Probabilités

Contents

On ne peut guère donner une définition satisfaisante de la probabilité. La définition complète de la probabilité est donc une sorte de pétition de principe.

Henri Poincaré (1854-1912) - Calcul des Probabilités

Quelques résultats utiles

0.1 Quelques formules

Démontrer par le calcul ou un raisonnement :

1) Si $k \geq 1$,

$$\frac{n}{k}.\binom{n-1}{k-1} = \binom{n}{k}$$

$$\binom{n}{k} = \binom{n}{n-k}$$

2) Si $n \geq 1$, et $0 \leq k \leq n$:

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

3)

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

$$\sum_{k=0}^{n} k.\binom{n}{k} = n.2^{n-1}$$

4)

$$\sum_{k=0}^{p} \binom{k}{n} = \sum_{k=n}^{p} \binom{k}{n} = \binom{p+1}{n+1}$$

5)

$$\forall x \in \left]-1,1\right[, \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

$$\forall x \in \left]-1,1\right[, \sum_{n=0}^{\infty} n.x^n = \frac{x}{(1-x)^2}$$

6) Si $0 \leq j \leq k \leq n$,

$$\binom{n}{k}.\binom{k}{j} = \binom{n}{j}.\binom{n-j}{k-j}$$

7)

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$$

8)

$$\sum_{i+j=k} \binom{n}{i}\binom{m}{j} = \binom{n+m}{k}$$

Démonstrations

1) Clair
2) Clair
3) $(1+1)^n$; dériver $(1+X)^n$
4)
$$\sum_{k=0}^{p} (1+X)^k = \frac{(1+X)^{p+1}-1}{X}$$

ou avec par une récurrence facile sur p avec

$$\binom{p+1}{n+1} + \binom{p+1}{n} = \binom{p+2}{n+1}$$

5) Dériver
$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$$

6) Par le calcul, ou en comptant le nombre de couples (A,B) tels que :

$$A \subset B \subset E,\ |A| = j,\ |B| = k,\ |E| = n$$

7)
$$(1+X)^n (1+X)^n = (1+X)^{2n}$$

8) Coefficient de X^k dans $(1+X)^{n+m}$.

0.2 Ensembles

Soit E, F, G des ensembles ; soit $f : E \to F$ et $g : F \to G$ des applications.

Sur les images réciproques

Soit $(F_i)_{i \in I}$ une famille de parties de F.

L'image réciproque de la réunion $\bigcup_{i \in I} F_i$ est la réunion des images réciproques. C'est vrai aussi pour l'intersection.

Cas de l'image

L'image d'une union est l'union des images ; ce n'est pas toujours le cas pour l'intersection.

Sur les complémentaires

Soit A est une partie de F et B son complémentaire dans F.

Alors $f^{-1}(A)$ et $f^{-1}(B)$ sont complémentaires dans E.

Que dire de $(g \circ f)^{-1}(B)$?

Si B est une partie de G :

$$(g \circ f)^{-1}(B) = f^{-1}\left(g^{-1}(B)\right)$$

Que dire de $(g \circ f)^{-1}(\{b\})$?

Soit $b \in G$.
$$(g \circ f)^{-1}(\{b\}) = \bigcup_{a \in g^{-1}(b)} f^{-1}(\{a\})$$

1 Espaces probabilisés

1.1 Introduction

1.1.1 Univers finis

Appellation	notation	exemple
un résultat possible, ou réalisation, ou issue	ω	3
un événement élémentaire		$\{2\}$
univers : tous les résultats possibles	Ω	$\{1, 2, 3, 4, 5, 6\}$
un événement	$A \subset \Omega$	$\{1, 3, 5\}$

1.1.2 Cas général

Dans le cas d'un univers fini Ω, un événement est une partie quelconque de Ω.

Mais si par exemple on souhaite modéliser le nombre de lancers de dé nécessaires à l'obtention d'un 6, on choisit naturellement $\Omega = \{1, 2, 3, 4, 5, 6\}^{\mathbb{N}}$, ensemble qui n'est pas dénombrable.

Lorsque Ω n'est pas dénombrable, pour pouvoir construire une probabilité satisfaisante, il est préférable de restreindre les événements que l'on considère à une partie de $\mathcal{P}(\Omega)$ appelée tribu.

1.2 Evénements

1.2.1 Définition

Une tribu T sur Ω est une partie de $\mathcal{P}(\Omega)$ (ensemble des parties de Ω) qui vérifie :

a) $\Omega \in T$

b) T est stable par passage au complémentaire

c) T est stable par union finie ou dénombrable

Les éléments de T sont appelés événements, (Ω, T) est appelé espace probabilisable.

Définitions

Evénement

- certain : Ω.
- impossible : $\varnothing$.
- contraire : $\Omega \setminus A$, ou A^c, ou $\overline{A}$.

Evénements incompatibles, ou disjoints : $A \cap B = \varnothing$.

Système complet d'événements, ou partition : famille d'événements deux à deux disjoints, dont la réunion est Ω. On supposera toujours la famille indexée par un ensemble fini ou dénombrable.

Propriété

Toute intersection finie ou dénombrable d'éléments de T est dans T.

Exercice

Si A et B sont des événements, $A \setminus B$ est un événement.

En effet :

$$A \setminus B = A \cap \overline{B}$$

où $\overline{B} = \Omega \setminus B$ désigne le complémentaire de B dans Ω.

Exercice

Toute intersection de tribus sur Ω en est une.

1.2.2 Exercice : 'transport' de tribus

Soit E et F deux ensembles, et $f : E \to F$ une application.

- L'image d'une tribu sur E n'est pas toujours une tribu sur $f(E)$, même si f est surjective.
- Soit T une tribu sur E ; $\left\{A/f^{-1}(A) \in T\right\}$ constitue une tribu sur F.
- Soit T une tribu sur F ; $\left\{f^{-1}(A)/A \in T\right\}$ constitue une tribu sur E.

Un contre-exemple

$E = \{a, b, c\}$, $F = \{0, 1\}$, $f(a) = f(b) = 0, f(c) = 1$ et

$$T = \{E, \{a, c\}, \{b\}, \phi\}$$

1.3 Probabilité

1.3.1 Définition

Si T est une tribu sur Ω, une probabilité sur (Ω, T) est une application $\mathbb{P}$ définie sur T, à valeurs dans $[0,\ 1]$, telle que

- $\mathbb{P}(\Omega) = 1$.
- Pour toute suite $(A_n)_{n\geq 0}$ d'événements deux à deux disjoints,

$$\mathbb{P}(\bigcup_{n=0}^{\infty} A_n) = \sum_{n=0}^{\infty} \mathbb{P}(A_n)$$

On dit que $(\Omega, T, \mathbb{P})$ est un espace probabilisé.

1.3.2 Exemple : la probabilité uniforme

Si Ω est fini, de cardinal n, on définit $\mathbb{P}$ par $\mathbb{P}(A) = ?$

Réponse

$$\mathbb{P}(A) = \frac{\mathrm{card} A}{n}$$

1.4 Cas d'un ensemble dénombrable

Si Ω est fini ou dénombrable et si $T = \mathcal{P}(\Omega)$, une probabilité $\mathbb{P}$ sur (Ω, T) s'identifie, via la formule

$$\mathbb{P}(\{\omega\}) = p_\omega$$

à une famille de réels positifs de somme 1 ; dans ce cas, $\mathbb{P}(A) = ?$

Réponse

$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$$

Exemples avec $\Omega = \mathbb{N}^*$

$p_n = \frac{1}{2^n}$, ou $p_n = \frac{1}{n(n+1)}$. Quelles expériences peuvent-elles être modélisées ainsi ?

Réponse

Pour

$$p_n = \frac{1}{2^n}$$

lancer une pièce jusqu'à obtenir pile.

Pour

$$p_n = \frac{1}{n(n+1)}$$

choisir un nombre x au hasard dans $]0,1[$, puis poser $n = \left\lfloor \frac{1}{x} \right\rfloor$.

En résumé :

$$\text{return floor}\left(\frac{1}{\text{random}\,()}\right)$$

Montrons qu'on définit ainsi une probabilité.

Démonstration

Soit $(A_n)_{n\geq 0}$ une suite d'événements deux à deux disjoints ; soit

$$A = \bigcup_{n=0}^{\infty} A_n$$

On veut montrer que

$$\mathbb{P}(A) = \sum_{n=0}^{\infty} \mathbb{P}(A_n)$$

$(A_n)_{n\geq 0}$ constituant une partition de A, on constate qu'il s'agit d'une application immédiate du théorème de sommation par paquets des familles de réels positifs.

1.5 Premières propriétés des probabilités

a) Si $A \subset B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$
b) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
c) $\mathbb{P}(\varnothing) = 0$; est-ce le seul événement de probabilité nulle ?
d) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Démonstration

Soit $A_1 = A \setminus (A \cap B)$

$$\mathbb{P}(A) = \mathbb{P}(A_1) + \mathbb{P}(A \cap B)$$

Et

$$\mathbb{P}(A \cup B) = \mathbb{P}(A_1) + \mathbb{P}(B)$$

1.6 Exercices

1.6.1 Exercice 1

On lance deux dés ; quel univers choisir ? On note S la somme des deux chiffres. Quelle est la probabilité que $S = 2$? Que $S = 3$? Que S soit paire ?

Réponses

$\{1,2,3,4,5,6\}^2$; $\frac{1}{36}$, $\frac{2}{36}$; $\frac{1}{2}$.

1.6.2 Exercice 2

On tire au hasard deux cartes sans remise dans un jeu de 52 cartes. Quelle est la probabilité pour

a) que la couleur des deux cartes soit pique ?
b) qu'il y ait au moins un coeur et un as ?
c) qu'il y ait exactement un coeur et un as ?

Réponses

$\frac{1}{17}$, $\frac{29}{26\times 17}$, $\frac{12}{13\times 17}$.

1.6.3 Exercice 3

On tire trois cartes au hasard, l'une après l'autre et sans remise, dans un jeu de 32 cartes.

Quelle est la probabilité que la troisième carte tirée soit une dame ?

1.6.4 Exercice 4

On tire deux cartes d'un jeu de 32.

Quelle est la probabilité que la deuxième soit l'as de coeur sachant que la première est un coeur ?

Réponse

Nombre de cas favorables : 7 ; nombre de cas possibles : 8×31.

Conclusion : $\frac{7}{8\times 31}$.

2 Propriétés élémentaires des probabilités

2.1 Lemme

Le complémentaire d'une intersection est la réunion des complémentaires ; et inversement.

Démonstration

Soit (A_j) des parties de E ; soit $x \in E$; il y a équivalence entre :

$$x \notin \bigcap_{j\in I} A_j$$

$$\exists j \in I, x \notin A_j$$

$$\exists j \in I, x \in \overline{A_j}$$

$$x \in \bigcup_{j\in I} \overline{A_j}$$

2.2 Continuité croissante

Théorème

Si $(A_n)_{n\geq 0}$ est une suite d'événements croissante pour l'inclusion, alors

$$\mathbb{P}(A_n) \underset{n}{\longrightarrow} \mathbb{P}(\bigcup_{n=0}^{\infty} A_n)$$

Démonstration

On note $B_0 = A_0$ et $B_n = A_n \setminus A_{n-1}$ pour $n \geq 1$; on constate que les (B_n) sont disjoints et que leur réunion est la réunion des (A_n) ; donc

$$\mathbb{P}(\bigcup_{n=0}^{\infty} A_n) = \mathbb{P}(\bigcup_{n=0}^{\infty} B_n) = \sum_{n=0}^{\infty} \mathbb{P}(B_n) = s$$

Pour les mêmes raisons, pour tout $n \geq 0$

$$\mathbb{P}(A_n) = \mathbb{P}(\bigcup_{k=0}^{n} B_k) = \sum_{k=0}^{n} \mathbb{P}(B_k) = s_n$$

La suite (s_n) converge vers s, d'où la conclusion.

