

Convergence de variables aléatoires.

M1EEF - Maths - Probabilités et Statistiques

On se propose dans ce problème de présenter et d'étudier différents modes de convergence d'une suite de variables aléatoires. La partie I est indépendante du reste du problème.

On notera $(\Omega, \mathcal{A}, p)$ un espace probabilisé à partir duquel seront définies les variables aléatoires considérées si ce n'est pas précisé. A toute variable aléatoire $X : \Omega \rightarrow \mathbb{R}$, on associe F_X la fonction de répartition de X , c'est-à-dire la fonction définie de $\mathbb{R}$ dans $[0, 1]$ telle que pour tout $x \in \mathbb{R}$ on ait $F_X(x) = p(X \leq x)$.

Pour un réel $\lambda > 0$, on dira qu'une variable aléatoire X suit une loi de Poisson de paramètre λ si X est une variable aléatoire discrète, d'image $\mathbb{N}$, et si $\forall k \in \mathbb{N}$, $p(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$.

Partie I - Une analyse de production d'élève

L'exercice suivant est donné à une classe. Une puce se déplace sur un axe gradué. Au temps $n = 0$, la puce est en 0. Si à l'instant n la puce est en x , alors à l'instant $n + 1$ elle sera en $x + 1$ avec probabilité $1/2$, en $x + 2$ avec probabilité $1/6$, et en $x - 3$ avec la probabilité $1/3$. On demande aux élèves un algorithme simulant la position de la puce après 50 itérations.

La proposition d'un élève :

$x \leftarrow 0$

Pour t allant de 0 à 50 Faire

 Si $(\text{random}()) \leq 1/2$ alors $x = x + 1$

 Si $(\text{random}()) \leq 1/6$ alors $x = x + 2$

 Si $(\text{random}()) \leq 1/3$ alors $x = x - 3$

Fin Pour

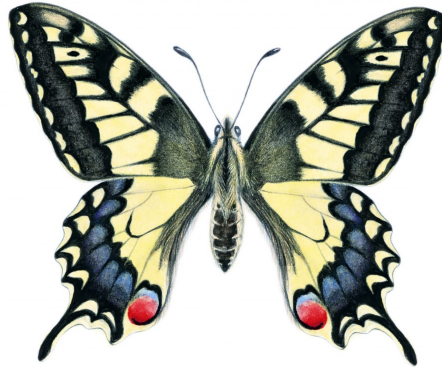
Afficher x .

1. Analyser l'algorithme proposé par l'élève, en vous concentrant sur ce qui vous semble important et en hiérarchisant vos remarques. On tiendra compte du fait que pour les élèves de cette classe la fonction $\text{random}()$ est une fonction qui renvoie de façon aléatoire un réel entre 0 et 1.

2. Proposer une correction.

Partie II - Des lois très classiques

Dans ce qui suit, r , r_1 et n sont des entiers tels que $0 \leq r_1 \leq r$ et $n \leq r$. On suppose que dans une vallée vivent r papillons dont r_1 *machaons*, une espèce particulière. On capture n papillons (on suppose que les différentes espèces se laissent indifféremment capturer) sans les relâcher. On utilise un modèle probabiliste, et on considère une variable aléatoire X dont la valeur correspond au nombre de papillons de l'espèce *machaon* attrapés.



- 1.a. Donner la loi de X . (Cette loi est appelée loi hypergéométrique de paramètres n , r et r_1 .)
 1.b. Pour $(k, n_1) \in \mathbb{N}^2$ tel que $k \leq n_1$, exprimer simplement $k \binom{n_1}{k}$ en fonction de $\binom{n_1-1}{k-1}$.
 1.c. Démontrer que pour tout $(a, b, c) \in \mathbb{N}^3$ tel que $c \leq a + b$ on a l'égalité

$$\sum_{j=0}^c \binom{a}{j} \binom{b}{c-j} = \binom{a+b}{c}.$$

Indication : on peut utiliser le fait que la somme des probabilités d'une loi hypergéométrique bien choisie doit valoir 1. Une autre façon de faire consiste à considérer le coefficient X^c dans le polynôme $(1+X)^{a+b} = (1+X)^a(1+X)^b$.

