donc une suite monotone et bornée (car $0 \leqslant \mathbb{P}(A_n) \leqslant 1$). Le théorème de la borne inférieure sur $\mathbb{R}$ nous dit que toute suite monotone bornée converge. Mais je ne met pas ce théorème en exercice pour une raison assez simple : il ne se démontre pas ! Il fait en fait partie des axiomes de $\mathbb{R}$, c'est à dire qu'on définit $\mathbb{R}$ comme un ensemble qui vérifie la propriété de la borne inférieure (ainsi que d'autres propriétés qui ne nous intéressent pas en probabilité).

On peut pourtant trouver des preuves du théorème de la borne inférieure et c'est pour une raison assez simple et très importante en maths : on a plusieurs définitions de $\mathbb{R}$ et on doit démontrer que ces définitions sont toutes équivalentes.

Exercice 1.7 (σ-sous-additivité). *On donne une suite $A_1, A_2, \dots$ d'événements mesurables pas forcément décroissante ni croissante, montrez que si la limite existe alors :*

$$\mathbb{P}\left(\bigcap A_n\right) \leqslant \lim_{n\to\infty} \mathbb{P}(A_n) \leqslant \mathbb{P}\left(\bigcup A_n\right).$$

Correction : On définit deux suites $I_n = A_1 \cap \dots \cap A_n$ (c'est l'intersection partielle de la suite (A_n)) et $U_n = A_1 \cap \dots \cap A_n$ (c'est l'union partielle de la suite (A_n)). On a alors pour tout n l'inclusion $I_n \subset A_n \subset U_n$. Donc d'après la propriété de croissance on a $\mathbb{P}(I_n) \leqslant \mathbb{P}(A_n) \leqslant \mathbb{P}(U_n)$. En plus on a $I_1 \supset I_2 \supset \dots$ c'est-à-dire que (I_n) est une suite croissante donc par σ-additivité, on a $\lim_{n\to\infty} \mathbb{P}(I_n) = \mathbb{P}(\bigcap I_n)$. De même (U_n) est une suite croissante donc par σ-additivité $\lim_{n\to\infty} \mathbb{P}(I_n) = \mathbb{P}(\bigcup U_n)$. Il suffit maintenant d'utiliser le théorème d'encadrement des limites (qu'on appelle théorème des gendarmes au lycée) et qui nous dit :

$$\lim_{n\to\infty} \mathbb{P}(I_n) \leqslant \lim_{n\to\infty} \mathbb{P}(A_n) \leqslant \lim_{n\to\infty} \mathbb{P}(U_n).$$

On remplace ensuite les termes de gauche et de droite par leur valeur et on trouve :

$$\mathbb{P}\left(\bigcap I_n\right) = \mathbb{P}\left(\bigcap A_n\right) \leqslant \lim_{n\to\infty} \mathbb{P}(A_n) \leqslant \mathbb{P}\left(\bigcup A_n\right) = \mathbb{P}\left(\bigcup U_n\right).$$

$\square$

2 Variables aléatoires

Maintenant qu'on a défini des événements aléatoires, on aimerait bien définir les variables aléatoires, c'est simplement un nombre qui dépend d'un événement aléatoire. Par exemple la gain en jetons d'un joueur qui suivrait une stratégie donnée après une manche de Poker. On se place dans le contexte qu'on vient de définir, on a un univers Ω (c'est-à-dire un ensemble) un ensemble $\mathcal{A}$ d'événements observables et une loi probabilité $\mathbb{P}$ (c'est à dire une fonction de $\mathcal{A}$ dans $[0, 1]$).

Définition 2.1 (Variable aléatoire réelle)**.** Une variable aléatoire réelle est une fonction $X : \Omega \to \mathbb{R}$ (c'est-à-dire que pour tout $\omega \in \Omega$ on a un réel $X(\omega) \in \mathbb{R}$). On demande en plus à ce que pour tout intervalle $[a, b]$ l'événement $(X \in [a, b]) = \{\omega \in \Omega, a \leqslant X(\omega) \leqslant b\}$ soit mesurable.

En probabilité on aime étudier des variables aléatoires indépendantes, par exemple les tirages des cartes à deux manches différentes sont indépendants car entre temps on a récupéré toutes les cartes et mélangé le paquet.

Définition 2.2 (Variables aléatoires indépendantes)**.** On dit que deux événements observables A et B sont indépendants si :

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

On peut interpréter cette définition avec la formule de Bayes qui nous définit la probabilité d'un événement A sachant qu'un autre événement B est réalisé. On note ce nombre $\mathbb{P}(A|B)$ et on le définit comme :

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

Dire que A et B sont indépendants revient donc à dire que $\mathbb{P}(A|B) = \mathbb{P}(A)$.