2.3 Continuité décroissante

Théorème

Si $(A_n)_{n\geq 0}$ est une suite d'événements décroissante pour l'inclusion, alors :

$$\mathbb{P}(A_n) \underset{n}{\longrightarrow} \mathbb{P}(\bigcap_{n=0}^{\infty} A_n)$$

Démonstration

Notons A l'intersection des (A_n) et $C_n = \Omega \backslash A_n$.

La suite (C_n) est croissante ; soit C sa réunion ; on est ramené au cas précédent :

$(1 - \mathbb{P}(A_n)) = (\mathbb{P}(C_n))$ converge vers $\mathbb{P}(C)$.

Or C est est la réunion des complémentaires, donc le complémentaire de l'intersection des (A_n).

On a montré que $(1 - \mathbb{P}(A_n)) = (\mathbb{P}(C_n))$ converge vers $1 - \mathbb{P}(A)$.

2.4 Réunion

Théorème

Si $(A_n)_{n\geq 0}$ est une suite d'événements, alors :

$$\mathbb{P}(\bigcup_{n=0}^{\infty} A_n) \leq \sum_{n=0}^{\infty} \mathbb{P}(A_n)$$

Démonstration

Par récurrence sur n, on montre que, pour tout $n \geq 0$

$$\mathbb{P}(\bigcup_{k=0}^{n} A_k) \leq \sum_{k=0}^{n} \mathbb{P}(A_k)$$

Ensuite, on applique le théorème de continuité croissante à la suite (B_n) croissante suivante

$$B_n = \bigcup_{k=0}^{n} A_k$$

2.5 Définitions

On dit qu'un événement A est négligeable si

$$\mathbb{P}(A) = 0$$

et presque sûr si

$$\mathbb{P}(A) = 1$$

Propriétés presque sûres.

Propriété

Une réunion dénombrable d'événements négligeables est négligeable. En effet, dans ce cas :

$$\mathbb{P}(\bigcup_{n=0}^{\infty} A_n) \leq \sum_{n=0}^{\infty} \mathbb{P}(A_n) = 0$$

3 Probabilités conditionnelles

3.1 $\mathbb{P}(A|B)$

3.1.1 Définition

Si $\mathbb{P}(B) > 0$, la probabilité conditionnelle de A sachant B est notée $\mathbb{P}_B(A)$ ou $\mathbb{P}(A|B)$ et définie par

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

On définit ainsi une nouvelle probabilité $\mathbb{P}_B$ sur (Ω, T).

3.1.2 Justification sur un exemple

Un lancer de dé ; univers de cardinal $n = 6$.
- événement A : le résultat est un nombre premier.
- événement B : le résultat est un nombre pair.

$$\frac{|A \cap B|}{|B|} = \frac{|A \cap B|}{n} \div \frac{|B|}{n}$$

$\frac{|A \cap B|}{|B|}$ est la proportion de nombres vérifiant A parmi ceux vérifiant B.

3.1.3 Cas de la probabilité uniforme

Si Ω est fini, et la probabilité uniforme, alors $\mathbb{P}_B$ est uniforme sur B.

3.2 Formule des probabilités composées

Si $\mathbb{P}(A) \neq 0$:

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

Généralisation

Si $\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_{n-1})$ est non nul :

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1) . \mathbb{P}(A_2|A_1) . \mathbb{P}(A_3|A_1 \cap A_2) ... \mathbb{P}(A_n|A_1 \cap ... \cap A_{n-1})$$

Démonstration

Par récurrence sur n, en remarquant que :

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1 \cap ... \cap A_{n-1}) . \mathbb{P}(A_n|A_1 \cap ... \cap A_{n-1})$$

3.3 Formule des probabilités totales

Soit A et B sont deux événements tels que $0 < \mathbb{P}(A) < 1$; alors :

$$\mathbb{P}(B) = \mathbb{P}(B|A) . \mathbb{P}(A) + \mathbb{P}(B|A^c) . \mathbb{P}(A^c)$$

Démonstration

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

Généralisation

Soit $(A_i)_{i \in I}$ un système dénombrable complet d'évènements de probabilités non nulles ; alors pour tout événement B,

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B|A_i) . \mathbb{P}(A_i)$$

3.4 Formule de Bayes

Si A et B sont deux événements tels que $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$:

$$\mathbb{P}\left(A|B\right) = \frac{\mathbb{P}\left(B\left|A\right.\right).\mathbb{P}\left(A\right)}{\mathbb{P}\left(B\right)} = \frac{\mathbb{P}\left(B\left|A\right.\right).\mathbb{P}\left(A\right)}{\mathbb{P}\left(B\left|A\right.\right).\mathbb{P}\left(A\right) + \mathbb{P}\left(B\left|A^c\right.\right).\mathbb{P}\left(A^c\right)}$$

Généralisation

Soit $(A_i)_{i \in I}$ un système complet d'événements de probabilités non nulles ; alors pour tout événement B de probabilité non nulle :

$$\mathbb{P}\left(A_j|B\right) = \frac{\mathbb{P}\left(B\left|A_j\right.\right).\mathbb{P}\left(A_j\right)}{\mathbb{P}\left(B\right)} = \frac{\mathbb{P}\left(B\left|A_j\right.\right).\mathbb{P}\left(A_j\right)}{\sum_{i \in I} \mathbb{P}\left(B|A_i\right).\mathbb{P}\left(A_i\right)}$$

3.5 Exemples

3.5.1 Ex 1

On dispose de trois jetons : un comporte deux faces blanches, un autre deux faces noires, le dernier une face blanche et une noire, notés BB, NN, BN.

On choisit un jeton au hasard ; sachant qu'on obtient une face blanche (FB), quelle est la probabilité que le jeton à deux faces blanches (BB) ait été choisi ?

Réponse 1

$$\mathbb{P}(BB|FB) = \frac{\mathbb{P}(BB \cap FB)}{\mathbb{P}(FB)} = \frac{1}{3} / \frac{1}{2} = \frac{2}{3}$$

Réponse 2

Les 6 faces sont équiprobables ; sachant qu'on obtient une face blanche, il y a deux chances sur 3 pour qu'elle appartienne au jeton blanc.

3.5.2 Ex 2

Lorsqu'un bon gérant (B) de portefeuille achète une valeur boursière, la probabilité que le cours de cette valeur monte (M) est de $0,8$; sinon, la probabilité que le cours descende (D) est de $0,6$. On sait par ailleurs que si l'on choisit au hasard un gérant, il y a une chance sur 10 que celui-ci soit un bon (B).

Un client choisit au hasard un gérant, et lui demande d'acheter une valeur. Sachant que le cours de cette valeur est monté, calculer la probabilité pour que le gérant soit un bon ; quel est l'univers ici ?

Réponse

$\Omega = \{B, N\} \times \{M, D\}$; $\frac{2}{11}$.

3.5.3 Ex 3

Un test médical se trompe une fois sur cent ; par ailleurs, 5% des personnes sont malades (M).

a) Si le test d'une personne est positif (P), quelle est la probabilité pour que cette personne soit effectivement malade (M) ? Pourquoi pas $\frac{99}{100}$?

b) Si le test d'une personne est négatif (N), quelle est la probabilité pour que cette personne soit malade ?

Réponses

a) $\mathbb{P}(P|M) = 0.99$; $\mathbb{P}(M) = 0.05$; $\mathbb{P}(P) = \frac{59}{1000}$; $\mathbb{P}(M|P) = \frac{99}{118}$; la probabilité d'une erreur est de 0.01 ; mais la probabilité d'une erreur sachant (P) est différente.
b) $\frac{1}{1882}$.

3.5.4 Ex 4

On lance un dé 3 fois ; on obtient 3 fois le chiffre 6 ; quelle est la probabilité qu'il soit non truqué ?

Réponse

Probabilité d'obtenir 3 fois le 6 sachant qu'il est non truqué : $\mathbb{P}(A|B) = \frac{1}{6^3}$.

$$\mathbb{P}\left(A\middle|B\right) = \frac{\mathbb{P}(B\,|A)\,.\mathbb{P}(A)}{\mathbb{P}(B)}$$

Mais on ne peut rien conclure ; que manque-t-il ?

Réponse

On suppose de plus qu'un dé sur cent est truqué, et que pour un tel dé, on obtient 6 pour un lancer sur deux ; quelle est la réponse dans ce cas ?

Réponse

$\mathbb{P}(B) = 0.99$; $\mathbb{P}(A) = \frac{99}{100}.\frac{1}{6^3} + \frac{1}{100}.\frac{1}{8}$; $\mathbb{P}(B|A) = \frac{11}{14}$.

3.6 Indépendance

3.6.1 Définition

On dit que A et B sont indépendants lorsque

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

Notation : $A \perp\!\!\!\perp B$.

3.6.2 Exemple avec un dé

On lance un dé ; D : ' X est pair ', et T : ' X est multiple de 3 ' sont-ils indépendants ?

Même question avec 2 et 5.

Réponses

Oui ; non.

3.6.3 Exemple avec des cartes

On tire une carte dans un jeu de 32.

' un as ' et ' un coeur ' sont-ils indépendants ?

Réponse

Oui.

3.6.4 Remarque

Si $\mathbb{P}(B) > 0$, l'indépendance de A et B s'écrit

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

Si $\mathbb{P}(B) = 0$ ou 1, A et B sont toujours indépendants.

Pour cette raison, dans le cas où $\mathbb{P}(B) = 0$, on peut définir $\mathbb{P}(A|B)$ par :

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

3.6.5 Indépendance et complémentaire

Exercice

Montrer que A et B sont indépendants si et seulement si A et $(\Omega \setminus B)$ sont indépendants.

Démonstration

Supposons que $\mathbb{P}(A \cap B) = \mathbb{P}(A).\mathbb{P}(B)$; soit $B' = \Omega \setminus B$.

$A \cap B$ et $A \cap B'$ constituent une partition de A, donc :

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

Donc

$$\mathbb{P}(A \cap B') = \mathbb{P}(A) . (1 - \mathbb{P}(B)) = \mathbb{P}(A).\mathbb{P}(B')$$

3.6.6 Exercice

Deux événements A et B disjoints peuvent-ils être indépendants ?

Réponse

Seulement si $\mathbb{P}(A) = 0$ ou $\mathbb{P}(B) = 0$.

3.7 Indépendance, généralisation

3.7.1 Définition

Des événements $(A_i)_{i \in I}$ sont mutuellement indépendants, si pour toute sous famille finie, la probabilité de l'intersection est le produit des probabilités.

3.7.2 Exemple

Considérons l'expérience aléatoire qui consiste à lancer un dé deux fois et les événements A, B, C définis respectivement par « le premier chiffre est pair », « le deuxième chiffre est impair », « la somme des chiffres est paire ».

A, B, C sont-ils deux à deux indépendants, forment-ils une famille indépendante ?

Réponse

Deux à deux indépendants, non mutuellement indépendants.

$$\mathbb{P}(A \cap B \cap C) = 0$$

3.7.3 Indépendance et complémentaire

Exercice

On suppose des événements $(A_i)_{i \in I}$ mutuellement indépendants ; si on remplace certains par leur complémentaire dans Ω, la famille reste indépendante.

4 Variables aléatoires discrètes

4.1 Un exemple

On lance deux dés, un blanc et un rouge ; l'univers est $\Omega = \{1, 2, 3, 4, 5, 6\}^2$; on s'intéresse à la somme des deux chiffres ; il est donc naturel d'introduire une application S de Ω dans $\mathbb{R}$, ou dans $\Omega' = \{2, ..., 12\}$; S est définie par ?

Réponse

$$S((x, y)) = x + y$$

Une telle fonction S est appelée variable aléatoire ; remarquons que ce n'est pas une variable, mais une fonction !

Ensuite, on s'intéresse à la probabilité que S prenne une valeur donnée ; $\mathbb{P}(S = 2)$, $\mathbb{P}(S = 3)$?

Réponse

$$\mathbb{P}(S = 2) = \frac{1}{36} \text{ ; } \mathbb{P}(S = 3) = \frac{2}{36}$$

On constate qu'on définit ainsi une probabilité sur Ω' ; cette probabilité est appelée la loi de S.