- 1.d. En déduire l'espérance de X : $E(X) = \frac{r_1 n}{r}$.

2.a. On considère une expérience différente : on capture les n papillons un par un et on note l'espèce à chaque fois avant de relâcher le papillon. On note Y le nombre de fois qu'un papillon de l'espèce *machaon* a été attrapé. Donner, en justifiant, la loi de Y . Quel est le nom de cette loi ?

2.b. Calculer l'espérance de Y . (Une méthode efficace consiste à écrire Y comme somme de variables aléatoires très simples.)

3. Sans calcul et sans démonstration, et en deux lignes, quelle est votre intuition quant au lien entre X et Y si r et r_1 sont très grands par rapport à n ?

Partie III - Convergence de variables aléatoires

- 1.a. Quelques rappels sur la fonction de répartition. Donner un exemple de variable aléatoire X telle que sa fonction de répartition ne soit pas continue.
 1.b. La fonction de répartition d'une variable aléatoire à densité est-elle toujours dérivable ?

On considère une première définition de convergence pour les variables aléatoires :

Définition de la convergence en loi : On considère un espace probabilisé $(\Omega, \mathcal{A}, p)$ et une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$, ainsi qu'une variable aléatoire X . On dit que (X_n) converge en loi vers la variable aléatoire X si pour tout $x \in \mathbb{R}$ en lequel F_X est continue on a $\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$.

2. Pour tout réel $\alpha \in \mathbb{R}$ on définit la variable aléatoire Δ_α telle que $\Delta_\alpha(\Omega) = \{\alpha\}$. On considère alors $a \in \mathbb{R}$. On souhaite montrer que la suite de variables aléatoires $(\Delta_{a+1/n})_{n \in \mathbb{N}^*}$ converge en loi

vers Δ_a .

2.a. On fixe $\alpha \in \mathbb{R}$. Déterminer F_{Δ_α} , puis tracer sa courbe représentative.

2.b. Etudier la continuité de F_{Δ_α} .

2.c. Conclure.

On propose un deuxième mode de convergence pour les suites de variables aléatoires :

Définition de la convergence en probabilité : On considère un espace probabilisé $(\Omega, \mathcal{A}, p)$, une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$, ainsi qu'une variable aléatoire X . On dit que (X_n) converge en probabilité vers la variable aléatoire X si pour tout $\epsilon > 0$ on a

$$\lim_{n \rightarrow \infty} p(|X_n - X| \geq \epsilon) = 0.$$

3. Pour $a \in \mathbb{R}$, étudier la convergence en probabilité de la suite de variables aléatoires $(\Delta_{a+1/n})_{n \in \mathbb{N}^*}$.

4. Dans cette question, on suppose que $\Omega = \{\omega_1, \omega_2, \dots\}$, et que $\forall i \in \{1, 2\} \ p(\omega_i) = \frac{1}{2}$. On considère une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ telle que pour tout $n \in \mathbb{N}$ on ait les égalités $X_{2n}(\omega_1) = 5$, $X_{2n}(\omega_2) = 7$, $X_{2n+1}(\omega_1) = 7$ et $X_{2n+1}(\omega_2) = 5$.

4.a. Montrer que la suite (X_n) converge en loi vers X_1 . (On pourra décrire F_{X_n} pour tout $n \in \mathbb{N}$.)

4.b. La suite (X_n) converge-t-elle vers X_1 en probabilités ?

Partie IV - Loi faible des grands nombres.

La loi faible des grands nombres est un théorème aux conséquences importantes qui peut s'énoncer ainsi :

Théorème 2 (Loi faible des grands nombres) : On considère un espace probabilisé $(\Omega, \mathcal{A}, p)$ et une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ indépendantes, admettant la même espérance et la même variance. Alors la suite de variables aléatoires $\frac{X_1 + \dots + X_n}{n}$ converge vers $E(X_1)$ en probabilité.

L'objectif de cette partie est de démontrer la loi faible des grands nombres dans le cas d'une variable aléatoire discrète.

1. On souhaite redémontrer l'inégalité de Markov dans le cas discret, c'est-à-dire dans le cas d'une variable aléatoire Z , admettant une espérance, discrète et positive (i.e telle que $Z(\Omega) \subset \mathbb{R}_+$ et telle que $Z(\Omega) = \{z_k \mid k \in \mathbb{N}\}$).

