Sommaire

1. Espaces probabilisés dénombrables

1.1. Ensemble dénombrables

Définition : Un ensemble Ω est dit dénombrable si et seulement si $\exists\varphi : \mathbb{N} \to \Omega$, bijective.

Remarque : Cela revient à : $\Omega = \{\omega_n,\ n \in \mathbb{N}\}$.

Exemple : $\mathbb{N}$, $\mathbb{N}^*$, $\mathbb{D}$, $\mathbb{Q}$, l'ensemble des entiers naturels pairs, celui des impairs, sont dénombrables !

1.2. Suite infinie d'événements

L'univers étant infini, on a une infinité d'événements élémentaires, donc une infinité d'événements.

Définition :
Si on a (A_n), une suite infinie d'événements, on définit : $\bigcup\limits_{n\in\mathbb{N}} A_n$ par : $\omega \in \bigcup\limits_{n\in\mathbb{N}} A_n \Leftrightarrow \exists n \in \mathbb{N},\ \omega \in A_n$

Définition :
Si on a (A_n), une suite infinie d'événements, on définit : $\bigcap\limits_{n\in\mathbb{N}} A_n$ par : $\omega \in \bigcap\limits_{n\in\mathbb{N}} A_n \Leftrightarrow \forall n \in \mathbb{N},\ \omega \in A_n$

Définition : On dit que (A_n), une suite infinie d'événements est un **système complet d'événements** si et seulement si : $i \neq j \Rightarrow A_i \cap A_j = \emptyset$, et, $\bigcup\limits_{n\in\mathbb{N}} A_n = \Omega$

1.3. Probabilité sur un univers dénombrable

Définition : On appelle **probabilité sur** Ω, une application : $\mathscr{P}(\Omega) \to [0,1]$ vérifiant $P(\Omega) = 1$, et, pour toute suite (A_n) d'événements 2 à 2 incompatibles : $P\left(\bigcup\limits_{n\in\mathbb{N}} A_n\right) = \sum\limits_{n=0}^{+\infty} P(A_n)$.

Exemple : On tire avec une pièce non truquée jusqu'à obtenir *pile*.
L'univers est $\mathbb{N}^*$, et pour $n \in \mathbb{N}^*$, on a évidemment $P(n) = \dfrac{1}{2^n}$.

2. Indépendance et conditionnement

2.1. Indépendance d'événements

On a toujours cette définition :

Définition : Les deux événements A et B sont **indépendants** $\Leftrightarrow$ $P(A \cap B) = P(A) \,.\, P(B)$.

Théorème : Si $P(B) > 0$, alors : Les deux événements A et B sont indépendants $\Leftrightarrow$ $P(A|B) = P(A)$.

Démonstration : On a : $P(A \cap B) = P(A|B) \,.\, P(B)$, comme $P(B) \neq 0$, on a aussi :
Les deux événements A et B sont indépendants $\Leftrightarrow$ $P(A \cap B) = P(A) \,.\, P(B) = P(A|B) \,.\, P(B)$.
Comme $P(B) \neq 0$, cela équivaut bien à $P(A|B) = P(A)$. ■

> On ne confondra pas événements indépendants et événements incompatibles !
> Avec un dé à jouer, les événements *nombre pair* et *nombre impair* sont incompatibles mais pas indépendants !

Définition : Les événements $(A_1, A_2, \ldots, A_n)$ sont **mutuellement indépendants**
$\Leftrightarrow P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \,.\, P(A_2) \,.\, \cdots \,.\, P(A_n)$.

Théorème : Deux événements d'une famille d'événements mutuellement indépendants sont indépendants.

> Dés événements 2 à 2 indépendants ne constituent pas une famille d'événements mutuellement indépendants !
> Prenons un double lancer d'une pièce non truquée.
> Posons $A = \{pp, pf\}$, $B = \{pp, fp\}$ et $C = \{fp, pf\}$.
> Chacun de ces événements est de probabilité 1/2.
> On a $A \cap B = \{pp\}$, $A \cap C = \{pf\}$ et $B \cap C = \{fp\}$, qui sont tous de probabilité 1/4,
> c'est à dire qu'on a l'indépendance 2 à 2 des événements A, B et C.
> Tandis que $A \cap B \cap C = \emptyset$, donc de probabilité nulle...

2.2. Probabilité conditionnelle

Définition : Si $P(B) > 0$, on appelle **probabilité conditionnelle** de A sachant B,
le réel : $P_B(A) = P(A|B) = \dfrac{P(A \cap B)}{P(B)}$.

