Statistique et Informatique (3I005)

2017-2018

Nicolas Baskiotis

Sorbonne Université - Université Pierre et Marie Curie (UPMC)
Laboratoire d'Informatique de Paris 6 (LIP6)
http://3I005.lip6.fr

Cours 4 :

Espérance, Variance
Loi des grands nombres

Résumé

On a vu comment :

- compter et dénombrer *(cours 1)*
- définir, manipuler et calculer une probabilité, *(cours 2)*
- définir une variable aléatoire *(cours 3)*.

A venir :

- manipuler une variable aléatoire
- modélisation, faire le lien entre la théorie et l'empirique
- probabilités réelles.

Loi d'une variable aléatoire

Propriété

Une variable aléatoire $X : \Omega \to F$ est totalement définie par sa loi de probabilité, caractérisé par :

- son image dans F : l'ensemble des valeurs qu'elle peut prendre,
- les probabilités attribuées à chacune de ses valeurs $P(X = x), x \in F$.

Loi conjointe

Soit (Ω, P) un espace de probabilité, et soient X et Y deux $v.a.$ sur cet espace, à valeur resp. dans F et G. (X, Y) est une $v.a.$, appelée loi conjointe de X et Y; les valeurs de (X, Y) sont dans $F \times G$.

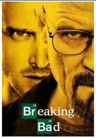
Propriétés

- la connaissance uniquement de X et de Y ne suffit pas à connaitre la loi jointe, sauf si X est indépendant de Y.
- $\forall x \in F, P(X = x) = \sum_{y \in G} P(X = x, Y = y)$

$\Rightarrow$ la connaisance de la loi jointe permet de déduire la loi de X, appelée dans ce cas *loi marginale*.

Tableau de contingence (intro)

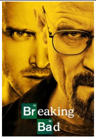
Etude de la loi jointe de deux variables aléatoires

Score, Film				
★	5	10	10	5
★★	10	2	5	5
★★★	15	20	15	10
★★★★	2	30	5	10

- Quelle est la loi jointe $P(S,F)$?
- La loi marginale du score ? du film ?
- Comment étudier l'indépendance ?

Tableau de contingence (intro)

Etude de la loi jointe de deux variables aléatoires

Score, Film					Total
★	5	10	10	5	30
★★	10	2	5	5	22
★★★	15	20	15	10	60
★★★★	2	30	5	10	47
Total	32	62	35	30	159

- Quelle est la loi jointe $P(S, F)$?
- La loi marginale du score ? du film ?
- Comment étudier l'indépendance ?

Plan

Caractéristiques d'une v.a.

Définitions

Soit une v.a. X,

- le *quantile* d'ordre α est la valeur x_α telle que $P(X < x_\alpha) = \alpha$.
- la médiane est le quantile d'ordre $\alpha = \frac{1}{2}$
- le *mode* (ou valeur dominante, valeur la plus probable) est la valeur de X associé à la plus grande probabilité.

Espérance d'une v.a. à valeurs réelles

Définition

Soit Ω un ensemble au plus dénombrable, et P une distribution de probabilité sur Ω.
Soit X une $v.a.$ sur l'espace probabilisé (Ω, P), à valeurs dans $\{x_1, x_2, ...\} \subset \mathbb{R}$.
on appelle *espérance* de X, notée $\mathbb{E}(X)$ la quantité :

$$\mathbb{E}(X) = \sum_i x_i P(X = x_i) = \sum_{\omega \in \Omega} X(\omega) P\big(\{\omega\}\big)$$

Remarques

- intuition (cf. loi des grands nombres, fin du cours) :

  l'espérance est la limite (quand $n \to \infty$) de la moyenne sur un échantillon de taille n.
- Si X suit une loi de Bernoulli de paramètre p, alors $\mathbb{E}(X) = p$.

Espérance d'une v.a. à valeurs réelles

Propriétés

Pour tout réel $a \in \mathbb{R}$ et tout couple de v.a. (X, Y) valeurs réelles :

- $\mathbb{E}(a) = a$,
- $\mathbb{E}(aX) = a\mathbb{E}(X)$,
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$,
  Plus généralement, si $X_1, ..., X_n$ sont n v.a. réelles:
  $$\mathbb{E}\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} \mathbb{E}(X_i)$$
- si X et Y sont indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

Espérance d'une v.a. à valeurs réelles

Propriétés

Pour tout réel $a \in \mathbb{R}$ et tout couple de v.a. (X, Y) valeurs réelles :