4.2 Variables aléatoires discrètes

4.2.1 Définition

Étant donné un ensemble E et un espace probabilisé $(\Omega, T, \mathbb{P})$, une variable aléatoire discrète définie sur Ω est une application X de Ω dans E telle que $E' = X(\Omega)$ soit fini ou dénombrable et que, pour tout x de $X(\Omega)$,

$$X^{-1}(\{x\}) \in T$$

Lorsque $E = \mathbb{R}$, la variable aléatoire est dite réelle.

Exemple

On lance un dé 10 fois ; exemples de variables aléatoires réelles :
- la somme des chiffres obtenus.
- le chiffre maximum.
- le nombre de 6.
- le rang du premier 6.

4.2.2 L'événement $\{X \in A\}$

Soit X une variable aléatoire discrète définie sur Ω à valeurs dans E.

Si A est une partie de E, l'ensemble

$$\{t \in \Omega / X(t) \in A\}$$

est noté $X^{-1}(A)$, ou $(X \in A)$; il s'agit d'un événement ; pourquoi ?

Réponse

C'est vrai si A est un singleton, et T est stable par union dénombrable.

Autres notations

Si $E = \mathbb{R}$, on note aussi $(X \leq x)$ pour $\{t \in \Omega / X(t) \leq x\}$; $(X \geq x)$; $(X = x)$...

4.2.3 Loi $\mathbb{P}_X$ de la variable aléatoire X

En résumé, $\mathbb{P}_X = \mathbb{P} \circ X^{-1}$; en détail :

$$\mathbb{P}_X(H) = \mathbb{P}(X \in H) = \mathbb{P}(\{x \in \Omega / X(x) \in H\}) = \mathbb{P}\left(X^{-1}(H)\right)$$

Propriété

$\mathbb{P}_X$ est une probabilité sur $E' = X(\Omega)$: la loi d'une variable aléatoire est une probabilité sur l'ensemble des valeurs prises par X.

Tribu : $P(E')$.

Notation

$X \sim Y$ si X et Y ont la même loi ; $X \sim \mathcal{L}$ si $\mathcal{L}$ est la loi de X.

4.3 Loi de Bernoulli de paramètre $p \in [0,1]$. Notation $B(p)$

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

Interprétation : succès d'une expérience.

4.4 Loi binomiale de paramètres $n \geq 1$ et $p \in [0,1]$. Notation $B(n,p)$

Interprétation

Nombre de succès lors de la répétition de n expériences de Bernoulli indépendantes, ou tirages avec remise dans un modèle d'urnes :

Une urne contient b boules blanches et r boules rouges ; si on tire une boule, la probabilité qu'elle soit blanche est

$$p = \frac{b}{b + r}$$

On tire n fois avec remise, et on appelle X le nombre de boules blanches obtenues ; la probabilité d'obtenir k boules blanches est

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

Démonstration

On peut choisir

$$\Omega = \{B, R\}^n$$

$p^k.(1-p)^{n-k}$ est par exemple la probabilité d'obtenir k boules blanches suivies de $n - k$ boules rouges.

$\binom{n}{k}$ est le nombre de choix des rangs de sortie des k boules blanches.

Remarque

Si $X \sim B(n,p)$,

$$X = X_1 + X_2 + ... + X_n$$

où les X_k suivent une loi de Bernoulli : $X \sim B(p)$,

et sont mutuellement indépendantes.

4.5 $f \circ X$

4.5.1 $f \circ X$ est une vad

Théorème

Si X est une variable aléatoire discrète, et f une fonction définie sur $X(\Omega)$, $Y = f \circ X$ est une variable aléatoire discrète.

Démonstration

Il est clair que $Y(\Omega) = f(X(\Omega))$ est dénombrable.

Soit $b \in Y(\Omega)$; soit (a_j) la famille des antécédents de b par f ; alors :

$$Y^{-1}(b) = \bigcup_{j \in J} X^{-1}(a_j)$$

événement car union dénombrable d'événements.

4.5.2 Exercice : $f \circ X \sim f \circ Y$

Montrer que si f est une fonction définie sur $X(\Omega) = Y(\Omega)$, et si $X \sim Y$, alors

$$f \circ X \sim f \circ Y$$

5 Couples de variables aléatoires

5.1 Généralités

On se donne ici deux variables aléatoires discrètes X, Y définies sur le même espace probabilisé $(\Omega, T, \mathbb{P})$.

5.1.1 Définition

On appelle couple (X, Y) l'application :

$$Z : \omega \to (X(\omega), Y(\omega))$$

C'est une variable aléatoire discrète ; pourquoi ?

Réponse

Le produit de deux ensembles dénombrables est dénombrable ; or

$$Z(\Omega) \subset X(\Omega) \times Y(\Omega)$$

donc $Z(\Omega)$ est fini ou dénombrable.

Soit $(u, v) \in Z(\Omega)$.

$$Z^{-1}(u, v) = X^{-1}(u) \cap Y^{-1}(v)$$

est un événement car intersection de deux événements.

5.1.2 Somme, produit

Théorème

La somme et le produit de deux variables aléatoires réelles discrètes sont aussi deux variables aléatoires réelles discrètes.

Démonstration

Découle de ce qui précède et de **4.5**.

5.1.3 Loi conjointe du couple (X, Y)

Pour tout couple (x, y) de réels, on note

$$p_{x,y} = \mathbb{P}((X = x) \cap (Y = y))$$

5.1.4 Lois marginales

Les lois de X et Y sont appelées lois marginales. La connaissance de la loi de X et de la loi de Y ne suffit pas à connaître la loi de (X, Y).

Remarque

Les lois marginales sont déterminées par la loi conjointe ; en effet :

$$p_x = \mathbb{P}(X = x) = \sum_y \mathbb{P}((X = x) \cap (Y = y)) = \sum_y p_{x,y}$$

Explication

$$(X = x) = \bigcup_{y \in Y(\Omega)} ((X = x) \cap (Y = y))$$

et il s'agit d'une union disjointe.

5.1.5 Exemple

On lance deux dés, un blanc (X) et un rouge (Y) ; comparer les lois conjointes des couples (X, Y), (Y, X), et (X, X).

Réponse

(X, Y) et (Y, X) ont la même loi...

5.1.6 Loi conditionnelle de Y sachant $(X = x)$

Si $p_x > 0$: $\mathbb{P}(Y = y | X = x) = \frac{p_{x,y}}{p_x}$.

Extension

Aux $n-$uplets de variables aléatoires. Vecteurs aléatoires discrets.

5.2 Couples de variables aléatoires indépendantes

5.2.1 Définition

Soit X, Y deux variables aléatoires discrètes définies sur $(\Omega, T, \mathbb{P})$.
On dit que X et Y sont indépendantes si :
pour tout couple $(x, y) \in X(\Omega) \times Y(\Omega)$,

$$p_{x,y} = \mathbb{P}((X = x) \cap (Y = y)) = \mathbb{P}(X = x) . \mathbb{P}(Y = y) = p_x . p_y$$

Propriété

Si X et Y sont indépendantes :

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

Donc les événements $X^{-1}(A) = (X \in A)$, $Y^{-1}(B) = (Y \in B)$ sont indépendants.

Démonstration sommaire

Avec le théorème de sommation par paquets :

$$\mathbb{P}((X, Y) \in A \times B) = \sum_{x,y} p_{x,y} = \sum_{x,y} p_x . p_y = \mathbb{P}(X \in A) . \mathbb{P}(Y \in B)$$

5.2.2 X et une constante Y

Exercice

On suppose Y constante : $Y = a$.
X et Y sont-elles indépendantes ?

Réponse

On vérifie

$$\mathbb{P}((X=x)\cap(Y=y))=\mathbb{P}(X=x).\mathbb{P}(Y=y)$$

avec $y=a$.

5.2.3 X et X ?

Exercice

X et X peuvent-elles être indépendantes ?

Réponse

Soit a tel que

$$\mathbb{P}(X=a)\neq 0$$

Dans ce cas, si $a\neq b$,

$$\mathbb{P}(X=a).\mathbb{P}(X=b)=\mathbb{P}((X=a)\cap(X=b))=0$$

Donc $\mathbb{P}(X=b)=0$.

Il existe donc un a unique tel que $\mathbb{P}(X=a)=1$: X est constante (presque sûrement).

5.2.4 Exemple de deux dés

On lance deux dés, un blanc et un rouge ; on pose $S=X+Y$, $P=X.Y$, $m=\min(X,Y)$, $M=\max(X,Y)$.

Parmi X,Y,S,P,m,M, quels sont les couples de variables aléatoires indépendantes ?

Réponse

Il n'y a que (X,Y). Par exemple

$$\mathbb{P}(m=6).\mathbb{P}(M=1)\neq 0=\mathbb{P}((m=6)\cap(M=1))$$

5.2.5 Exemple de deux lois de Bernoulli

On suppose que X et Y suivent la loi de Bernoulli $B\left(\frac{1}{2}\right)$ et sont indépendantes ; soit

$$Z=1-Y$$

Que dire de $T_1=(X,X)$, $T_2=(X,Y)$, $T_3=(X,Z)$?

Réponse

Toutes les lois marginales sont identiques ; T_2 et T_3 ont la même loi conjointe.

5.2.6 Théorème

Si X et Y sont indépendantes, et f et g deux fonctions quelconques, définies sur $X(\Omega)$ et $Y(\Omega)$, les variables aléatoires $f(X)$ et $g(Y)$ le sont aussi.

Conditions sur f et g

Aucune condition sur f et g :

On sait que si X est une variable aléatoire discrète, $f\circ X$ est automatiquement une variable aléatoire discrète.

Démonstration

Notons $X_1 = f \circ X$ et $Y_1 = g \circ Y$. Soit $a \in X_1(\Omega)$ et $b \in Y_1(\Omega)$; notons

$$A_1 = f^{-1}(a), \ B_1 = g^{-1}(b)$$

Un dessin !

$$\{\omega \in \Omega / f \circ X(\omega) = a\} = \left\{\omega \in \Omega / X(\omega) \in f^{-1}(a)\right\}$$

Ou encore :

$$(X_1 = a) = X_1^{-1}(a) = (f \circ X)^{-1}(a) = X^{-1}\left(f^{-1}(a)\right)$$

Donc

$$(X_1 = a) = X^{-1}\left(f^{-1}(a)\right) = X^{-1}(A_1) = (X \in A_1)$$

De même :

$$(Y_1 = b) = Y_1^{-1}(b) = Y^{-1}(B_1) = (Y \in B_1)$$

Or,

$$\mathbb{P}((X \in A_1) \cap (Y \in B_1)) = \mathbb{P}(X \in A_1) . \mathbb{P}(Y \in B_1)$$

Donc

$$\mathbb{P}((X_1 = a) \cap (Y_1 = b)) = \mathbb{P}(X_1 = a) . \mathbb{P}(Y_1 = b)$$

5.3 Famille de variables aléatoires mutuellement indépendantes

5.3.1 Définition

Soit $(X_k)_{0 \leq k \leq n}$ une suite finie de variables aléatoires discrètes définies sur le même espace probabilisé $(\Omega, T, \mathbb{P})$; on dit qu'elles sont indépendantes, ou mutuellement indépendantes, lorsque pour tout $(x_0, ..., x_n)$

$$\mathbb{P}\left(\bigcap_{k=0}^{n} (X_k = x_k)\right) = \prod_{k=0}^{n} \mathbb{P}(X_k = x_k)$$

Remarque

Il n'est pas nécessaire de parler de sous-famille dans la définition.

Par contre, toute sous-famille d'une famille de variables aléatoires discrètes mutuellement indépendantes est également mutuellement indépendante.

5.3.2 Définition

Famille de variables aléatoires mutuellement indépendantes : si c'est vérifié pour toute sous-famille finie.

5.3.3 Lemme des coalitions

Si $X_1, X_2, ...X_n$ sont des variables aléatoires mutuellement indépendantes.

Alors pour tout m compris entre 1 et $n-1$, et toutes fonctions f et g, les variables $f(X_1, ..., X_m)$ et $g(X_{m+1}, ..., X_n)$ sont indépendantes.

