Probabilités discrètes

Cyrille GUIEU , Lycée Antoine Kela

Prérequis :

1. **Ensembles**
   (a) Si les A_j sont des ensembles indexés sur un ensemble J alors :
   $$(\bigcup_{j\in J} A_j)^c = \bigcap_{j\in J} A_j^c$$
   (b) Un ensemble I est fini et de cardinal n si il est en bijection avec $\{1,...,n\}$
   (c) Un ensemble J est dénombrable si il est bijection avec $\mathbb{N}$.
2. **Propriétés topologiques de $\mathbb{R}$**
   (a) $\sum u_n$ converge équivaut à
   $$\lim_{p\to+\infty} \sum_{n>p} u_n = 0$$
   (b) Une série majorée à termes positifs converge.
   (c) Si (u_n) est une suite positive et ϕ une bijection $\mathbb{N}$ dans $\mathbb{N}$ alors :
   $$\sum_{i=1}^{+\infty} u_i = \sum_{i=1}^{+\infty} u_{\phi(i)}$$
   (d) Si (u_n) est une suite telle que la série $\sum_{i=1}^{+\infty} u_i$ est absolument convergente et ϕ une bijection $\mathbb{N}$ dans $\mathbb{N}$ alors :
   $$\sum_{i=1}^{+\infty} u_i = \sum_{i=1}^{+\infty} u_{\phi(i)}$$
   On peut donc, dans le cas d'une série absolument convergente, donner un sens à $\sum_{n\in\mathbb{N}} u_n$

1 Espace probabilisé

Définition 1 : Ω est un ensemble et on note $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω. On dit alors que $\mathcal{A}$ est une tribu si les conditions suivantes sont réunies :

1. $\Omega \in \mathcal{A}$
2. Si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$
3. $\forall n \in \mathbb{N}, A_n \in \mathcal{A} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$

Proposition 1 : Si $\mathcal{A}$ est une tribu alors :

1. $\emptyset \in \mathcal{A}$
2. Si les $A_1, A_2, ..., A_n$ sont n éléments de la tribu $\mathcal{A}$ alors $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$
3. Une intersection finie ou dénombrable d'éléments de $\mathcal{A}$ est un élément de $\mathcal{A}$.

Exemple des deux dés indiscernables.
Définition 2 : Si $\mathcal{A}$ est une tribu sur Ω on dit que $(\Omega, \mathcal{A})$ est un espace probabilisable. Dans ce cas, la tribu $\mathcal{A}$ est appelée *tribu des évènements*.

Définition 3 : Si $(\Omega, \mathcal{A})$ est un espace probabilisable, une *probabilité* sur cet espace est une application p de $\mathcal{A}$ dans $\mathbb{R}^+$ qui vérifie:

1. $p(\Omega) = 1$
2. Si pour tout $n \in \mathbb{N}$, $A_n \mathcal{A}$ et que les A_n sont disjoints deux à deux alors : $p(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} p(A_n)$.

Remarque: cette définition est cohérente grâce au prérequis 2.b
Définition 4 : Dans ce cas, on dit alors que $(\Omega, \mathcal{A}, p)$est un *espace probabilisé.*
Proposition 2 : Si $(\Omega, \mathcal{A}, p)$ est un espace probabilisé alors :

1. $p(\emptyset) = 0$
2. Si $A \in \mathcal{A}$ alors $p(A^c) = 1 - p(A)$
3. Pour tout événement A, on a : $p(A) \in [0, 1]$
4. Si $A \subset B$ alors $p(A) \leq p(B)$
5. Si $A, B \in \mathcal{A}$ alors $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

Proposition 3 (Caractérisation de p sur $(\Omega, \mathcal{P}(\Omega))$ lorsque Ω est dénombrable)
Si Ω est dénombrable et que g est une fonction définie de Ω dans $\mathbb{R}^+$ telle que

$$\sum_{\omega\in\Omega} g(\omega) = 1$$

alors il existe une une unique probabilité p sur $(\Omega, \mathcal{P}(\Omega))$ telle que pour tout $\omega \in \Omega$,

$$p(\omega) = g(\omega)$$

On dit alors que g est le germe de p.
Proposition 4 : Si $(\Omega, \mathcal{A}, p)$ est un espace probabilisé alors :

1. Si A_n est une suite croissante d'évènements alors $p(\bigcup_{n\in\mathbb{N}} A_n) = \lim p(A_n)$
2. Si A_n est une suite décroissante d'évènements alors $p(\bigcap_{n\in\mathbb{N}} A_n) = \lim p(A_n)$