- $\mathbb{E}(a) = a$,
  Soit X la fonction constante $X(\omega) = a$,
  $\mathbb{E}(X) = \sum_\omega X(\omega)P(\{\omega\}) = \sum_\omega aP(\{\omega\}) = a\sum_\omega P(\{\omega\}) = a$
- $\mathbb{E}(aX) = a\mathbb{E}(X)$,
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$,
  Plus généralement, si $X_1, ..., X_n$ sont n v.a. réelles:

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i)$$

- si X et Y sont indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

Espérance d'une v.a. à valeurs réelles

Propriétés

Pour tout réel $a \in \mathbb{R}$ et tout couple de v.a. (X, Y) valeurs réelles :

- $\mathbb{E}(a) = a$,
- $\mathbb{E}(aX) = a\mathbb{E}(X)$,
  Soit $Y(\omega) = aX(\omega)$, $\mathbb{E}(Y) = \sum_\omega Y(\omega)P(\{\omega\}) = \sum_\omega aX(\omega)P(\{\omega\}) = a\mathbb{E}(X)$
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$,
  Plus généralement, si $X_1, ..., X_n$ sont n v.a. réelles:

$$\mathbb{E}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathbb{E}(X_i)$$

- si X et Y sont indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

Espérance d'une v.a. à valeurs réelles

Propriétés

Pour tout réel $a \in \mathbb{R}$ et tout couple de v.a. (X, Y) valeurs réelles :

- $\mathbb{E}(a) = a$,
- $\mathbb{E}(aX) = a\mathbb{E}(X)$,
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$,
  Soit $Z = X + Y$, $\mathbb{E}(Z) = \sum_{i,j}(x_i + y_j)P(X = x_i, Y = y_j)$
  $= \sum_{i,j} x_i P(X = x_i, Y = y_j) + \sum_{i,j} y_i P(X = x_i, Y = y_j)$
  $= \sum_i x_i(\sum_j P(X = x_i, Y = y_j)) + \sum_j y_j(\sum_i P(X = x_i, Y = y_j)) = \mathbb{E}(X) + \mathbb{E}(Y)$
  Plus généralement, si $X_1, ..., X_n$ sont n v.a. réelles:

$$\mathbb{E}\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} \mathbb{E}(X_i)$$

- si X et Y sont indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

Espérance d'une v.a. à valeurs réelles

Propriétés

Pour tout réel $a \in \mathbb{R}$ et tout couple de v.a. (X, Y) valeurs réelles :

- $\mathbb{E}(a) = a$,
- $\mathbb{E}(aX) = a\mathbb{E}(X)$,
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$,
  Plus généralement, si $X_1, ..., X_n$ sont n v.a. réelles:

$$\mathbb{E}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathbb{E}(X_i)$$

- si X et Y sont indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

Soit $Z = XY$, $\mathbb{E}(Z) = \sum_{i,j} x_i y_j P(X = x_i, Y = y_j) = \sum_{i,j} x_i y_j P(X = x_i)P(Y = y_j) = (\sum_i x_i P(X = x_i))(\sum_j y_j P(Y = y_j)) = \mathbb{E}(X)\mathbb{E}(Y)$

Espérance : exemple

- Espérance du score si les films et les profils sont équiprobables ?
- Si $P(\text{joyeux}) = 0.5,\ P(\text{grincheux} = 0.2)$?

Espérance : exemple

- Espérance du score si les films et les profils sont équiprobables ?
- Si $P(\text{joyeux}) = 0.5,\ P(\text{grincheux} = 0.2)$?

$$E(\text{Score}) = \sum_{f \in \text{Film}, p \in \text{Profil}} \text{Score}_{f,p} P(\text{F} = f, \text{P} = p)$$

Espérance d'une v.a. à valeurs réelles

Exemples

- Combien de cartes restent à une place inchangée après un mélange aléatoire du paquet ?
- Algorithmes de tri :
  on suppose que les éléments dans le tableau à trier sont placés aléatoirement :
  - l'espérance du nombre de comparaisons du *tri rapide* est en $O(n \ln n)$,
  - l'espérance du nombre de comparaisons du *tri par insertion* est en $O(n^2)$.
- Roulette : 37 numéros (0 à 36) :
  - pari sur les nombres impairs :
    on donne au joueur 2 fois sa mise s'il gagne ;
  - pari sur un nombre particulier :
    on donne au joueur 36 fois sa mise s'il gagne ;
  - vaut-il mieux jouer sur les nombres impairs ou sur un nombre particulier ?