On va maintenant définir l'espérance d'une variable aléatoire, l'intuition est de dire que l'espérance de X est la valeur "moyenne" de X. On a par exemple envie que l'espérance soit additive.

Définition 2.3. On se donne X une variable aléatoire réelle positive, on définit :

$$\mathbb{E}(X) = \int_{t=0}^{\infty} \mathbb{P}(X \geqslant t) dt$$

On rappelle la définition de l'intégrale de Riemann :

$$\int_{t=0}^{\infty} \mathbb{P}(X \geqslant t) dt = \lim_{n \to \infty} \sum_{k=1}^{\infty} \frac{1}{2^n} \mathbb{P}\left(X \geqslant \frac{k}{2^n}\right) = \lim_{n \to \infty} \lim_{m \to \infty} \sum_{k=1}^{m} \frac{1}{2^n} \mathbb{P}\left(X \geqslant \frac{k}{2^n}\right)$$

Remarque. Ici les limites existent car $\lim_{m \to \infty} \sum_{k=1}^{m} \frac{1}{n} \mathbb{P}\left(X \geqslant \frac{k}{n}\right)$ est une limite croissante donc elle est bien définie mais elle peut parfois prendre la valeur $+\infty$. Et comm la fonction $t \mapsto \mathbb{P}(X \geqslant t)$ est décroissante, $\lim_{n \to \infty} \sum_{k=1}^{\infty} \frac{1}{2^n} \mathbb{P}\left(X \geqslant \frac{k}{2^n}\right)$ est elle aussi une limite croissante. Dessine une courbe décroissante puis dessine un histogramme sous une courbe décroissante puis dessine un autre histogramme avec des barres deux fois plus fines. L'histogramme le plus fin sera plus haut que l'autre.

Theorem 2.4 (Inégalité de Markov)**.** *Si X est une variable aléatoire réelle positive et C un nombre réel positif alors :*

$$\mathbb{P}(X \geqslant C) \leqslant \frac{\mathbb{E}(X)}{C}.$$

Démonstration. On décompose la formule de l'espérance :

$$\mathbb{E}(X) = \int_{t=0}^{\infty} \mathbb{P}(X \geqslant t)dt = \int_{t=0}^{C} \mathbb{P}(t \leqslant X \leqslant C) + \mathbb{P}(X \geqslant C)dt + \int_{t=C}^{\infty} \mathbb{P}(X \geqslant t)dt.$$

Et $\int_{t=0}^{C} \mathbb{P}(t \leqslant X \leqslant C)dt + \int_{t=C}^{\infty} \mathbb{P}(X \geqslant t)dt$ est un nombre positif donc

$$\mathbb{E}(X) \geqslant \int_{t=0}^{C} \mathbb{P}(X \geqslant C)dt = C\mathbb{P}(X \geqslant C).$$

□

Si X est à valeurs entières, on a :

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

Pour démontrer ça, on commence par remarquer que $\mathbb{P}(X \geqslant k) = \sum_{j=k}^{\infty} \mathbb{P}(X = j)$. et on ré indexe la somme

$$\sum_{k=1}^{\infty} \mathbb{P}(X \geqslant k) = \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} \mathbb{P}(X = j) = \sum_{j=1}^{\infty} \sum_{k=1}^{j} \mathbb{P}(X = j) = \sum_{j=1}^{\infty} j\mathbb{P}(X = j).$$

Définition 2.5. On dit que deux variables aléatoires réelles X et Y sont indépendantes. Si pour tous intervalles $[a, b]$ et $[c, d]$, les événements $(a \leqslant X \leqslant b)$ et $(c \leqslant Y \leqslant d)$ sont indépendants.

Exercice 2.6. *Si X est une variable aléatoire à valeurs entières, alors $\mathbb{E}(nX) = n\mathbb{E}(X)$.*

Correction : On utilise la distributivité de la somme :

$$\mathbb{E}(nX) = \sum_{k=1}^{\infty} \mathbb{P}(nX \geqslant k) = \sum_{k=1}^{\infty} \mathbb{P}(X \geqslant k/n) = \sum_{j=1}^{\infty} \sum_{k=nj-n+1}^{nj} \mathbb{P}(X \geqslant j) = n \sum_{j=1}^{\infty} \mathbb{P}(X \geqslant j)$$

□

Exercice 2.7. *On se donne deux variables aléatoires entières positives X et Y. Montrez que si X et Y sont indépendantes alors :*