Démonstration non exigible

Par exemple

Si $X_1, X_2, ...X_n$ sont des variables aléatoires mutuellement indépendantes, alors

X_1 et $X_2 + X_3$ sont indépendantes,
$X_1.X_4$ et $X_2 + X_3$ sont indépendantes,
$X_1.X_2.X_3$ et $\exp(X_4.X_5)$ sont indépendantes...

6 Espérance

6.1 Définition

6.1.1 Variables positives

Si X est une variable aléatoire à valeurs dans $\mathbb{R}+$, l'espérance $E(X)$ est la somme dans $[0,+\infty]$ (finie ou non), de la famille $(x.\mathbb{P}(X=x))_{x\in X(\Omega)}$.

$$E(X)=\sum_{x\in X(\Omega)} x.\mathbb{P}(X=x)$$

6.1.2 Variables réelles

Si X est une variable aléatoire réelle, la variable aléatoire X est d'espérance finie si la famille

$$(x.\mathbb{P}(X=x))_{x\in X(\Omega)}$$

est sommable ; dans ce cas, la somme de cette famille est l'espérance de X.

6.2 Propriétés

6.2.1 Positivité

Théorème

Si $X\geq 0$, alors $E(X)\geq 0$: découle de la définition.

6.2.2 Valeur absolue

Théorème

X est d'espérance finie si et seulement si $|X|$ est d'espérance finie.

Démonstration

Notons $Y=|X|$ et (y_j) les valeurs prises par Y.

Y est d'espérance finie si la somme suivante est finie :

$$E(Y)=\sum_j y_j.\mathbb{P}(Y=y_j)=\sum_j y_j.\left(\mathbb{P}(X=y_j)+\mathbb{P}(X=-y_j)\right)=\sum_i |x_i|.\mathbb{P}(X=x_i)$$

Donc si X est d'espérance finie, d'après le théorème de sommation par paquets.

Remarque

Dans ce cas, $|E(X)|\leq E(|X|)$.

Démonstration

$$|E(X)|=\left|\sum_i x_i.\mathbb{P}(X=x_i)\right|\leq\sum_i |x_i|.\mathbb{P}(X=x_i)=E(|X|)$$

A nouveau d'après le théorème de sommation par paquets.

C'est un cas particulier très simple de la formule de transfert.

6.2.3 Comparaison

Théorème

On suppose que $0\leq X\leq Y$ et Y est d'espérance finie.

Alors X est d'espérance finie et

$$0\leq E(X)\leq E(Y)$$

Démonstration

On note $(x_i)_{i\in I}$ les valeurs prises par X et $(y_j)_{j\in J}$ les valeurs prises par Y.

En utilisant la sommation par paquets :

$$E(Y) = \sum_{i,j} b_{i,j}$$

où

$$b_{i,j} = y_j.\mathbb{P}\left((X = x_i) \cap (Y = y_j)\right)$$

De même :

$$E(X) = \sum_{i,j} a_{i,j}$$

où

$$a_{i,j} = x_i.\mathbb{P}\left((X = x_i) \cap (Y = y_j)\right)$$

Il reste à justifier que pour tout couple (i, j), $0 \leq a_{i,j} \leq b_{i,j}$.

Corollaire

Si $|X| \leq Y$ et si Y est d'espérance finie, alors X est d'espérance finie.

6.2.4 Linéarité

Théorème

L'espérance est linéaire sur l'espace vectoriel des variables aléatoires d'espérance finie définies sur Ω.

Démonstration

$E(\lambda.X) = \lambda.E(X)$: immédiat ; pour l'espérance d'une somme, voir la démonstration plus loin.

6.2.5 Monotonie

Théorème

L'espérance est croissante sur l'espace vectoriel des variables aléatoires d'espérance finie définies sur Ω.

Ce qui signifie que si $X \leq Y$, alors $E(X) \leq E(Y)$. Découle de la linéarité et de la positivité.

6.2.6 Remarque

Si X est constante : $X = c$, alors $E(X) = c$.

En particulier

Si a et b sont des réels :

$$E(aX + b) = a.E(X) + b$$

6.2.7 Variables centrées

Une variable X est centrée si son espérance est nulle ; pour toute variable aléatoire X, la variable aléatoire suivante est centrée :

$$X_c = X - E(X)$$

6.3 L'exemple des deux dés

On lance deux dés, un blanc et un rouge ; l'univers est $\Omega = \{1, 2, 3, 4, 5, 6\}^2$.

S est la somme des deux chiffres.

$\mathbb{P}(S = 4)$? $\mathbb{P}(S = 5)$? loi de S ? $E(S)$?

Réponse

$$\mathbb{P}(S=4)=\frac{3}{36}\,;\,\mathbb{P}(S=5)=\frac{4}{36}$$

Pour un dé, $E(X)=3,5$; d'où $E(S)=7$.

6.4 Exercice

Un sauteur en hauteur tente de franchir successivement les hauteurs 1,2,3,... ; la probabilité de succès à la hauteur n est $\frac{1}{n}$; il est éliminé au premier échec ; X est le nombre de sauts réussis. Déterminer la loi de X et son espérance.

Réponse

$\mathbb{P}(X\geq n)=\frac{1}{n!}$; puis $\mathbb{P}(X=n)=\frac{1}{n!}-\frac{1}{(n+1)!}=\frac{n}{(n+1)!}$

$$E(X)=\sum_{n=1}^{\infty}n\left(\frac{1}{n!}-\frac{1}{(n+1)!}\right)=\sum_{n=1}^{\infty}\frac{n}{n!}-\sum_{n=1}^{\infty}\frac{n}{(n+1)!}$$

$$E(X)=\sum_{n=1}^{\infty}\frac{n}{n!}-\sum_{n=2}^{\infty}\frac{n-1}{n!}=\sum_{n=1}^{\infty}\frac{1}{n!}=e-1$$

Voir paragraphe **15**.

6.5 Cas de la loi binomiale

6.5.1 Loi de Bernoulli

$\mathbb{P}(X=k)=p$ si $k=1$, $1-p$ si $k=0$.
$E(X)=1.p+0.(1-p)$, donc

$$E(X)=p$$

6.5.2 Loi binomiale

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

Pour calculer l'espérance, on peut utiliser la définition :

$$E(X)=\sum_{k=0}^{n}k.\mathbb{P}(X=k)=\sum_{k=0}^{n}k.\binom{n}{k}.p^k.(1-p)^{n-k}$$

mais il est plus simple d'utiliser $X=X_1+X_2+...+X_n$.

Conclusion

$$E(X)=np$$

6.6 Formule de transfert

6.6.1 Enoncé

Soit X une variable aléatoire discrète et f une fonction définie sur $X(\Omega)$ à valeurs dans $\mathbb{R}$; alors $Y=f(X)$ est d'espérance finie si et seulement si la famille $f(x).\mathbb{P}(X=x)$ est sommable ; si tel est le cas :

$$E(f(X))=\sum_{x\in X(\Omega)}f(x).\mathbb{P}(X=x)$$

6.6.2 Remarque

La variable décrit $X(\Omega)$ au lieu de $Y(\Omega)$ dans la définition de l'espérance : la somme est plus 'éclatée'.

6.6.3 Un cas particulier : $X = \mathrm{Id}$

On suppose Ω fini ou dénombrable, et $X = \mathrm{Id}$; donc $Y = f \circ X = f$.

Dans ce cas :

$$E(Y) = \sum_{y \in Y(\Omega)} y.\mathbb{P}(Y = y) = \sum_{\omega \in \Omega} Y(\omega).p_\omega$$

Dans ce cas particulier, on a donc une autre définition de l'espérance.

6.6.4 Démonstration rapide (non exigible)

Il s'agit d'une application de la sommation par paquets :

- soit $Y = f(X) = f \circ X$; soit $y \in Y(\Omega)$.
- soit $(x_i)_{i \in I}$ la famille des antécédents de y par f.

$$(Y = y) = \bigcup_{i \in I} (X = x_i)$$

Union disjointe, donc

$$\mathbb{P}(Y = y) = \sum_{i \in I} \mathbb{P}(X = x_i)$$

puis

$$y.\mathbb{P}(Y = y) = \sum_{i \in I} f(x_i).\mathbb{P}(X = x_i)$$

Il reste à sommer sur y...

6.7 Espérance d'une somme

Théorème

Soit X, Y deux variables aléatoires réelles d'espérances finies ; soit

$$S = X + Y$$

Alors S est d'espérance finie, et

$$E(X + Y) = E(X) + E(Y)$$

Démonstration

Commençons par le cas où $X \geq 0$ et $Y \geq 0$. Notons

$$Z = (X, Y)$$

et

$$f : (x, y) \to x + y$$

de $\mathbb{R}^2$ dans $\mathbb{R}$; donc $S = f(Z)$.

$$E(f(Z)) = \sum_{z \in Z(\Omega)} f(z).\mathbb{P}(Z = z)$$

ou encore :

$$E(S) = \sum_{x,y} (x + y).\mathbb{P}[(X, Y) = (x, y)]$$

On sépare la somme en deux. Ensuite :

$$E(X) = E(p_1(Z)) = \sum_{x,y} x.\mathbb{P}[(X, Y) = (x, y)]$$

pour les mêmes raisons :

on a simplement remplacé f par

$$p_1 : (x, y) \to x$$

L'autre somme vaut $E(Y)$.

Le cas général

Le même calcul est valable car les familles sont sommables.

6.8 Espérance d'un produit

Si X et Y sont deux variables aléatoires indépendantes d'espérances finies, alors XY est d'espérance finie, et

$$E(XY) = E(X)E(Y)$$

6.8.1 Démonstration

Supposons X et Y positives ; on utilise la formule de transfert : soit

$$Z = (X, Y)$$

Soit $f : (x, y) \to x.y$; on sait que

$$E(f(Z)) = \sum_{z \in Z(\Omega)} f(z) . \mathbb{P}(Z = z)$$

ou encore :

$$E(f(Z)) = \sum_{x,y} f(x, y) . \mathbb{P}[(X, Y) = (x, y)]$$

X et Y étant indépendantes :

$$E(f(Z)) = \sum_{x,y} x.y.\mathbb{P}(X = x) . \mathbb{P}(Y = y)$$

On termine avec le théorème de sommation par paquets.

Le cas général est analogue ; ne pas oublier de vérifier la sommabilité.

6.8.2 Réciproque

Fausse ! Exemple : soit X uniforme sur $\{0, 1\}$ et Z uniforme sur $\{-1, 1\}$, indépendantes, et

$$Y = ZX$$

a) Décrire la loi de Y.
b) Montrer que X et Y ne sont pas indépendantes.
c) Montrer que $E(XY) = E(X)E(Y)$.
d) Y et Z sont-elles indépendantes ?

Réponses

a) $\mathbb{P}(Y = 0) = \frac{1}{2}$; $\mathbb{P}(Y = 1) = \mathbb{P}(Y = -1) = \frac{1}{4}$.
b) $\mathbb{P}(X = 0, Y = 0) = \mathbb{P}(X = 0) = \frac{1}{2}$, mais $\mathbb{P}(X = 0) . \mathbb{P}(Y = 0) = \frac{1}{4}$.
c) $E(Y) = 0$; $XY = Y$.
d) Non : $\mathbb{P}[(Z = 1), (Y = -1)] = 0$.

6.9 Fonction indicatrice

6.9.1 Espérance

Soit A un événement ; on appelle fonction indicatrice de A :

$$1_A : \Omega \to \{0, 1\}$$

définie par $1_A(\omega) = 1$ si $\omega \in A$ et 0 si $\omega \notin A$.

Montrer que c'est une variable aléatoire et trouver son espérance.

Réponse

Il est clair que $(1_A = 1) = A$ et $(1_A = 0) = \Omega \setminus A$; d'où :

$$E(1_A) = \mathbb{P}(A)$$

6.9.2 Indépendance

Soit A et B deux événements ; montrer qu'ils sont indépendants si et seulement si 1_A et 1_B le sont.