Variance et écart-type d'une $v.a.$ à valeurs réelles

Définition

Soit X une $v.a.$ d'espérance m à valeurs dans $\{x_1, x_2, ...\}$, On appelle variance de X, la quantité:

$$V(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] = \sum_{i=1}^{\infty} (x_i - m)^2 P(X = x_i)$$

$\sigma(X) = \sqrt{V(X)}$ est appelé *l'écart-type* de X.

Propriétés

- $V(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$,
- $\forall a \in \mathbb{R}, V(a) = 0$,
- $\forall a \in \mathbb{R}, V(X + a) = V(X)$,
- $\forall a \in \mathbb{R}, V(aX + b) = a^2 V(X)$
- Si $X_1, ..., X_n$ sont n v.a.r. indépendantes : $V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$

Variance d'une v.a. à valeurs réelles

Exemples

- Roulette : 37 numéros (0 à 36) :
  - pari sur les nombres impairs :
    on donne au joueur 2 fois sa mise s'il gagne ;
  - pari sur un nombre particulier :
    on donne au joueur 36 fois sa mise s'il gagne.

  - L'espérance des gains est la même ;
  - quelle est la variance du gain dans chacun des cas ?

- Complexité des algorithmes :
  à complexité moyenne équivalente, on préfèrera souvent l'algorithme de plus faible variance.

Covariance de deux $v.a.$

Définitions

On appelle covariance de deux variables $v.a.$ X et Y, l'expréssion:

$$Cov(X,Y) = \mathbb{E}\left[(X - \mathbb{E}(X)) \times (Y - \mathbb{E}(Y))\right] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Propriétés

- $V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$,
- Si deux variables X et Y sont centrées on a alors $Cov(X,Y) = \mathbb{E}(XY)$
- Si deux variables sont indépendants alors:
  - $Cov(X,Y) = 0$, la réciproque est fausse
  - $V(X+Y) = V(X) + V(Y)$

Covariance de deux $v.a.$

Exemple

Soient X et Y deux variables aléatoires discrètes prenant leurs valeurs dans respectivement $\mathcal{X} = \{1,2,3\}$ et $\mathcal{Y} = \{-1,1\}$. La distribtuion de probabilité du couple est donnée par:

$y \backslash x$	1	2	3
-1	0.1	0.3	0.1
$+1$	0.2	0.1	0.2

On a dans ce cas $Cov(X,Y) = 0$ mais comme par exemple

$$\underbrace{P(X=1, Y=1)}_{=0.2} \neq \underbrace{P(X=1) \times P(Y=1)}_{=0.15}$$

les variables X et Y ne sont pas indépendantes!

Coefficient de corrélation

Propriété

Soient deux variables $v.a.$ X et Y définies sur un même espace probabilisé (Ω, C, P), on a

$$|Cov(X,Y)| \leq \sqrt{V(X)V(Y)}$$

De plus **l'égalité** a lieu si et seulement si Y et X sont linéairement dépendantes, ou si X est une constante

Définintion

Soient deux variables $v.a.$ X et Y définies sur un même espace probabilisé (Ω, C, P), on appelle le coefficient de corrélation la version normalisée de la covariance entre X et Y

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X) \times V(Y)}}$$

On a dans ce cas

$$-1 \leq \rho(X,Y) \leq +1$$

Plan

1. Caractéristiques d'une variable aléatoire

2. Lois usuelles

3. Loi des grands nombres

Loi de Bernouilli

Question :

Quelle est la probabilité de :

- obtenir pile avec une pièce équilibrée ?
- de deviner correctement une date d'anniversaire ?,
- d'obtenir un 6 sur un dé non pipé ?
- de gagner au loto ?

Loi de Bernouilli

Question :

Quelle est la probabilité de :

- obtenir pile avec une pièce équilibrée ?
- de deviner correctement une date d'anniversaire ?,
- d'obtenir un 6 sur un dé non pipé ?
- de gagner au loto ?