1.a. Avec les notations introduites, montrer que

$$E(Z) \geq \sum_{k \in \mathbb{N} \mid z_k \geq a} z_k p(Z = z_k).$$

1.b. En déduire

$$\forall a > 0, \ p(Z \geq a) \leq \frac{E(Z)}{a}.$$

1.c. *Une première application.* On lance plusieurs fois une pièce équilibrée : utiliser l'inégalité de Markov pour majorer la probabilité de devoir attendre au moins 40 tirages pour obtenir face pour la première fois.

1.d. Calculer cette probabilité. Que pensez-vous de la majoration obtenue en **1.c.** avec l'inégalité de Markov ?

2. En appliquant l'inégalité de Markov à une variable aléatoire bien choisie, montrer l'inégalité de Bienaymé-Tchebychev, c'est-à-dire que dans le cas où une variable aléatoire X admet une espérance et une variance, on a

$$\forall a > 0, p(|X - E(X)| \geq a) \leq \frac{Var(X)}{a^2}.$$

3. En déduire une démonstration de la loi faible des grands nombres.

Partie V - Convergence en loi de variables aléatoires discrètes

1. Dans le cas de variables aléatoires prenant leurs valeurs dans $\mathbb{N}$, on dispose d'une interprétation simple de la convergence en loi :

Théorème 1 : On considère un espace probabilisé $(\Omega, \mathcal{A}, p)$ et une suite de variables aléatoires discrètes $(X_n)_{n \in \mathbb{N}}$ à valeurs dans $\mathbb{N}$, ainsi qu'une variable aléatoire discrète X à valeurs dans $\mathbb{N}$. La suite de variables aléatoires (X_n) converge en loi vers la variable aléatoire X si et seulement si pour tout $k \in \mathbb{N}$ on a

$$\lim_{n \rightarrow \infty} p(X_n = k) = p(X = k).$$

Démontrer ce théorème. Nous l'utiliserons dans cette partie.

2. *Convergence de la loi binomiale vers la loi de Poisson* On considère un réel λ strictement positif, une suite $(p_n)_{n \in \mathbb{N}}$ d'éléments de $]0, 1[$ telle que $\lim_{n \rightarrow \infty} np_n = \lambda$, et la suite de variables aléatoires $(S_n)_{n \in \mathbb{N}^*}$ telle que pour tout $n \in \mathbb{N}^*$ on ait S_n de loi $\mathcal{B}(n, p_n)$. On souhaite montrer que (S_n) converge en loi vers une variable aléatoire suivant une loi de Poisson de paramètre λ .

2.a. Montrer que pour tout $k \in \mathbb{N}$ et $n \geq k$ on a

$$p(S_n = k) = \frac{n(n-1) \cdots (n-k+1)}{k!} \left(\frac{\lambda}{n} + o\left(\frac{1}{n}\right) \right)^k \left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right) \right)^{n-k}.$$

2.b. Conclure.

3. *Convergence de la loi hypergéométrique vers la loi binomiale* On considère $(X_j)_{j \in \mathbb{N}}$ une suite de variables aléatoires de loi hypergéométrique de paramètres $n, r^{(j)}$ et $r_1^{(j)}$ telle que $r^{(j)}$ diverge vers $+\infty$, et telle qu'il existe $p \in]0, 1[$ tel que $\lim_{j \rightarrow \infty} \frac{r_1^{(j)}}{r^{(j)}} = p$. On souhaite montrer que sous ces conditions la suite (X_j) converge en loi vers une variable aléatoire Y de loi binomiale $\mathcal{B}(n, p)$.

3.a. Montrer que si des entiers $k, n, r_1^{(j)}$ et $r^{(j)}$ sont tels que $k \leq r_1^{(j)}$ et $n - k \leq r^{(j)} - r_1^{(j)}$ alors on a

$$p(X_j = k) = \binom{n}{k} \prod_{l=0}^{k-1} \frac{r_1^{(j)} - l}{r^{(j)} - l} \prod_{l=0}^{n-k-1} \frac{r^{(j)} - r_1^{(j)} - l}{r^{(j)} - l}.$$

3.b. Conclure.