7 Variance, écart-type

7.1 Très utile

Quelques inégalités :
-$\forall x, y \in \mathbb{R}, |x.y| \leq \frac{1}{2}\left(x^2 + y^2\right)$
-$\forall x \in \mathbb{R}, |x| \leq 1 + x^2$
-$\forall x, y \in \mathbb{R}, (x+y)^2 \leq 2\left(x^2 + y^2\right)$

7.2 Moments

7.2.1 Définition

Soit X une variable aléatoire réelle. Pour $k \in \mathbb{N}^*$, le moment d'ordre k de X est $E\left(X^k\right)$, s'il existe.

D'après la formule de transfert :

$$E\left(X^k\right) = \sum_{x \in X(\Omega)} x^k.\mathbb{P}(X = x)$$

7.2.2 Théorème

Si $E\left(X^2\right)$ existe, alors $E(X)$ existe.

Démonstration

$\forall x \in \mathbb{R}, |x| \leq 1 + x^2$.

Remarque

C'est un cas particulier de l'inégalité de Cauchy-Schwarz qu'on verra plus loin.

7.2.3 Exercice

Soit $p, q \in \mathbb{N}$; si $0 \leq q \leq p$ et si $E(X^p)$ existe, alors $E(X^q)$ existe.

Démonstration

$\forall x \in \mathbb{R}, |x|^q \leq 1 + |x|^p$.

7.3 Variance

7.3.1 Définition

Soit X une variable aléatoire réelle ; on suppose que $E\left(X^2\right)$ existe ; alors $m = E(X)$ existe.

$E\left(X_c^2\right)$ existe également car

$$|X_c|^2 \leq 2\left(X^2 + m^2\right)$$

On pose :

$$V(X) = E\left[(X - E(X))^2\right] = E\left(X_c^2\right)$$

Remarque : $(X_c)_c = X_c$, donc

$$V(X) = V(X_c)$$

7.3.2 Ecart-type

$$\sigma(X) = \sqrt{V(X)}$$

On dit qu'une variable aléatoire réelle est réduite si $V(X) = 1$, ou $\sigma(X) = 1$.

7.3.3 Propriétés

Quelques propriétés de la variance :
- $V(X) = E\left(X^2\right) - E(X)^2$
- $V(aX+b) = a^2.V(X)$ si a et b sont des constantes réelles
- Si $V(X) > 0$, alors

$$\frac{X - E(X)}{\sigma(X)} = \frac{X_c}{\sigma(X)}$$

est centrée réduite

7.3.4 Exercice : que signifie $V(X) = 0$?

Réponse

$\{\omega \in \Omega / X(\omega) \neq m\}$ est de probabilité nulle, soit $\mathbb{P}(X \neq m) = 0$.

On dit que X est constante presque partout.

7.4 Inégalité de Cauchy-Schwarz

7.4.1 Théorème

Si X et Y admettent chacune un moment d'ordre 2, alors XY est d'espérance finie et

$$[E(XY)]^2 \leq E\left(X^2\right) E\left(Y^2\right)$$

Remarque

L'existence de $E(XY)$ découle facilement de la formule :

$$\forall x, y \in \mathbb{R}, |x.y| \leq \frac{1}{2}\left(x^2 + y^2\right)$$

Démonstration

Le cas où $E\left(X^2\right) = 0$ ou $E\left(Y^2\right) = 0$ est facile.

Dans le cas contraire, on se ramène au cas où $E\left(X^2\right) = E\left(Y^2\right) = 1$; des relations
$E\left[(X+Y)^2\right] \geq 0$ et $E\left[(X-Y)^2\right] \geq 0$, on déduit

$$-1 \leq E(XY) \leq 1$$

ce qui permet de conclure.

7.4.2 Cas d'égalité

Il y a égalité :

$$[E(XY)]^2 = E\left(X^2\right) E\left(Y^2\right)$$

si et seulement si, presque sûrement, $X = 0$, ou

$$\exists \lambda \in \mathbb{R}, Y = \lambda X$$

7.4.3 Corollaire

L'ensemble des variables aléatoires définies sur Ω admettant un moment d'ordre 2 constitue un espace vectoriel.

Remarque

Ce résultat découle aussi directement de :

$$\forall x, y \in \mathbb{R}, (x+y)^2 \leq 2\left(x^2+y^2\right)$$

8 Fonction génératrice de la variable aléatoire X à valeurs dans $\mathbb{N}$

8.1 Définition

$$G_X(t) = E\left(t^X\right) = \sum_{k=0}^{\infty} \mathbb{P}(X=k)\, t^k$$

Plus précisément

D'après la formule de transfert, pour tout réel t :

$E\left(t^X\right)$ existe si et seulement si la série $\sum_{k=0}^{\infty} \mathbb{P}(X=k)\, t^k$ est absolument convergente, et dans ce cas il y a égalité.

Le rayon de convergence

Le rayon de convergence R est supérieur ou égal à 1 ; la série converge normalement sur le disque fermé de centre 0 et de rayon 1 ; G_X est donc continue sur $[-1,1]$. Réciproquement :

$$\mathbb{P}(X=k) = \frac{1}{k!} G_X^{(k)}(0)$$

La loi de X est donc déterminée par G_X.

8.2 Un lemme sur les séries entières

Soit (a_n) une suite de réels *positifs* ; soit

$$f(x) = \sum_{n=0}^{\infty} a_n . x^n$$

On suppose $R=1$.

- Si $\sum a_n$ converge, la série converge normalement sur $[0,1]$.
- Si $\sum a_n$ diverge, f tend vers $+\infty$ en 1^-.

Démonstration

Notons $s_n = \sum_{k=0}^{n} a_k$ et $f_n(x) = \sum_{k=0}^{n} a_k . x^k$. Supposons que $\sum a_n$ diverge.

Fixons $M>0$. Il existe alors n tel que $s_n > M$; donc $f_n(1) > M$; f_n étant continue sur $\mathbb{R}$, $f_n > M$ sur un voisinage de 1 ; or $f \geq f_n$ sur $]0,1[$; donc :

$$\exists \delta > 0, \forall t \in \left]1-\delta, 1\right[, f(t) \geq M$$

Ce qui prouve que f tend vers $+\infty$ en 1^-.

8.3 Calcul des moments de X

On notera $a_k = \mathbb{P}(X = k)$.

$$\forall t \in \left]-R, R\right[, \forall q \in \mathbb{N}, G_X^{(q)}(t) = \sum_{k=q}^{\infty} k(k-1)...(k-q+1).a_k.t^{k-q}$$

$G_X^{(q)}(1) = E\left(X\left(X-1\right)...\left(X-q+1\right)\right)$ si $R > 1$.

8.3.1 Espérance

Théorème

La variable aléatoire X est d'espérance finie si et seulement si G_X est dérivable en 1 ; dans ce cas

$$E(X) = G_X'(1)$$

Démonstration

Si $E(X)$ existe, $\sum ka_k$ converge ; la série dérivée converge normalement sur $[0,1]$. On en déduit que G_X est de classe C^1 sur $[0,1]$.
Si $\sum ka_k$ diverge : G_X' tend vers $+\infty$ en 1^- d'après le lemme ; on termine avec le théorème de la limite de la dérivée qui prouve que G_X n'est pas dérivable en 1.

8.3.2 Variance

Théorème

La variable aléatoire X admet un second moment si et seulement si G_X est deux fois dérivable en 1 ; dans ce cas :

$$G_X''(1) = E\left(X\left(X-1\right)\right)$$

$$V(X) = E\left(X.\left(X-1\right)\right) + E(X) - E(X)^2$$

Donc

$$V(X) = G_X''(1) + G_X'(1) - G_X'(1)^2$$

Démonstration

Analogue.

Commentaire

Les étudiants doivent savoir retrouver l'expression de la variance de X à l'aide de $G_X''(1)$ et $G_X'(1)$.

8.4 Fonction génératrice d'une somme finie de variables aléatoires indépendantes à valeurs dans $\mathbb{N}$

8.4.1 Théorème

Soit X, Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$.

Alors, sur $[-1,1]$:

$$G_{X+Y} = G_X.G_Y$$

Démonstration

$$\mathbb{P}(X+Y=n) = \sum_{k=0}^{n} \mathbb{P}\left[(X=k) \cap (Y=n-k)\right] = \sum_{k=0}^{n} \mathbb{P}(X=k).\mathbb{P}(Y=n-k)$$

Puis produit de Cauchy de deux séries entières.

Validité

La formule
$$G_{X+Y}(t) = G_X(t).G_Y(t)$$
est vérifiée au moins pour $t \in [-1, 1]$; on rappelle que le rayon de convergence d'un produit est ???

8.4.2 Autre démonstration

$G_{X+Y}(t) = E\left(t^{X+Y}\right) = E\left(t^X\right).E\left(t^Y\right)$ car ?

8.4.3 Généralisation

Se généralise à un nombre fini de variables aléatoires mutuellement indépendantes à valeurs dans $\mathbb{N}$.

8.5 Cas de la loi binomiale

8.5.1 Loi de Bernoulli $B(p)$

$\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$.

On note $q = 1 - p$.

$$E(X) = p$$
$$V(X) = pq$$
$$G_X(t) = pt + q$$

8.5.2 Loi binomiale $B(n, p)$

$$E(X) = np$$
$$V(X) = npq$$
$$G_X(t) = (pt + q)^n$$

9 Loi géométrique

9.1 Définition

9.1.1 Pour p dans]0,1[, loi géométrique de paramètre p

La variable aléatoire X suit la loi $G(p)$ si :
$$\forall k \geq 1, \mathbb{P}(X = k) = p.(1 - p)^{k-1} = p.q^{k-1}$$
On note souvent $q = 1 - p$.

Interprétation : instant de premier succès dans une suite d'épreuves de Bernoulli mutuellement indépendantes de paramètre p.

9.1.2 Exemple

On lance un dé ; X est le nombre de lancers nécessaires pour obtenir le 6.
$$\mathbb{P}(X = n) = \frac{1}{6}.\left(1 - \frac{1}{6}\right)^{n-1}$$

9.2 Caractéristiques

$G_X(t) = \frac{pt}{1-qt}$; $E(X) = \frac{1}{p}$, $V(X) = \frac{q}{p^2}$.

9.3 Exercice

Un logiciel comporte 5 erreurs. A chaque operation de maintenance, toute erreur a la probabilité $p = \frac{1}{3}$ d'être corrigée. Les corrections sont indépendantes les unes des autres ; quand elles sont réussies, elles font disparaître l'erreur et n'en font pas réapparaître d'autres.

Combien d'opérations de maintenance faut-il faire sur le logiciel pour que la probabilité qu'il ne reste aucune faute soit au moins de 0.9 ?

Réponse

$\left(1 - \left(\frac{2}{3}\right)^n\right)^5 \geq 0.9$, soit $n \geq 10$.

9.4 Absence de mémoire des lois géométriques

9.4.1 $\mathbb{P}(X > k)$

$$\forall k \geq 0, \mathbb{P}(X > k) = q^k$$

Démonstration

$$\mathbb{P}(X > k) = 1 - \sum_{j=1}^{k} p.q^{j-1} = q^k$$

9.4.2 Théorème

$$\forall n \geq 0, \forall k \geq 0, \mathbb{P}(X > n + k \,|\, X > n) = \mathbb{P}(X > k)$$

Démonstration

$$\frac{\mathbb{P}((X > n + k) \cap (X > n))}{\mathbb{P}(X > n)} = \frac{q^{n+k}}{q^n} = q^k$$

9.4.3 Exercice

Ce sont les seules lois sur $\mathbb{N}^*$ sans mémoire.

Démonstration

Pour $k = 1$, on trouve, pour tout $n \geq 1$:

$$\mathbb{P}(X > n + 1) = \mathbb{P}(X > n).\mathbb{P}(X > 1)$$

On note $p = \mathbb{P}(X = 1)$, $q = \mathbb{P}(X > 1)$; on obtient

$$\mathbb{P}(X > n) = q^n.\mathbb{P}(X > 0) = q^n$$

D'où

$$\mathbb{P}(X = n) = q^{n-1} - q^n = p.q^{n-1}$$

10 Loi de Poisson

10.1 Définition

Loi de Poisson de paramètre $\lambda > 0$:

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}.\lambda^k}{k!}$$

$$G_X(t) = e^{\lambda(t-1)},\ E(X) = \lambda = V(X)$$

Notation :

$$X \sim \mathcal{P}(\lambda)$$

Remarque

Les étudiants doivent savoir calculer la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson.