Exemple: Si on joue à pile ou face une infinité de fois, la probabilité de n'obtenir que des piles est 0.
Remarque : illustre bien la leçon 202 d'analyse

2 Variables aléatoires discrètes

Définition 1 : Soit $(\Omega, \mathcal{A})$ et $(E, \mathcal{E})$ des espaces probabilisables. Une application X de Ω dans E est dite *variable aléatoire* si pour tout $A \in \mathcal{E}$, $X^{-1}(A) \in \mathcal{A}$ avec

$$X^{-1}(A) = \{\omega \in \Omega / X(\omega) \in A\}$$

On note alors $X \in A$ l'évènement $X^{-1}(A)$
Proposition 1 : Soit X une application de Ω dans E telle que $X(\Omega)$ soit dénombrable. On suppose de plus que $\forall x \in E, x \in \mathcal{E}$. Dans ce cas, X est une variable aléatoire si et seulement si :

$$\forall x \in E, X^{-1}(x) \in \mathcal{A}$$

Définition 2 : Soit $(\Omega, \mathcal{A})$ et $(E, \mathcal{E})$ deux espaces probabilisables . On dit que X de Ω dans E est une *variable aléatoire discrète* si les deux conditions suivantes sont satisfaites:

- $X(\Omega)$ est dénombrable.

- Pour tout $x \in E$, $X^{-1}(x) \in \mathcal{A}$

Proposition 2 : Soit X une variable aléatoire de l'espace probabilisé $(\Omega, \mathcal{A}, p)$ dans l'espace probabilisable $(E, \mathcal{E})$.L'application $P_X : A \to p(X^{-1}(A))$ de E dans $[0,1]$ est une probabilité sur E. C'est la loi de la variable X.

Exemple : le modèle hypergéométrique

On considère un ensemble U partitionné en deux sous ensemble U_1 et U_2. On note r ,r_1 et r_2 les cardinaux de ces ensembles. L'expérience aléatoire consisite à tirer au hasard n éléments de U.On cherche la probabilité d'obtenir k_1 éléments de U_1 (pour $0 \leq k_1 \leq n$).
L'univers Ω est $\{A \in \mathcal{P}(U) || A| = n\}$. On a donc $|\Omega| = \binom{r}{n}$. On munit Ω de la tribu $\mathcal{P}(\Omega)$ et de la probabilité uniforme. On cherche $p(A_{k_1})$ où

$$A_{k_1} = \{A \in \Omega || A \cap U_1| = k_1\}$$

A_{k_1} est non vide si et seulement si : $n - (r - r_1) \leq k_1 \leq r_1$ dans tous les cas :

$$p(A_{k_1}) = \frac{\binom{r_1}{k_1}\binom{r-r_1}{n-k_1}}{\binom{r}{n}}$$

(Convention: $\binom{n}{p}$=0 si $p > n$ ou $n < 0$ ou $p < 0$).
Ainsi si on note X la variable aléatoire égale au nombre d'éléments de U_1 chosis on a:

$$p(X = k_1) = \frac{\binom{r_1}{k_1}\binom{r-r_1}{n-k_1}}{\binom{r}{n}}$$

3 Indépendance

Définition 1 :

1. Deux évènements A et B d'un espace probabilisé $(\Omega, \mathcal{A}, p)$ sont dits *indépendants* si $p(A \cap B) = p(A)p(B)$.
2. Des évènements $(A_i)_I$ indexés sur un ensemble quelqconque I sont dits indépendants si pour tout sous-ensemble fini $J \subset I$:
$$p(\bigcap_{j \in J} A_j) = \prod_{j \in J} p(A_j)$$

Rq: l'indépendance 2 à 2 n'implique pas l'indépendance.

Définition 2 :

1. Si X_1 est une variable aléatoire de $(\Omega, \mathcal{A}, p)$ dans $(E_1, \mathcal{E}_1)$ et X_2 est une variable aléatoire de $(\Omega, \mathcal{A}, p)$ dans $(E_2, \mathcal{E}_2)$, on dit que X_1 et X_2 sont indépendantes si pour tout $A_1 \in \mathcal{A}_1$ et $A_2 \in \mathcal{A}_2$, les évènements $X_1^{-1}(A_1)$ et $X_2^{-1}(A_2)$ sont indépendants.
2. Si pour $i \in I$, X_i est une variable aléatoire de $(\Omega, \mathcal{A}, p)$ dans $(E_i, \mathcal{E}_i)$, on dit que $(X_i)_{i \in I}$ sont indépendantes si pour tout ensemble d'évènements, $(A_i)_{i_i n I}$ tels que pour tout $i \in I$ on ait $A_i \in \mathcal{E}_i$, les évènements $(X_i^{-1}(A_i))_{i \in I}$ sont indépendants.