2.3. Formule des probabilités composées

On adapte la formule des probabilités composées, qu'on avait vue dans le cas des univers finis, dans le cas d'une suite infinie d'événements :

Théorème : Si on a (A_n) une suite d'événements de conjonction non impossible, c'est à dire telle que :
$P\left(\bigcap_{n \in \mathbb{N}} A_n\right) \neq 0$, alors :

$$P\left(\bigcap_{n \in \mathbb{N}} A_n\right) = P(A_0) \prod_{n=1}^{+\infty} P\left(A_n \,\middle|\, \bigcap_{1 \leqslant k \leqslant n-1} A_k\right)$$

2.4. Formule des probabilités totales

Encore une fois, c'est la généralisation de cette formule qu'on a vue dans un univers fini :

Définition : Si on a (A_n) un système complet d'événements de probabilités non nulles, alors, la série $\sum P(B \cap A_n)$ converge et : $P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n) = \sum_{n=0}^{+\infty} P(B|A_n)\, P(A_n)$

2.5. Formules de Bayes

Ce qui nous donne encore une fois la formule de Bayes :

Théorème : Si $P(A) > 0$ et $P(B) > 0$, alors : $P_B(A) = P(A|B) = \dfrac{P_A(B) \cdot P(A)}{P(B)} = \dfrac{P(B|A) \cdot P(A)}{P(B)}$

Démonstration : Il suffit d'écrire : $P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$. ■

> Cette dernière formule s'utilise souvent aussi avec l'événement A et son contraire $\overline{A}$.
> On obtient : $P(A|B) = \dfrac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P\left(B|\overline{A}\right) \cdot P\left(\overline{A}\right)}$.

3. Variables aléatoires discrètes

3.1. Variable aléatoire réelle

Définition : Une **variable aléatoire réelle** est une application : $\Omega \to \mathbb{R}$.

Définition : La **loi** de la variable aléatoire X est, pour tous les $x \in X(\Omega)$, la donnée de $P(X = x)$.
On la note P_X, ainsi : $P_X(x) = P(X = x)$.

Exemple : Dans l'exemple du tirage d'une pièce jusqu'à obtenir *pile*, on définit la variable aléatoire X par $X(n) = n$.
La loi de cette variable aléatoire est alors bien sûr : $P(X = n) = \dfrac{1}{2^n}$ pour $n \in \mathbb{N}^*$.

Définition : La **fonction de répartition** de la variable aléatoire *réelle* X sur l'espace probabilisé (Ω, P), notée F_X, est une application : $\mathbb{R} \to [0, 1]$ définie par : $F_X(x) = P(X \leqslant x)$.

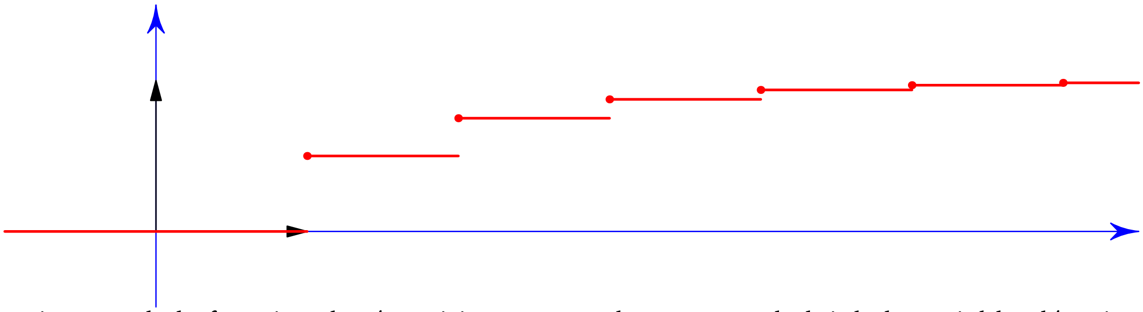
> Cette fonction de répartition est bien entendu croissante.

Exemple : Dans notre exemple, $(X \leqslant n) = (X = 1) \cup (X = 2) \cup \cdots \cup (X = n)$, qui sont des événements incompatibles 2 à 2, donc on a pour $n \in \mathbb{N}^*$:

$$F_X(n) = P(X \leqslant n) = P(X = 1) + P(X = 2) + \cdots + P(X = n) = \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n} = \frac{\frac{1}{2} - \frac{1}{2^{n+1}}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^n}.$$

Maintenant, si $x \notin \mathbb{N}^*$, alors, on a : $F_X(x) = F_X(\mathrm{Ent}(x))$

On a représenté le graphe de cette fonction de répartition.



La connaissance de la fonction de répartition permet de retrouver la loi de la variable aléatoire X.
Il y a deux cas :

- Cas simple : dans $X(\Omega)$, on a un plus grand y strictement plus petit que x.
  Alors $(X \leqslant x) = (X = x) \cup (X \leqslant y)$, ces deux derniers événements étant incompatibles.
  Ce qui donne $P(X = x) = P(X \leqslant x) - P(X \leqslant y) = F_X(x) - F_X(y)$.
  C'est le cas dans notre exemple !
- Cas contraire, beaucoup plus complexe : il faut remplacer $F_X(y)$ par la borne supérieure des $F_X(y)$ avec $y < x$.