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

Correction : On écrit :

$$\mathbb{E}(X \times Y) = \sum_{k=1}^{\infty} \mathbb{P}(X \times Y \geqslant k).$$

Puis on décompose $(X \times Y \geqslant k) = \bigsqcup_{n=1}^{\infty} (X = n \text{ et } Y \geqslant k/n)$ (le symbole $\bigsqcup$ signifie que l'union est disjointe). On a alors par la σ-additivité et par indépendance :

$$\mathbb{E}(X \times Y) = \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \mathbb{P}(X = n \text{ et } Y \geqslant k/n) = \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \mathbb{P}(X = n)\mathbb{P}(Y \geqslant k/n).$$

On réordonne la somme et on trouve :

$$\mathbb{E}(X \times Y) = \sum_{n=1}^{\infty} \mathbb{P}(X = n) \sum_{k=0}^{\infty} \mathbb{P}(Y \geqslant k/n) = \sum_{n=1}^{\infty} n\mathbb{P}(X = n)\mathbb{E}(Y) = \mathbb{E}(X) \times \mathbb{E}(Y).$$

□

En approchant les variables aléatoires à valeur entière, on peut montrer la même égalité pour les variables aléatoires réelles.

Exercice 2.8. *Tu joues au pile ou face contre Elon Musk. À chaque manche si tu fais face tu gagnes et Elon te donne 2\$ et si tu fais pile tu perds et tu dois 1\$ à Elon. Tu commences avec 1\$ et si tu tombes à 0 la partie s'arrête. Est-il possible que tu arrives à plumer Elon Musk ? Donnes une borne inférieure sur la probabilité que ça arrive. Avec Python, calcules à 10^{-16} près la probabilité de l'événement "ruiner Elon".*

Remarque. La précision 10^{-16} demandée est simplement la précision de calcul de Python qui code les nombres réels sur 64-bits.

Indice. On définit une variable aléatoire G_n qui représente ton gain en dollars à l'étape n. Donner la loi de G_n.

On définit $S_n = 1 + G_1 + \cdots + G_n$. On se donne un nombre réel $0 < \alpha < 1$, En utilisant le fait que les (G_n) sont indépendants, donner l'espérance de α^{S_n}. En utilisant l'inégalité de Markov, majorer $\mathbb{P}(S_n \leqslant 0)$. En déduire une borne inférieure sur $\mathbb{P}(\forall n, S_n \geqslant 1)$.

Pour l'algorithme, il faut remarquer que

$$\mathbb{P}(\forall k \leqslant n+1, S_k > 0) = \mathbb{P}(\forall k \leqslant n, S_k > 0 \text{ et } G_{n+1} = 2) + \mathbb{P}(\forall k \leqslant n, S_k > 0 \text{ et } G_{n+1} = -1 \text{ et } S_n > 1).$$

On note $P_{n,i} = \mathbb{P}(\forall k \leqslant n, S_k > 0 \text{ et } S_n > i)$. Trouver une relation de récurrence (style triangle de Pascal) sur les $P_{n,i}$ et l'implémenter dans un programme python.

Correction : La variable aléatoire G_n prend la valeur -1 avec probabilité $\frac{1}{2}$ et la valeur $+2$ avec probabilité $\frac{1}{2}$. On écrit $\alpha^{S_n} = \alpha \times \alpha^{G_1} \times \cdots \times \alpha^{G_n}$. On calcule $\mathbb{E}(\alpha^{G_1}) = \cdots = \mathbb{E}(\alpha^{G_n}) = \frac{\alpha^2+\alpha^{-1}}{2}$. Comme $G_1, \ldots, G_n$ sont indépendants, on a :

$$\mathbb{E}(\alpha^{S_n}) = \alpha \left(\frac{\alpha^2 + \alpha^{-1}}{2} \right)^n .$$

On a dit que $0 < \alpha < 1$ donc $(S_n \leqslant 0) = (\alpha^{S_n} \geqslant 1)$. Et par inégalité de Markov on a :

$$\mathbb{P}(\alpha^{S_n} \geqslant 1) \leqslant \alpha \left(\frac{\alpha^2 + \alpha^{-1}}{2} \right)^n .$$