Proposition 1 : Soit $(X_i)_{i \in \{1,...,n\}}$ une famille finie de variables aléatoires discrètes. Les (X_i) sont indépendantes si et seulement si pour tout $(x_1, ..., x_n) \in E_1 \times .... \times E_2$, on a :

$$p(\bigcap_{i=1}^{n} X_i = x_i) = \prod_{i=1}^{n} p(X_i = x_i)$$

4 Exercices : Expérience répétées, modèle géométrique, modèle binomial, loi de Poisson

On considère un espace probabilisé $(\Omega, \mathcal{A}, p)$ et des évènements indépendants $(A_i)_{i \in \mathbb{N}}$ ayant la même probabilité p. Cette situation est un modèle pour des expériences ayant deux issues possibles répétées de manière indépendante.
On note alors N la fonction qui à ω associe le nombre $inf\{k \in \mathbb{N} | \omega \in A_k\}$ si $\{k \in \mathbb{N} | \omega \in A_k\}$ est non vide et $+\infty$ sinon.
De même on note $S_n = \sum_{i=0}^{n} 1_{A_i}$ la fonction égale à $|\{k \in \mathbb{N} | \omega \in A_k\}|$.

1. **Expériences répétées**

   (a) Montrer que ce sont deux variables aléatoires de $(\Omega, \mathcal{A}, p)$ dans $\mathbb{N} \cup \{\infty\}$

   (b) Montrer que la loi de N est donnée par :
   $$p(N = k) = q^k p$$
   avec $q = 1 - p$.
   On note $\mathcal{G}(p)$ cette loi.

   (c) Montrer que la loi de S_n est donnée par
   $$p(S_n = k) = (\begin{matrix} n \\ k \end{matrix}) p^k q^{n-k}$$
   On note $\mathcal{B}(n, p)$ cette loi.

   (d) **Exemple** Un joueur jette une pièce truquée ayant une probabilité p de tomber sur "‘pile"’ jusqu'à ce qu'il obtienne "‘pile"’. Si ceci se passe au k-ième lancé, il lance alors k fois un dé bien équilibré. Il gagne si il obtient exactement un six. Quel est la probabilité pour que le joueur gagne à ce jeu ? Comment choisir p pour que cette probabilité soit maximale ?

2. **Théorème de Poisson** On considère une suite $(p_n)_{\mathbb{N}}$ de réels dans l'intervalle $]0; 1[$ telle que $\lim np_n = \lambda$ (avec $\lambda > 0$). Soit S_n une variable aléatoire de loi $\mathcal{B}(n, p_n)$.

   (a) Démontrer que
   $$\lim_{n \to +\infty} (1 - \frac{\lambda}{n} + o(\frac{1}{n}))^{n-k} = e^{-\lambda}$$

   (b) En déduire que
   $$\lim_{n \to +\infty} p(S_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

   (c) Montrer que la fonction $\mathcal{P}(\lambda) : k \to e^{-\lambda} \frac{\lambda^k}{k!}$ définit une loi de probabilité sur $\mathbb{N}$. C'est la *loi de Poisson de paramètre* λ

5 Moments d'une variable aléatoire

Définition 1 Soit X une variable aléatoire discrète à valeurs réelles. Si la série $\sum_{x \in \mathbb{R}} xp(X = x)$ converge absolument, on dit que la variable aléatoire X possède une moyenne et le nombre $\sum_{x \in \mathbb{R}} xp(X = x)$ est appelé *moyenne* ou *espérance* de la variable aléatoire X. On note ce nombre $E(X)$.

On note $\mathcal{L}_d^1(\Omega, \mathcal{A}, p)$ l'ensemble de variables aléatoires discrètes à valeurs réelles qui possèdent une moyenne.

Théorème de transfert Soit X une variable aléatoire réelle discrète prenant ses valeurs dans l'ensemble E. Soi f une application de E dans $\mathbb{R}$. L'application composée $Y = f \circ X$ notée $f(X)$ est alors une variable aléatoire discrète. Pour que Y admette une moyenne, il faut que la série
$$\sum_{x \in E} |f(x)| p(X = x)$$
soit finie et dans ce cas :
$$E(X) = \sum_{x \in E} f(x) p(X = x)$$

Proposition 1 $\mathcal{L}_d^1(\Omega, \mathcal{A}, p)$ est un espace vectoriel et l'application E est une forme linéaire sur $\mathcal{L}_d^1(\Omega, \mathcal{A}, p)$.