Définition : f étant une application définie, entre autres sur $X(\Omega)$ et à valeurs réelles, on définit la variable aléatoire $Y = f(X)$ par : $P(Y = y) = P(f(X) = y)$.

3.2. Espérance

La définition découle de celle vue dans le cas des univers finis. Il faudra quand même au départ vérifier que la série $\sum x_n P(X = x_n)$ converge absolument !

Définition : Soit X, une variable aléatoire *réelle*, sur espace probabilisé dénombrable, et prenant les valeurs (x_n) avec $n \in \mathbb{N}$.
On dit que X est d'espérance finie si et seulement si la série $\sum x_n P(X = x_n)$ converge absolument.
Dans ce cas, l'espérance de X vaut : $E(X) = \sum\limits_{n=0}^{+\infty} x_n P(X = x_n)$.

> L'espérance est la moyenne des valeurs de X pondérée par leur probabilité.

Exemple : Toujours avec notre exemple, notre variable aléatoire X est d'espérance finie car la série $\sum \dfrac{n}{2^n}$ converge, par exemple en utilisant le critère de d'Alembert.
On peut facilement calculer son espérance en utilisant des séries entières.
Pour $x \in]-1, 1[$, $\sum\limits_{n=0}^{+\infty} x^n = \dfrac{1}{1-x}$.
Toujours sur le même ouvert, on dérive cette relation : $\sum\limits_{n=1}^{+\infty} n\, x^{n-1} = \dfrac{1}{(1-x)^2}$,
on en déduit immédiatement : $\sum\limits_{n=1}^{+\infty} n\, x^n = \dfrac{x}{(1-x)^2}$, toujours sur le même ouvert.
On prend maintenant $x = \dfrac{1}{2}$, et on en déduit : $E(X) = \dfrac{\dfrac{1}{2}}{\left(1 - \dfrac{1}{2}\right)^2} = 2$.

Théorème : Si X et Y sont des variables aléatoires réelles d'espérance finie et λ, μ, deux réels, alors : $\lambda X + \mu Y$ est d'espérance finie et, de plus, $E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$.
Ce qui signifie que l'espérance est linéaire !

L'espérance de la fonction constante 1 étant égale à 1, on a aussi : $E(aX + b) = a\,E(X) + b$.

On a maintenant le théorème de transfert :

Théorème : Soit X, une variable aléatoire *réelle*, sur espace probabilisé dénombrable (Ω, P).
On considère aussi l'application φ de $\mathbb{R} \to \mathbb{R}$.
Alors, si $\varphi(X)$ est d'espérance finie, la série $\sum \varphi(x_n)\, P(X = x_n)$ converge absolument.
Son espérance vaut alors : $E(\varphi(X)) = \sum_{n=0}^{+\infty} \varphi(x_n)\, P(X = x_n)$.

3.3. Variance et écart type

La définition de la variance et de l'écart-type prolonge celles vues dans le cas des univers finis.
On commence par un théorème préliminaire un peu technique :

Théorème : Si la variable aléatoire X^2 est d'espérance finie, alors la variable aléatoire X est aussi d'espérance finie.

Démonstration : On a donc $\sum x_n^2 P(X = x_n)$ qui converge absolument. Il en est de même de $\sum P(X = x_n)$ puisque la somme de cette série positive est 1.
Enfin, $\sum \left(1 + x_n^2\right) P(X = x_n)$ converge absolument aussi.
Comme, on a toujours $|x_n| \leqslant 1 + x_n^2$ et que $P(X = x_n) \geqslant 0$, la série $\sum x_n\, P(X = x_n)$ converge absolument.
■

Définition : Lorsque la variable aléatoire X^2 est d'espérance finie,
on appelle variance de X le réel : $V(X) = E(X^2) - E(X)^2$.

Théorème : La variance est aussi l'espérance de $\left(X - E(X)\right)^2$.

Exemple : Reprenons notre exemple de la première occurrence d'un *pile* avec une pièce non truquée.
Pour $n \in \mathbb{N}^*$, on a : $P(X = k) = \dfrac{1}{2^k}$. Il nous faut calculer l'espérance de X^2.
On sait que $\sum_{n=1}^{+\infty} n\, x^n = \dfrac{x}{(1-x)^2}$ sur $]-1, 1[$. On dérive et on obtient $\sum_{n=1}^{+\infty} n^2 x^{n-1} = \dfrac{1+x}{(1-x)^3}$.
Ce qui donne : $\sum_{n=1}^{+\infty} n^2 x^n = \dfrac{x(1+x)}{(1-x)^3}$, on prend $x = \dfrac{1}{2}$ pour obtenir : $E\left(x^2\right) = 6$.
Finalement : $V(X) = 6 - 2^2 = 2$.