Réponse : loi de Bernouilli

C'est la loi d'une $v.a.$ X à valeur dans $\{0, 1\}$.
$X = 1$ représente le "succès" de l'expérience, et $X = 0$ l'"échec".

$$\forall x \in \{0, 1\}, P(X = x) = p^x(1-p)^{1-x}$$

La probabilité de succès $p = P(X = 1)$ est le paramètre de la loi.
$E(X) = p$, $V(X) = p(1-p)$

Loi binomiale

Question :

Quelle est la probabilité de :

- d'obtenir 10 piles lors de 30 tirages ?
- d'obtenir quatre nombre paires sur 10 tirages de dé ?,
- de faire plus de 2 erreurs

Loi binomiale

Question :

Quelle est la probabilité de :

- d'obtenir 10 piles lors de 30 tirages ?
- d'obtenir quatre nombre paires sur 10 tirages de dé ?,
- de faire plus de 2 erreurs

Réponse : loi binomiale

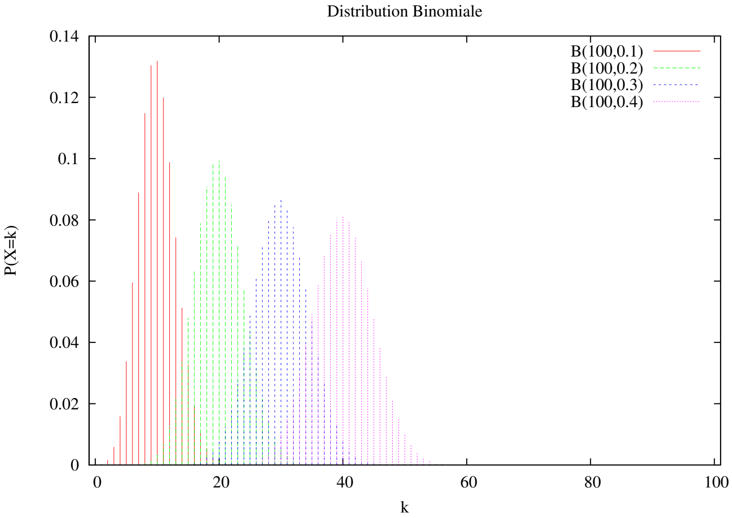
Soit X, le nombre de succès d'une épreuve de Bernoulli de paramètre p, répétée n fois indépendamment. La loi de X est appelée la *loi binomiale* de paramètres n et p :

$$\forall k \in \{0, ..., n\}, P(X = k) = C_n^k p^k (1-p)^{n-k}$$

$$E(X) = np, \quad V(X) = np(1-p)$$

Variance de la loi binomiale

Si X suit une loi binomiale, $V(X) = np(1-p)$.



Lois de probabilités discrètes : loi de Poisson

Loi de Poisson

Une variable aléatoire suit une loi de Poisson de paramètre λ si elle vérifie :

$$\forall k \geq 0, P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Propriétés

$\mathbb{E}(X) = \lambda$, $V(X) = \lambda$.
La loi de poisson est la seule loi discrète vérifiant $\mathbb{E}(X) = V(X)$.

Domaine d'application

La loi de poisson est la loi des petites probabilités ou loi des événements rares. On l'utilise pour modéliser le nombre d'occurrences d'un événement durant une période donnée quand il s'en produit en moyenne λ.

Loi de Poisson (2)

Propriétés

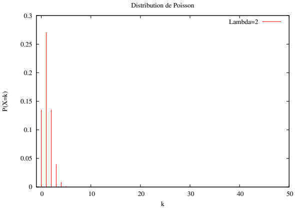
Soit $(Y_n)_{n>0}$ une suite de v.a.r. sur (Ω, P) telles que Y_n suit la loi binomiale de paramètres n et λ/n.
Soit X une v.a.r. sur (Ω, P) suivant une loi de Poisson de paramètre λ.

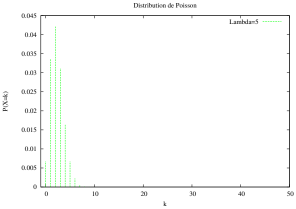
- pour tout entier k, on a:

$$\lim_{n\to\infty} P(Y_n = k) = P(X = k).$$

- On peut donc approximer une loi binomiale de paramètre n et p par une loi de Poisson de paramètre np lorsque n est grand par rapport à p.
En pratique, on peut utiliser la règle $n \geq 100$ et $np \leq 10$.