Proposition 2

- Si X suit une loi de Bernoulli telle que $X = 1_A$ alors $E(X) = p(A)$
- Si S_n suit la loi binomiale $\mathcal{B}(n, p)$ alors $E(S_n) = np$

- Si N suit la loi géométrique $\mathcal{G}(p)$ alors $E(N) = \frac{q}{p}$
- Si X suit la loi de Poisson $\mathcal{P}(\lambda)$ alors $E(X) = \lambda$

Définition 2 On note $\mathcal{L}_d^p(\Omega, \mathcal{A}, p)$ l'ensemble des variables aléatoires discrètes à valeurs réelles X telles que $X^p \in \mathcal{L}_d^1(\Omega, \mathcal{A}, p)$

Proposition 3 (inégalité de Schwarz)

1. $\mathcal{L}_d^2(\Omega, \mathcal{A}, p) \subset \mathcal{L}_d^2(\Omega, \mathcal{A}, p)$ et
$$E(|X|) \leq (E(X^2))^{1/2}$$
2. Si X et Y sont dans $\mathcal{L}_d^2(\Omega, \mathcal{A}, p)$ alors $XY \in \mathcal{L}_d^2(\Omega, \mathcal{A}, p)$ et
$$E(|XY|) \leq (E(X^2))^{1/2}(E(Y^2))^{1/2}$$

Définition 3 Si $X \in \mathcal{L}_d^2(\Omega, \mathcal{A}, p)$ alors X admet une moyenne et on appelle variance le réel : $var(X) = E((X - E(X))^2)$. Sa racine carrée positive est notée σ_X. On a donc $\sigma_X^2 = var(X)$. D'après le théorème de transfert on a donc :
$$\sigma_X^2 = \sum_{x \in val(X)} (x - E(X))^2 p(X = x)$$

Proposition 4 Si X et Y sont indépendantes alors :

1. E(XY)=E(X)E(Y)
2. $\sigma_{(X+Y)^2} = \sigma_X^2 + \sigma_Y^2$

Proposition 5

1. $\sigma_{1_A}^2 = p(1-p)$
2. Si X suit la loi binomiale $\mathcal{B}(n, p)$ alors $\sigma_X^2 = np(1-p)$
3. Si X suit la loi géométrique $\mathcal{G}(p)$ alors $\sigma_X^2 = \frac{q}{p^2}$
4. Si X suit la loi de Poisson $\mathcal{P}(\lambda)$ alors $\sigma_X^2 = \lambda$

6 Approximation de lois

Proposition 1 : (Inégalité de Markov) Si $X \in \mathcal{L}_d^1(\Omega, \mathcal{A}, p)$ on a, pour tout $\epsilon > 0$:
$$p(|X| > \epsilon) \leq \frac{E(|X|)}{\epsilon}$$

Proposition 2: (Inégalité de Tchebichev) Si $X \in \mathcal{L}_d^2(\Omega, \mathcal{A}, p)$ on a, pour tout $\epsilon > 0$:
$$p(|X - E(X)| > \epsilon) \leq \frac{\sigma_X^2}{\epsilon^2}$$

Définition : Une suite de variables aléatoires (X_n) est dite *converger en probabilité* vers une varaiable alétoire X si pour tout $\epsilon > 0$,
$$\lim_n p(|X_n - X| > \epsilon) = 0$$

Théorème (loi faible des grands nombres) Soit (X_n) une suite de variables aléatoires discrètes, définies sur une espace probabilisé $(\Omega, \mathcal{A}, p)$, indépendantes et admettant une variance. On suppose la convergence des suites :
$$\lim_{n \to +\infty} \frac{1}{n} \sum_{j=1}^{n} E(X_j) = m$$
et
$$\lim_{n \to +\infty} \frac{1}{n^2} \sum_{j=1}^{n} \sigma_{X_j}^2 = 0$$

Alors la suite de variables aléatoires $\overline{X_n} = \frac{1}{n} \sum_{j=1}^{n} X_j$ converge en probabilité vers m

Problème : formule de Poincaré et applications

On considère un espace probabilisé $(\Omega, \mathcal{A}, p)$.