Définition : Lorsque la variable aléatoire X^2 est d'espérance finie,
on appelle écart type de X le réel : $\sigma(X) = \sqrt{V(X)} = \sqrt{E(X^2) - E(X)^2}$.

La variance est une moyenne des écarts à la moyenne de la variable aléatoire : elle traduit la *dispersion* de ses valeurs.

Enfin, on a encore l'inégalité de Bienaymé-Tchebychef :

Théorème :
Soit X, une variable aléatoire *réelle* d'espérance finie, sur espace probabilisé dénombrable (Ω, P).
Alors, $\forall t \in \mathbb{R},\ P\Big(|X - E(X)| \geqslant t\Big) \leqslant \dfrac{V(X)}{t^2}$.

3.4. Lois usuelles

a/ Loi géométrique

Définition : Soit $p \in]0,1[$, la variable aléatoire X suit la loi géométrique de paramètre p, notée $\mathscr{G}(p)$, si et seulement si

$$\forall k \in \mathbb{N}^*,\ \mathrm{P}(\mathrm{X} = k) = p(1-p)^{k-1}.$$

Exemple : Notre exemple de tirage d'une pièce non truquée suit la loi géométrique de paramètre $\frac{1}{2}$. Plus généralement, le rang du premier succès dans une répétition infinie d'épreuves de Bernoulli indépendantes de paramètre p suit la loi géométrique de paramètre p aussi.

Théorème : Soit X une variable aléatoire géométrique de paramètre p.

Alors : $\mathrm{E}(\mathrm{X}) = \dfrac{1}{p}$ et $\mathrm{V}(\mathrm{X}) = \dfrac{1-p}{p^2}$.

b/ Loi de Poisson

Définition : Soit $\lambda \in \mathbb{R}_+^*$, la variable aléatoire X suit la loi de Poisson de paramètre λ, notée $\mathscr{P}(\lambda)$, si et seulement si

$$\forall k \in \mathbb{N},\ \mathrm{P}(\mathrm{X} = k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}.$$

La loi de Poisson est destinée à modéliser tout phénomène de temps d'attente comme :
- le nombre de voyageurs arrivant sur un quai de métro en une minute ;
- le nombre d'atomes radioactifs se désintégrant en une seconde ;
- le nombre d'erreurs de transmission sur une ligne ADSL par heure...

Théorème : Soit X une variable aléatoire de Poisson de paramètre λ.

Alors : $\mathrm{E}(\mathrm{X}) = \lambda$ et $\mathrm{V}(\mathrm{X}) = \lambda$.

Théorème :
Si les variables aléatoires X_n suivent des lois binomiales de paramètres (n, p_n) avec $\lim\limits_{n\to\infty} np_n = \lambda$,
et si X suit la loi de Poisson de paramètre λ,
alors la suite de variables aléatoires X_n tend vers la variable aléatoire X.
C'est à dire : $\lim\limits_{n\to\infty} \mathrm{P}(\mathrm{X}_n = k) = \mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!} = \mathrm{P}(\mathrm{X} = k)$

Démonstration : Cette démonstration n'est pas rigoureuse !...
On travaille avec un k fixé et un n grand.

$$\mathrm{P}(\mathrm{X}_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k} \underset{+\infty}{\sim} \frac{(np_n)^k}{k!}(1-p_n)^{n-k}$$

$$\underset{+\infty}{\sim} \frac{\lambda^k}{k!}(1-p_n)^n = \frac{\lambda^k}{k!}\exp(n\ln(1-p_n)) = \frac{\lambda^k}{k!}\exp(-np_n + \mathrm{o}(1)) \underset{+\infty}{\sim} \frac{\lambda^k}{k!}\exp(-np_n + \mathrm{o}(1)) \underset{+\infty}{\sim} \frac{\lambda^k}{k!}\exp(-\lambda) \ \blacksquare$$

c/ Espérance et variance des lois usuelles

Nom	Valeurs	Paramètre	Loi	Espérance	Variance
Bernoulli	$(0,1)$	$p \in [0,1]$	$\mathrm{P}(\mathrm{X}=1)=p$	p	$p(1-p)$
Binomiale	$(0,1,\ldots,n)$	$n \in \mathbb{N}^*,\ p \in [0,1]$	$\mathrm{P}(\mathrm{X}=k)=\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
Géométrique	$\mathbb{N}^*$	$p \in [0,1]$	$\mathrm{P}(\mathrm{X}=k)=p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
Poisson	$\mathbb{N}$	$\lambda \in \mathbb{R}_+^*$	$\mathrm{P}(\mathrm{X}=k)=\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}$	λ	λ