Cette inégalité est vraie pour toutes les valeurs de α, pour avoir la meilleure inégalité possible, on va tout simplement minimiser $\alpha^{-1} + \alpha^2$ sur l'intervalle $]0, 1[$. Pour ça on commence par dériver :

$$\frac{d}{d\alpha}(\alpha^{-1} + \alpha^2) = 2\alpha - \alpha^{-2}$$

La dérivée est une fonction croissante donc $\alpha^{-1} + \alpha^2$ atteint son minimum au point où $2\alpha - \alpha^{-2} = 0$, c'est-à-dire en $\alpha = 2^{-1/3} \approx 0,8$. Et donc on a avec l'aide de la calculatrice :

$$\mathbb{P}(S_n \leqslant 0) \leqslant 2^{-\frac{1}{3}} \left(\frac{2^{-\frac{2}{3}} 2^{\frac{1}{3}}}{2} \right)^n \leqslant 0,8 \times 0,95^n$$

On décompose ensuite l'événement (On n'est jamais ruiné) comme l'union $\bigcap_{n=1}^{\infty} (S_n \geqslant 0)$ et donc :

$$\mathbb{P}(\text{On n'est jamais ruiné}) \geqslant 1 - \sum_{n=0}^{\infty} \mathbb{P}(S_n \geqslant 0) \geqslant 1 - \sum_{n=N}^{\infty} 0,8 \times 0,95^n \geqslant 1 - 0,8 \times 20.$$

En effet $\sum_{n=0}^{\infty} 0,95^n = \frac{1}{1-0,95} = 20$. Maintenant on veut que ce nombre soit positif. Pour ça on calcule la probabilité de ruiner Elon Musk sachant qu'on a réussi à jouer jusqu'à l'étape N pour un nombre N très grand :

$$\mathbb{P}(\text{On n'est jamais ruiné}|\text{On joue la manche } N) \geqslant 1 - \sum_{n=N}^{\infty} \mathbb{P}(S_n \geqslant 0) \geqslant 1 - 16 \times 0,95^N.$$

On trouve à la calculatrice : $0,95^100 \leqslant 0,006$ donc pour $N = 100$ la probabilité de ruiner Elon Musk sachant qu'on n'a toujours pas perdu à la 100^e manche est d'au moins $1 - 0,006 * 16 \geqslant 0,9$. Maintenant on sait que la probabilité de tenir jusqu'à la 100^e manche est d'au moins 2^{-33} car si on gagne les 33 premières manches alors on a 67$ et donc à l'étape 99 on a au moins un dollar et on peut jouer la 100^e manche. On a donc d'après la formule de Bayes

$$\mathbb{P}(\text{on ruine Elon Musk}) \geqslant 2^{-33} \times 0,9.$$

C'est vraiment une très mauvaise borne mais ce n'est pas zéro. Si on veut une borne plus précise, on calcule la probabilité de tenir jusqu'à la N^e manche avec le programme python suivant

```python
N=100
P=[(2*n+1)*[0] for  n in range(N)]
#On veut que P[n][i] represente la proba
#d'etre toujours en jeu apres l'etape n et
#d'avoir strictement plus de i dollars
P[0][0]=1
P[1][0]=1/2
P[1][1]=1/2
P[1][2]=1/2
P[2][0]=1/2
P[2][1]=1/2
P[2][2]=1/4
P[2][3]=1/4
P[2][4]=1/4
#on a rentre les premieres valeurs a la main
#on calcule les suivantes avec une boucle
for n in range(2,N-1):
    for i in range(0,2):
        P[n+1][i]=P[n][1+i]*0.5+P[n][0]*0.5
    for i in range(2,2*n-1):
        P[n+1][i]=P[n][i+1]*0.5+P[n][i-2]*0.5
        #On a une chance sur deux de perdre 1
        #et une chance sur deux de gagner 2
    for i in range(2*n-1,2*n+1):
        P[n+1][i]=P[n][i-2]*0.5
    print(P[n+1][0])
```

On trouve alors $\mathbb{P}(\text{on n'est toujours pas ruiné à l'étape } 100) \approx 0,36$. Et si on prend $N = 1000$ on trouve que la probabilité stagne à la valeur $0,36372875703131613$. En effet $16 \times 0,95^{1000} \approx 10^{-21}$ et l'ordinateur n'a une précision que de 17 chiffres donc le calcul théorique nous garanti qu'on a bien $36,372875703131613\%$ de chances de ruiner Elon Musk avec ce jeu. Attention, des fois quand on calcule la valeur d'une suite avec un programme on peut avoir l'impression que la suite converge alors que ce n'est pas le cas (par exemple si on calcule $\sum_{n=0}^{\infty} \frac{1}{n \log(n)}$ on a l'impression que ça converge alors qu'en fait non). Ici on a trouvé une formule qui nous dit que l'écart entre la probabilité d'être enjeu à l'étape N est la probabilité d'être toujours en jeu est d'au plus $16 \times 0,95^N$. Donc en prenant N assez grand on peut vraiment trouver la probabilité de victoire avec une précision . $\square$