Les espérances et variances doivent être connues.

10.2 Approximation de la loi binomiale par la loi de Poisson

Théorème

Si, pour tout n, $X_n \sim B(n, p_n)$ et si $(n.p_n)$ converge vers $\lambda > 0$, alors :

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

Remarque

On dit que (X_n) converge en loi : $X_n \xrightarrow{\mathcal{L}} X$, où $X \sim \mathcal{P}(\lambda)$.

Démonstration

Si $\geq k$:

$$\mathbb{P}(X_n = k) = \frac{1}{k!}.\frac{n(n-1)...(n-k+1)}{n^k}.(n.p_n)^k.(1-p_n)^{n-k}$$

10.3 Interprétation de la loi de Poisson comme loi des événements rares

Un très grand nombre N de personnes jouent au loto ; chacune a une probabilité p très faible de gagner ; que dire du nombre de gagnants ?

Réponse

Il suit une loi binomiale d'espérance

$$\lambda = Np$$

Loi qu'on peut approcher par une loi de Poisson de paramètre λ.

Que dire de l'espérance et de la variance de ces deux lois ?

Réponse

Les deux lois ont même espérance λ.

Variance : λ pour la loi de Poisson, Npq pour la loi binomiale, très proche de λ.

10.4 Somme de deux variables aléatoires réelles indépendantes suivant une loi de Poisson

Exercice

Si X et Y sont indépendantes, $X \sim \mathcal{P}(\lambda)$, $Y \sim \mathcal{P}(\mu)$, alors

$$X + Y \sim \mathcal{P}(\lambda + \mu)$$

1e méthode

Simple calcul, voir banque CCP.

2e méthode

On utilise $G_{X+Y} = G_X.G_Y$, et $G_X(t) = e^{\lambda(t-1)}$.

10.5 Exercice : les points fixes et les dérangements

10.5.1 Dérangements

Définition

On appelle dérangement toute permutation sans point fixe.

Exercice

Le nombre de dérangements de $[\![1,n]\!]$ est

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

Démonstration rapide

On pose

$$f(x) = \sum_{n=0}^{\infty}\frac{d_n}{n!}x^n$$

avec la convention $d_0 = 1$.

On vérifie que

$$\forall n \geq 0,\ n! = \sum_{k=0}^{n}\binom{n}{k}d_k$$

On en déduit :

$$\forall n \geq 0,\ 1 = \sum_{k=0}^{n}\frac{d_k}{k!}.\frac{1}{(n-k)!}$$

puis

$$\forall x \in \left]-1,1\right[,\ \frac{1}{1-x} = e^x.f(x)$$

Conclusion :

$$\forall x \in \left]-1,1\right[,\ f(x) = \frac{e^{-x}}{1-x}$$

10.5.2 Un problème de cheval

n élèves se rendent au lycée à cheval ; en partant, saoulés par le cours de maths, ils prennent un cheval au hasard ; quelle est la probabilité p_n pour qu'aucun ne parte avec son propre canasson ?

Quelle est la limite de (p_n) ?

Réponse

$$p_n = 1 - \frac{1}{1!} + \frac{1}{2!} + ... + \frac{(-1)^n}{n!} = \sum_{j=0}^{n}\frac{(-1)^j}{j!}$$

C'est la proportion de permutations de $[1,n]$ sans points fixes. Limite : $\frac{1}{e}$.

10.5.3 Généralisation

Soit X_n le nombre d'élèves qui partent avec le bon cheval ; quelle est l'espérance de X_n ?

On note

$$q_n(k) = \mathbb{P}(X_n = k)$$

Déterminer $q_n(k)$ et sa limite, pour $k \geq 0$ fixé, quand n tend vers l'infini.

Réponse

Pour n fixé, on définit Y_j par $Y_j = 1$ si le $j-$ième élève a pris le bon cheval et $Y_j = 0$ sinon.

$E\left(Y_j\right) = \frac{1}{n}$. On en déduit que

$$E\left(X_n\right) = 1$$

$$q_n\left(k\right) = \frac{1}{k!}\left(1 - \frac{1}{1!} + ... + \frac{\left(-1\right)^{n-k}}{\left(n-k\right)!}\right) = \frac{1}{k!}.p_{n-k}$$

Limite :

$$\frac{1}{e.k!}$$

On reconnaît une loi de Poisson. On dit que $\left(X_n\right)$ converge en loi.

10.5.4 Variance et fonction génératrice

Calculer $V\left(X_n\right)$ à l'aide des Y_i, puis à l'aide de $G_{X_n} = G_n$.

Réponse

Si $i \neq j$: $E\left(Y_i.Y_j\right) = \frac{1}{n(n-1)}$, d'où $E\left(X_n^2\right) = 2$, d'où

$$V\left(X_n\right) = 1$$

Autre méthode :

$$G_n\left(t\right) = \sum_{k=0}^{n} \frac{p_{n-k}}{k!}.t^k$$

On en déduit facilement que

$$G_n^{'} = G_{n-1}$$

D'où $G_n\left(1\right) = G_n'\left(1\right) = G_n''\left(1\right) = 1$.

On retrouve $V\left(X_n\right)$ et

$$G_n\left(t\right) = \sum_{k=0}^{n} \frac{\left(t-1\right)^k}{k!}$$

ce qu'on peut aussi vérifier avec la formule du binôme en permutant une somme double.

11 Covariance

11.1 Définition

Soit X, Y deux variables aléatoires réelles définies sur Ω possèdant une variance.

Soit $X_c = X - E(X)$ et $Y_c = Y - E(Y)$; on pose

$$\text{Cov}(X, Y) = E\left(X_c Y_c\right)$$

Remarque

$\text{Cov}(X, Y) = \text{Cov}(X_c, Y_c)$; $\text{Cov}(X, X) = V(X)$.

11.1.1 Propriétés

Soit X, Y deux variables aléatoires réelles définies sur Ω.

a) $\text{Cov}(X, Y) = \text{Cov}(X + a, Y)$ pour tout réel a

b) La covariance est bilinéaire symétrique

c) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

11.1.2 Théorème

Si X et Y sont indépendantes, $\mathrm{Cov}(X,Y)=0$; la réciproque est fausse.

11.2 Exemples

11.2.1 Exemple 1

Soit $n\geq 1$, X uniforme sur $[\![-n,n]\!]$, et $Y=|X|$.
Montrer que X et Y ne sont pas indépendantes, mais que $\mathrm{Cov}(X,Y)=0$.

Réponse

$\mathbb{P}\left(X=0,Y=1\right)=0\neq\mathbb{P}\left(X=0\right).\mathbb{P}\left(Y=1\right)$.

$$E\left(X\right)=0=E\left(XY\right)$$

donc $\mathrm{Cov}\left(X,Y\right)=0$.

11.2.2 Exemple 2

$\Omega=\left\{\left(1,0\right),\left(0,1\right),\left(-1,0\right),\left(0,-1\right)\right\}$, probabilité uniforme ; X et Y les fonctions coordonnées ; $\mathrm{Cov}(X,Y)$? X et Y sont-elles indépendantes ?
Que dire de $X+Y$ et $X-Y$?

Réponse

$E\left(X\right)=E\left(Y\right)=0$; $XY=0$; $\mathrm{Cov}\left(X,Y\right)=0$.
Pour l'indépendance, examiner

$$\mathbb{P}\left[\left(X=0\right)\cap\left(Y=0\right)\right]$$

$X+Y$ et $X-Y$ sont indépendantes ; par exemple :

$$\mathbb{P}\left[\left(X+Y=1\right)\cap\left(X-Y=1\right)\right]=\frac{1}{4}=\mathbb{P}\left(X+Y=1\right).\mathbb{P}\left(X-Y=1\right)$$

11.2.3 Exemple 3

T uniforme sur

$$\left\{\frac{k\pi}{2n}/-n\leq k\leq n\right\}$$

$X=\cos(T)$, $Y=\sin(T)$.
Montrer que X et Y ne sont pas indépendantes, mais que $\mathrm{Cov}(X,Y)=0$.
Calculer $E\left(X\right)$.

Réponse

Par symétrie :

$$E\left(Y\right)=\frac{1}{2n+1}.\sum_{k=-n}^{n}\sin\frac{k\pi}{2n}=0$$

$$E\left(XY\right)=E\left(\frac{1}{2}\sin 2T\right)=0$$

Pour l'indépendance, examiner

$$\mathbb{P}\left[\left(X=0\right)\cap\left(Y=0\right)\right]$$

$$E\left(X\right)=\frac{1}{2n+1}\cot\frac{\pi}{4n}$$

11.3 Variance d'une somme

11.3.1 Théorème

$V(X+Y) = V(X) + V(Y) + 2.\mathrm{Cov}(X,Y)$

Si X et Y sont indépendantes, $\mathrm{Cov}(X,Y) = 0$, donc

$$V(X+Y) = V(X) + V(Y)$$

Remarque

Espérance d'une somme ?

11.3.2 Généralisation

Se généralise à une somme de variables aléatoires réelles indépendantes **deux à deux** :

$$V(X_1 + X_2 + ...X_n) = \sum_{k=1}^{n} V(X_k)$$

Démonstration

Par récurrence sur n ; la covariance étant bilinéaire,

$$\mathrm{Cov}(X_1 + ...X_{n-1}, X_n) = 0$$

Néanmoins, $X_1 + ... + X_{n-1}$ et X_n peuvent ne pas être indépendantes.

11.3.3 Variance d'une variable aléatoire binomiale

$$V(X) = npq$$

Explication : c'est la variance de la somme de n variables de Bernoulli indépendantes.

11.4 Complément : coefficient de corrélation

D'après l'inégalité de Cauchy-Schwarz :

$$[\mathrm{Cov}(X,Y)]^2 = [E(X_c Y_c)]^2 \leq E\left(X_c^2\right) E\left(Y_c^2\right) = V(X).V(Y)$$

Si ces variances sont non nulles, on pose

$$\mathrm{Cor}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

qui est un nombre dans $[-1, 1]$.

12 Inégalités

12.1 Inégalité de Markov

Théorème

Si $X \geq 0$, et si $E(X)$ existe, alors, pour tout réel $a > 0$:

$$\mathbb{P}(X \geq a) \leq \frac{E(X)}{a}$$

Démonstration

$$E(X) \geq \sum_{x \geq a} x.\mathbb{P}(X = x) \geq a. \sum_{x \geq a} \mathbb{P}(X = x) = a.\mathbb{P}(X \geq a)$$

12.2 Inégalité de Bienaymé-Tchebychev

12.2.1 Enoncé

Théorème

On suppose que X admet une espérance $m = E(X)$ et une variance σ^2 ; alors, pour tout $a > 0$

$$\mathbb{P}(|X - m| \geq a) \leq \frac{\sigma^2}{a^2}$$

12.2.2 Démonstration

Application du précédent à $|X - m|^2$.

12.2.3 Généralisation

Exercice

Pour tout réel c

$$\mathbb{P}(|X - c| \geq a) \leq \frac{\sigma^2 + (c - m)^2}{a^2}$$

Démonstration

$$\mathbb{P}(|X - c| \geq \varepsilon) = \mathbb{P}(|X - c|^2 \geq \varepsilon^2) \leq \frac{E\left[(X-c)^2\right]}{\varepsilon^2}$$

Or

$(X - c)^2 = (X - m + m - c)^2 = (X - m)^2 + (m - c)^2 + 2.(X - m)(m - c)$

et $E(X - m) = 0$.

12.3 Complément : l'inégalité de Cantelli-Tchebychev (one-sided)

Exercice

Si X admet une espérance $m = E(X)$ et une variance σ^2, et si $a > 0$, alors

$$\mathbb{P}(X - m \geq a) \leq \frac{\sigma^2}{a^2 + \sigma^2}$$

Remarque

Il suffit de traiter le cas où $E(X) = m = 0$.