1. **La formule de Poincaré**

   (a) Si A_1, A_2, A_3 sont des évènements, montrer que

   $p(A_1 \cup A_2 \cup A_3) = p(A_1)+p(A_2)+p(A_3)-p(A_1\cap A_2)-p(A_1\cap A_3)-p(A_2\cap A_3)+p(A_1\cap A_2\cap A_3)$

   (b) Démontrer la formule de Poincaré : Si $n \geq 2$, pour toute suite finie $(A_i)_{1\leq i\leq n}$ d'éléments de $\mathcal{A}$ alors :

   $$p(\bigcup_{i=1}^{n} A_i) = \sum_{k=1}^{n}((-1)^{k-1} \sum_{\substack{J\subset I\\ |J|=k}} p(\bigcap_{j\in J} A_j))$$

   en notant $|J|$ le cardinal de l'ensemble J et $I = \{1, ..., n\}$.

2. **Manipulation d'un ensemble fini d'évènements** Pour toute partie J non vide de I on note : $\hat{A}_J = \bigcap_{j\in J} A_j$ et $B^J = \hat{A}_J \cap (\bigcap_{j\in J^c} A_j^c)$ puis pour si $1 \leq m \leq n$:

   $$B_m = \bigcup_{\substack{J\subset I\\ |J|=m}} B^J$$

   et enfin :

   $$B_0 = \bigcap_{j=1}^{n} A_j^c$$

   (a) Montrer que si J et J' sont distinctes et non vides, les ensembles B^J et $B^{J'}$ sont disjoints.

   (b) Montrer que pour tout $m \in I$, $\omega \in B_m \Leftrightarrow \omega$ appartient à exactement m évènements A_i.

   On note $S_0 = 1$ et, pour $1 \leq r \leq n$, $S_r = \sum_{\substack{J\subset I\\ |J|=r}} p(\hat{A}_J)$

   (c) Pour tout $A \in \mathcal{A}$,

   $$p(B_0 \cup A) = p(A) + \sum_{k=1}^{n}(-1)^k \sum_{\substack{J\subset I\\ |J|=k}} p(\bigcap_{j\in J} A_j \cap A)$$

   (d) Montrer que si $1 \leq m \leq n$,

   $$p(B_m) = S_m - \binom{m+1}{m} S_{m+1} + ... + (-1)^{n-m}\binom{n}{m} S_n$$

3. **Une application concrète** Au cours d'une soirée, n personnes déposent dans un chapeau une carte portant leur nom (aucune personne ne portant le même nom qu'une autre). A la fin de la soirée, on brasse bien les cartes puis chacun tire une carte.

   (a) Quelle est la probabilité qu'aucune personne ne tire une carte portant son nom ?

   (b) Quelle est la probabilité qu'exactement m personnes tirent une carte portant leur nom ?

4. **Une application à un problème d'indépendance** $(\Omega, \mathcal{A}, p)$ est un espace probabilisé

   (a) Si A et B sont deux évènmements indépendants, démontrer que :

   - A et B^c sont indépendants.
   - A^c et B^c sont indépendants.

(b) Si $(A_i)_I$ est une famille d'évènements indépendants et que I est partitionné en deux parties I_1 et I_2 alors on définit la famille $(B_i)_I$ par :

$$B_i = \{ \begin{array}{l} A_i \ si \ i \in I_1 \\ A_i^c \ si \ i \in I_2 \end{array}$$

Démontrer que les évènements $(B_i)_I$ sont indépendants.

(c) En déduire que si une famille d'évènement est indépendante alors la famille des évènements contraires est une famille d'évènements indépendants.

(d) **Calcul de l'indicateur d'Euler par les probabilités** Soit Ω l'ensemble des entiers $\{1, ..., n\}$ où n est un entier non premier supérieur ou égal à 2. On considère l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$ où p est la probabilité uniforme. Si d divise n on note : $A_d = \{kd | k \in \Omega et kd \in \Omega\}$

i. Quelle est la probabilité de A_d ?
Soit $p_1 < p_2 < ... < p_r$ la suite des diviseurs premiers de n rangés par ordre croissant.

ii. Démontrer que $(A_{p_i})_{1 \leq i \leq r}$ est une famille d'évènements indépendants.

iii. En déduire le cardinal $\varphi(n)$ de l'ensemble A des nombres inférieurs ou égaux à n et premiers avec n (indicateur d'Euler).

7 Bibliographie

- *Cours d'analayse*, Alain Pommellet, 1997, Ellipses
- *Probabilités 1*, Jean-Yves Ouvrard, 1998, Cassini
- *Cours de probabilités et de statistiques*, Christian Leboeuf Jean-Louis Roque et Jean Guegand, 1996, Ellipses
- *Rapport de jury de l'agrégation interne de mathématiques 2006*