12.3.1 Méthode 1

Premier essai :

$$\mathbb{P}(X \geq a) \leq \mathbb{P}(|X| \geq a) = \mathbb{P}\left(X^2 \geq a^2\right) \leq \frac{\sigma^2}{a^2}$$

d'après ?

Résultat intéressant mais pas assez précis. On va généraliser le calcul précédent en remplaçant l'intervalle $I = [-a, a]$ par

$$I = [c - r, c + r]$$

avec $c < a$ et $r = a - c$, de sorte que

$$c + r = a$$

Le calcul devient

$$\mathbb{P}(X \geq a) \leq \mathbb{P}(|X-c| \geq r) = \mathbb{P}\left((X-c)^2 \geq r^2\right) \leq \frac{E\left((X-c)^2\right)}{r^2} = \frac{\sigma^2+c^2}{(a-c)^2}$$

Que faire ensuite ?

On peut par dérivation chercher le minimum de

$$c \to \frac{\sigma^2+c^2}{(a-c)^2}$$

On peut aussi directement chercher c tel que

$$\frac{\sigma^2+c^2}{(a-c)^2} = \frac{\sigma^2}{a^2+\sigma^2}$$

On trouve

$$c = -\frac{\sigma^2}{a}$$

12.3.2 Méthode 2

On applique l'inégalité de Cauchy-Schwarz à $Y = 1_{X<a}$ et $Z = a - X$.

$$a = E(a-X) \leq E((a-X)Y)$$

Avec l'inégalité de Cauchy-Schwarz :

$$a^2 \leq E\left((a-X)^2\right) E\left(Y^2\right)$$

$E\left(Y^2\right) = E(Y) = \mathbb{P}(X<a)$; $E\left(Z^2\right) = a^2+\sigma^2$.

12.4 Exemple

On lance un dé. Majorer $\mathbb{P}(X=6)$ en utilisant les propriétés précédentes.

a) $V(X) = \frac{35}{12}$; $\mathbb{P}(X=6)$? $\frac{1}{6} \approx 0,17$

b) D'après l'inégalité de Markov :

$$\mathbb{P}(X=6) = \mathbb{P}(X \geq 6) \leq \frac{7}{12} \approx 0,58$$

c) Inégalité de Bienaymé-Tchebychev :

$$\mathbb{P}(X=6) \leq \mathbb{P}(X=6 \cup X=1) \leq \frac{7}{15} \approx 0,47$$

Mieux :

$$\mathbb{P}(X=6) = \frac{1}{2}.\mathbb{P}(X=6 \cup X=1) \leq \frac{7}{30} \approx 0,23$$

d) D'après l'inégalité de Cantelli :

$$\mathbb{P}(X=6) \leq \frac{7}{22} \approx 0,32$$

12.5 Loi faible des grands nombres

12.5.1 Théorème

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires deux à deux indépendantes, de même loi et admettant un moment d'ordre 2.

Soit $S_n = \sum_{k=1}^{n} X_k$ et $m = E(X_1)$; alors, pour tout $\varepsilon > 0$:

$$\lim_n \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) = 0$$

On dit que $\left(\frac{S_n}{n}\right)_{n\geq 1}$ converge vers m en probabilité.

12.5.2 Démonstration

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n.\varepsilon^2}$$

12.5.3 Exemple : lancers de dés

Que dire de $\mathbb{P}\left(\left|\frac{S_n}{n} - 3,5\right| \geq 0.01\right)$?

Peut-on dire que $\frac{S_n}{n} \to 3,5$?

Le théorème ne répond pas à cette question ; il existe un autre théorème qui répond que $\frac{S_n}{n} \to 3,5$ presque sûrement.

12.6 L'inégalité de Chernov

Exercice

Soit (X_n) une suite de variables aléatoires indépendantes valant ± 1 avec la probabilité $\frac{1}{2}$; soit $S_n = \sum_{k=1}^{n} X_k$.

1- Montrer que $\forall t \in \mathbb{R}, \text{ch}t \leq \exp\left(\frac{t^2}{2}\right)$.

2- Calculer $E\left(\exp\left(t.S_n\right)\right)$ pour t réel.

3- Montrer que

$$\forall R > 0,\, \mathbb{P}\left(\left|\frac{S_n}{n}\right| \geq R\right) \leq 2.\exp\left(-\frac{1}{2}nR^2\right)$$

4- Comparer au résultat donné par l'inégalité de Bienaymé-Tchebychev.

Réponse

1- Utiliser les développements en série entière.

2-

$$E\left(\exp\left(t.S_n\right)\right) = \text{ch}^n t$$

3- Appliquer l'inégalité de Markov à $\exp\left(t.S_n\right)$ avec $t > 0$.

13 Simulation de lois

13.1 Loi de Bernoulli $B(p)$

```
def bernoulli(p):
    return random() < p
```

13.2 Loi binomiale $B(n, p)$

```
def binomiale(n,p):
    return sum([bernoulli(p) for k in range(n)])
```

13.3 Loi géométrique $G(p)$

13.3.1 Méthode 1

```
def geometrique(p):
    geo = 1
    while not(bernoulli(p)):
      geo += 1
    return geo
```

13.3.2 Méthode 2

Soit U de loi uniforme sur $[0,1]$; remarquons que $V = 1 - U$ est aussi de loi uniforme sur $[0,1]$.

Notons

$$p_k = \mathbb{P}(X = k) = p.q^{k-1}$$

et

$$s_n = \sum_{k=1}^{n} p_k = \mathbb{P}(X \leq n) = 1 - q^n$$

Alors $0 < s_1 < s_2 < s_3 < ... < 1$, et

$$p_n = s_n - s_{n-1} = \mathbb{P}(s_{n-1} \leq U < s_n) = \mathbb{P}(s_{n-1} \leq V < s_n)$$

Donc

$$p_n = \mathbb{P}\left(q^n \leq U < q^{n-1}\right) = \mathbb{P}\left(n - 1 < \frac{\ln U}{\ln q} \leq n\right)$$

Donc

$$p_n = \mathbb{P}\left(\left\lceil \frac{\ln U}{\ln q} \right\rceil = n\right)$$

D'où le programme très simple :

```
def geometrique(p):
    return ceil(log(random()) / log(1-p))
```

13.4 Loi de Poisson

Analogue à ce qui précède ; mais ici, on ne dispose pas de formule simple pour s_n.

```python
from random import *
from math import *

def bernoulli(p):
    return int(random() < p)

def binomiale(n,p):
    return sum([bernoulli(p) for k in range(n)])

def geometrique_1(p):
    geo = 1
    while not(bernoulli(p)):
        geo += 1
    return geo

def geometrique_2(p):
    return ceil(log(random()) / log(1-p))

def poisson(ld):
    x = random()
    n = 0
    p = exp(-ld)
    while x > p:
        x -= p
        n += 1
        p *= ld / n
    return n

s = 0
for k in range(10000):
    s += poisson(5)
print(s/10000)  # estimation de l'espérance
```

14 Existence d'espaces probabilisés portant une suite de variables indépendantes de lois discrètes données

14.1 Théorème

Soit $(\mathcal{L}_n)$ une suite de lois discrètes ; il existe un espace probabilisé $(\Omega, T, \mathbb{P})$, et une suite (X_n) de variables aléatoires mutuellement indépendantes définies sur Ω tels que pour tout n, $\mathcal{L}_n$ soit la loi de X_n.

La démonstration est hors programme.

14.2 La mesure de Lebesgue

Ici, $\Omega = [0,1]$; on appelle tribu borélienne la plus petite tribu contenant les ouverts de $[0,1]$; existence ? On peut montrer qu'il existe une probabilité μ sur $[0,1]$ telle que pour tout intervalle I, $\mu(I)$ est sa largeur ; on l'appelle la mesure de Lebesgue.

Exercice

Toute partie A de $[0,1]$ dénombrable est un borélien ; $\mu(A)$?

14.3 Modélisation du jeu de pile ou face infini : suite d'épreuves de Bernoulli mutuellement indépendantes

Idée de démonstration

Problème : construire $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, suivant une loi de Bernoulli $B\left(\frac{1}{2}\right)$.

Soit $\Omega = [0,1]$, T : la tribu borélienne, et μ la mesure de Lebesgue ; on définit X_n par

$$X_n(x) = E(2^n x) \,\%\, 2$$

c'est-à-dire le $n-$ième chiffre de x dans l'écriture en base 2.

Plus clairement :

$$x = \sum_{n=1}^{\infty} \frac{X_n(x)}{2^n}$$

X_1 et X_2 sont indépendantes, pourquoi ?

15 Complément : espérance d'une variable aléatoire entière

On suppose que X est une variable aléatoire à valeurs dans $\mathbb{N}$ qui possède une espérance :

$$E(X) = \sum_{n=0}^{\infty} n.\mathbb{P}(X = n)$$

Dans ce cas, l'espérance est aussi donnée par la formule :

$$E(X) = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n)$$

Notons $a_n = \mathbb{P}(X = n)$ et $b_n = \mathbb{P}(X \geq n)$; pour tout $n \geq 0$:

$$a_n = b_n - b_{n+1}$$

et

$$E(X) = \sum_{n=0}^{\infty} n.a_n = \sum_{n=1}^{\infty} b_n$$

15.1 Démonstration dans le cas fini

On suppose que X est à valeurs dans $[0, N]$.

$$E(X) = \sum_{k=0}^{N} k.a_k = \sum_{k=1}^{N} k.(b_k - b_{k+1}) = \sum_{k=1}^{N} b_k - N.b_{N+1}$$

Conclusion ?

15.2 Démonstration élémentaire

On suppose ici que X est à valeurs dans $\mathbb{N}$, et que l'espérance existe.
Soit $N \geq 1$.

$$\sum_{k=0}^{N} k.a_k = \sum_{k=1}^{N} k.(b_k - b_{k+1}) = \sum_{k=1}^{N} b_k - N.b_{N+1}$$

Il reste à montrer que $(N.b_{N+1})$ tend vers 0. Or :

$$0 \leq N.b_{N+1} = N. \sum_{k=N+1}^{\infty} a_k \leq \sum_{k=N+1}^{\infty} k.a_k$$

qui tend vers 0 (reste d'une série convergente).

15.3 Avec le théorème de Fubini

On suppose ici que X est à valeurs dans $\mathbb{N}$, et que l'espérance existe.

$$E(X) = \sum_{p=1}^{\infty} a_p.p = \sum_{p=1}^{\infty} a_p. \sum_{q=1}^{\infty} c_{p,q}$$

où $c_{p,q} = ?$

Réponse

$c_{p,q} = 1$ si $q \leq p$ et 0 sinon. Donc :

$$E(X) = \sum_{q=1}^{\infty} \sum_{p=1}^{\infty} c_{p,q}.a_p = \sum_{q=1}^{\infty} \sum_{p=q}^{\infty} a_p = \sum_{q=1}^{\infty} b_q$$

Justification

L'application du théorème de Fubini est justifiée car tous les termes sont positifs et toutes les séries convergent.
$b_q = \sum_{p=q}^{\infty} a_p$: car suite d'évènements deux à deux disjoints.

15.4 Une illustration

L'exercice **6.4**

16 Complément : la fonction de répartition

16.1 Définition

Soit X une variable aléatoire réelle discrète ; on pose

$$F_X(t) = \mathbb{P}(X \leq t)$$

16.2 Monotonie

F_X est croissante sur $\mathbb{R}$.

Démonstration

Si $x < y$, alors $(X \leq x) \subset (X \leq y)$, donc $F_X(x) \leq F_X(y)$.

Remarque

F_X a donc en tout point une limite finie à gauche et à droite.

16.3 Limites

$\lim\limits_{+\infty} F = 1$ et $\lim\limits_{-\infty} F = 0$.

Démonstration

Soit $p_n = F_X(n) = \mathbb{P}(X \leq n)$; (p_n) converge vers 1 d'après...
???
le théorème de continuité croissante.

16.4 Continuité à droite

F_X est continue à droite sur $\mathbb{R}$.

Démonstration

Soit $a \in \mathbb{R}$, $a_n = a + \frac{1}{n}$, et

$$p_n = F_X(a_n) - F_X(a) = \mathbb{P}(a < X \leq a_n)$$

(p_n) converge vers 0 d'après...
???
le théorème de continuité décroissante.

16.5 Limite à gauche

Soit $a \in \mathbb{R}$.

$$F_X(a) - F_X\left(a^-\right) = \mathbb{P}(X = a)$$

Démonstration

Analogue au précédent.

Remarque

F_X est continue au point a si et seulement si $\mathbb{P}(X = a) = 0$.

17 Complément : le lemme de Borel-Cantelli

17.1 Introduction

Soit $(\Omega, T, \mathbb{P})$ un espace de probabilité.
Soit $(A_n)_{n \geq 0}$ une suite d'événements ; on pose

$$B_n = \bigcup_{k=n}^{\infty} A_k$$

On note

$$B = \bigcap_{n \geq 0} B_n$$

Que dire de la suite (B_n) ?
Que signifie $x \in B$? $x \in F = \Omega \setminus B$?
Que dire de $\mathbb{P}(B)$?

Réponse

(B_n) est décroissante ; on note $B = \limsup A_n$.

Notation :

$$\mathbb{N}_\omega = \{k \in \mathbb{N}/\omega \in A_k\}$$

B est l'ensemble des éléments de Ω qui appartiennent à A_k pour une infinité de valeurs de k.

F est donc l'ensemble des ω tels que $\mathbb{N}_\omega$ est fini.

On sait que

$$\mathbb{P}(B) = \lim_n \mathbb{P}(B_n)$$

Analogie avec les suites de réels

Pour une suite (u_n) de réels bornée, on note

$$v_n = \sup\{u_k/\, k \geq n\}$$

(v_n) est décroissante, convergente, et sa limite est notée limsup u_n ; c'est la plus grande valeur d'adhérence de (u_n).

17.2 Cas convergent

Soit $(\Omega, T, \mathbb{P})$ un espace de probabilité.

Soit $(A_n)_{n\geq 1}$ une suite d'événements telle que $\sum \mathbb{P}(A_n)$ converge.

Montrer que

$$\mathbb{P}(\limsup A_n) = 0$$

Démonstration

Soit

$$r_n = \sum_{k=n}^{\infty} \mathbb{P}(A_k)$$

Alors :

$$\forall n \geq 0,\, 0 \leq \mathbb{P}(B) \leq \mathbb{P}(B_n) \leq r_n$$

et (r_n) tend vers 0.

17.3 Cas divergent

Soit $(A_n)_{n\geq 1}$ une suite d'événements indépendants telle que $\sum \mathbb{P}(A_n)$ diverge.

On veut démontrer que

$$\mathbb{P}(\limsup A_n) = 1$$

Remarque : l'indépendance n'était pas supposée dans le cas convergent.

17.3.1 Lemme

Soit (u_n) le terme général d'une série divergente, à valeurs dans $[0, 1]$.

Soit

$$v_n = \prod_{k=0}^{n} (1 - u_k)$$

Montrer que $\lim_n v_n = 0$; on pourra utiliser :

$$\forall x \in \mathbb{R},\, 1 + x \leq e^x$$

Démonstration

$$\forall n \in \mathbb{N},\ 0 \leq v_n \leq \prod_{k=0}^{n} e^{-u_k} = e^{-s_n}$$

où (s_n) est la suite des sommes partielles de $\sum u_n$. Donc

$$s_n \to +\infty$$

Donc

$$v_n \to 0$$

17.3.2 1e étape

On veut montrer que

$$\mathbb{P}(B_0) = \mathbb{P}\left(\bigcup_{k=0}^{\infty} A_k\right) = 1$$

On note $C_k = \Omega \setminus A_k$, $D_n = \Omega \setminus B_n$ et $u_k = \mathbb{P}(A_k)$.

On doit donc montrer que $\mathbb{P}(D_0) = 0$.

$$D_0 = \bigcap_{k=0}^{\infty} C_k$$

$$\forall m \geq 0,\ 0 \leq \mathbb{P}(D_0) \leq \mathbb{P}(C_0 \cap ... \cap C_m) = \prod_{k=0}^{m} (1 - u_k)$$

Par passage à la limite, à l'aide du lemme,

$$\mathbb{P}(D_0) = 0,\ \mathbb{P}(B_0) = 1$$

17.3.3 2e étape

On constate que la même démonstration, et le même résultat s'appliquent à $(A_k)_{k \geq n}$; donc :

$$\forall n \geq 0,\ \mathbb{P}(B_n) = 1$$

On en déduit que $\mathbb{P}(B) = 1$ à l'aide du...

???

???

théorème de continuité décroissante.

17.4 Application

Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre $p \in]0, 1[$.

Montrer qu'avec probabilité 1, la suite $(X_n)_{n \geq 0}$ prend une infinité de fois la valeur 1 et une infinité de fois la valeur 0.

Ici

$$A_n = (X_n = 1)$$

Même chose si $(X_n)_{n \geq 1}$ est une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre $p_n = \frac{1}{n}$.

Quelle est la conclusion si $(X_n)_{n \geq 1}$ est une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre $p_n = \frac{1}{n^2}$?

18 Complément : retour à 0 lors d'une marche aléatoire

18.1 Introduction

$(X_i)_{i\geq 1}$ est une suite de variables aléatoires de Rademacher indépendantes :

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$$

On pose $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$. Calculer

$$\mathbb{P}(S_n = 0)$$

Réponse

On peut se ramener au cas d'une loi binomiale ainsi : $Y_i = \frac{1}{2}(X_i + 1)$

$$T_n = \sum_{i=1}^n Y_i = \frac{S_n + n}{2}$$

Pour $n = 2m$:

$$p_{2m} = \mathbb{P}(S_{2m} = 0) = \mathbb{P}(T_{2m} = m) = \frac{1}{2^{2m}}.\binom{2m}{m} \sim \frac{1}{\sqrt{\pi m}}$$

Que peut-on déduire du lemme de Borel-Cantelli ?

18.2 Le problème du scrutin

p personnes votent pour C_1 ; q personnes votent pour C_2. On suppose

$$p > q \geq 0$$

On cherche la probabilité pour que C_1 reste strictement en tête lors du dépouillement.

Notations

$$n = p + q,\ d = p - q$$

$$O = (0,0)\,,\ A = (1,1)\,,\ A' = (1,-1)\,,\ B = (n,d)\,,\ B' = (n,-d)$$

Nombre de chemins

- de O à B : $\binom{n}{q}$
- de A à B : $\binom{n-1}{q}$
- de A à B', ou de A' à B : $\binom{n-1}{q-1}$

Conclusion

Le nombre de cas favorables (nombre de chemins de O à B ne retouchant pas l'axe) :

$$f = \binom{n-1}{q} - \binom{n-1}{q-1}$$

Nombre de cas total :

$$t = \binom{n}{q}$$

D'où le résultat :

$$\frac{f}{t} = \frac{p-q}{p+q} = \frac{d}{n}$$

18.3 Le retour à 0 est presque sûr

Ici, seul $n \geq 1$ est fixé ; le nombre de marches aléatoires de longueur n est 2^n. Notons

$$m = \left\lfloor \frac{n-1}{2} \right\rfloor$$

Nombre de marches restant toujours strictement positives :

$$\sum_{q=0}^{m} \binom{n-1}{q} - \binom{n-1}{q-1} = \binom{n-1}{m}$$

La probabilité α que le marcheur ne repasse jamais par 0 est majorée par

$$\alpha_n = 2.2^{-n}.\binom{n-1}{m}$$

On vérifie que

$$\lim_n 2^{-n}.\binom{n-1}{m} = 0$$

par exemple avec la formule de Stirling.

Conclusion : le marcheur repassera presque sûrement par 0.

18.4 Le retour à 0 n'est pas unique

La probabilité que le marcheur repasse un nombre fini de fois par 0 est nulle.

Indication : étudier

$$\mathbb{P}\left(S_n = 0, S_{n+1} > 0, S_{n+2} > 0, ..., S_{n+m} > 0\right)$$

18.5 Utilisation de séries entières

18.5.1 Un développement en série entière

$$\forall x \in \left]-1, 1\right[, \ \frac{1}{\sqrt{1-x^2}} = \sum_{n=0}^{\infty} u_n x^n$$

avec u_n ?

Réponse

$$\forall x \in \left]-1, 1\right[, \ \frac{1}{\sqrt{1-x^2}} = \sum_{m=0}^{\infty} 2^{-2m}.\binom{2m}{m}.x^{2m} = \sum_{m=0}^{\infty} \mathbb{P}\left(S_{2m} = 0\right) x^{2m}$$

En déduire que $\sum \mathbb{P}\left(S_m = 0\right)$ diverge.

18.5.2 Le temps de retour

On note

$$T = \min\left\{n \geq 1 / \, S_n = 0\right\}$$

Montrer que T est une variable aléatoire.

18.5.3 Un produit de Coch-coch

Montrer que

$$\forall n \in \mathbb{N}^*, \mathbb{P}\left(S_n = 0\right) = \sum_{k=0}^{n} \mathbb{P}\left(T = k\right).\mathbb{P}\left(S_{n-k} = 0\right)$$

En déduire G_T et $\mathbb{P}\left(T = n\right)$.

Réponse

$$\forall t \in \left]-1,1\right[, \sum_{n=0}^{\infty} \mathbb{P}(T=n)\, x^n = 1 - \sqrt{1-x^2}$$

$$\forall n \in \mathbb{N}^*, \mathbb{P}(T=2m) = \frac{p_{2m}}{2m-1} = \frac{2^{-2m}}{2m-1}.\binom{2m}{m}$$

18.5.4 Qu'en déduire ?

En déduire que :
- T est fini presque sûrement.
- $E(T)$ est infini (par deux méthodes).

19 Complément : la formule de Wald

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{N}$ de même loi. Soit N une variable aléatoire à valeurs dans $\mathbb{N}$; on suppose $(N, X_1, X_2, ...)$ mutuellement indépendantes. On notera

$$S_0 = 0,\ S_n = \sum_{k=1}^{n} X_k$$

On s'intéresse à S définie par

$$S(\omega) = \sum_{k=1}^{N(\omega)} X_k(\omega) = S_{N(\omega)}(\omega)$$

En résumé,

$$S = S_N$$

Interprétation

Un individu donne naissance à un nombre aléatoire N de fils ; chacun à son tour a X fils ; S représente donc le nombre de petits-fils.

19.1 S est une variable aléatoire

$$\forall k \in \mathbb{N},\ S^{-1}(\{k\}) = \bigcup_{n=0}^{\infty} (N=n, S_n=k)$$

19.2 $G_S = G_N \circ G_X$

Fixons $t \in [0,1]$.

$$G_S(t) = \sum_{k=0}^{\infty} t^k \mathbb{P}(S=k)$$

Donc

$$G_S(t) = \sum_{k=0}^{\infty} t^k.\left[\sum_{n=0}^{\infty} \mathbb{P}(N=n)\,\mathbb{P}(S_n=k)\right] = \sum_{n=0}^{\infty} \mathbb{P}(N=n)\left[\sum_{k=0}^{\infty} \mathbb{P}(S_n=k)\, t^k\right]$$

En effet, le théorème de Fubini s'applique car tous les termes sont positifs et les séries convergent.

D'où

$$\forall t \in [0,1],\ G_S(t) = \sum_{n=0}^{\infty} \mathbb{P}(N=n).(G_X(t))^n = G_N \circ G_X(t)$$

19.3 $E(S) = E(N)\,E(X)$

Si N et X ont une espérance, S aussi, et $E(S) = E(N)\,E(X)$.

Il suffit de dériver et d'évaluer en 1.

19.4 Cas particuliers

Beaucoup d'exercices sont des cas particuliers de celui-ci.

Exemples : 6, 7, 15.