Loi de Poisson





Exemple

Le tableau ci-dessous répertorie le nombre de plantage d'un serveur sur une période de 200 jours

# plantage	0	1	2	3	4	5
# de jours	86	82	22	7	2	1

Les plantages sont survenus indépendamment les uns des autres, on peut approcher le nombre de plantage par jours avec une v.a. qui suit une loi de poisson de paramètre λ, où λ est le nombre moyen de plantage par jour.
Ainsi, l'estimation du nombre de jours où il s'est produit moins de 3 plantages est:

$$200 * P(X < 3) \approx 190$$

Lois discrètes usuelles : résumé

Nom	Loi	$\mathbb{E}(X)$	$Var(X)$
Bernoulli	$\mathbb{P}(X=k)=p^k(1-p)^{(1-k)}$	p	$p(1-p)$
binomiale	$\mathbb{P}(X=k)=C_n^k p^k(1-p)^{(n-k)}$	np	$np(1-p)$
Poisson	$\mathbb{P}(X=n)=exp(-\lambda)\lambda^n/n!$	λ	λ
géométrique	$\mathbb{P}(X=n)=p(1-p)^{n-1}$	$1/p$	$(1-p)/p^2$

Plan

Loi des grands nombres : préliminaires

Inégalité de Markov

Soit X une $v.a.r.$ non négative, et a un réel strictement positif, on a alors

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

Loi des grands nombres : préliminaires

Inégalité de Markov

Soit X une $v.a.r.$ non négative, et a un réel strictement positif, on a alors

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

Démonstration

on suppose que X est à valeurs dans $\{x_1, x_2, ...\}$, $0 \leq x_i < x_{i+1}$.
On note j, le plus petit indice tel que $x_j \geq a$.

$$aP(X \geq a) = a \sum_{i=j}^{\infty} P(X = x_i) \leq \sum_{i=j}^{\infty} x_i P(X = x_i) \leq \mathbb{E}(X)$$

Loi des grands nombres : préliminaires

Inégalité de Tchebychev

Soit X une $v.a.r$ d'espérance $\mathbb{E}(X)$ et de variance $V(X)$ et soit $\lambda \in \mathbb{R}_+^*$. On a:

$$P(|X - \mathbb{E}(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$$

Loi des grands nombres : préliminaires

Inégalité de Tchebychev

Soit X une $v.a.r$ d'espérance $\mathbb{E}(X)$ et de variance $V(X)$ et soit $\lambda \in \mathbb{R}_+^*$. On a:

$$P(|X - \mathbb{E}(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$$

Démonstration

$\forall \lambda \in \mathbb{R}_+^*$, l'événement $\{|X - \mathbb{E}(X)| \geq \lambda\}$ est égal à $\{(X - \mathbb{E}(X))^2 \geq \lambda^2\}$.
D'après l'inégalité de Markov :

$$P(|X - \mathbb{E}(X)| \geq \lambda) = P((X - \mathbb{E}(X))^2 \geq \lambda^2) \leq \frac{V(X)}{\lambda^2}.$$

Loi des grands nombres : préliminaires

Inégalité de Tchebychev : corollaire

- Inégalité : $P(|X - \mathbb{E}(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$,
- Cas particulier : si $V(X)$ est petit, alors X est proche de son espérance avec une grande probabilité.

Cas particulier

Soient $X_1, ..., X_n$, n v.a.r. indépendantes et identiquement distribuées, d'espérance m et de variance v.
Soit $Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Alors :

- $\mathbb{E}(Z_n) = m$,
- $\mathbb{V}(Z_n) = \frac{v}{n}$

La variance tend vers 0 avec $n \Rightarrow$ la moyenne est proche de l'espérance.

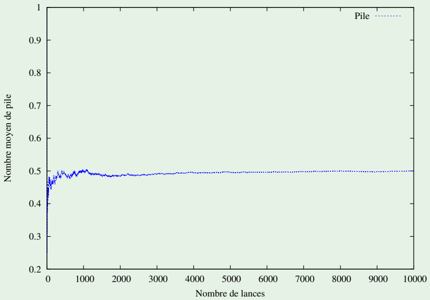
Loi des grands nombres

Énoncé

Soit $(X_n)_{n\geq 1}$ une suite de $v.a.r.$ indépendantes et identiquement distribuées, d'espérance E et de variance v finies. Alors:

$$\forall \epsilon > 0, \lim_{n\to+\infty} P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - E\right| \geq \epsilon\right) = 0$$

Exemple: Pile ou Face avec une pièce non-truquée



Loi des grands nombres

Applications

- Estimation des paramètres d'une loi binomiale ou d'une loi de poisson en calculant la moyenne des valeurs sur un échantillon.

- Estimation de probabilités par comptage :
  soit (Ω, P) un espace probabilisé, où P est *inconnue*.
  Soit E un événement. On peut estimer $P(E)$ en estimant l'espérance de la v.a. de bernoulli Y :

$$\forall \omega \in \Omega, Y(\omega) = 1 \Leftrightarrow \omega \in E$$

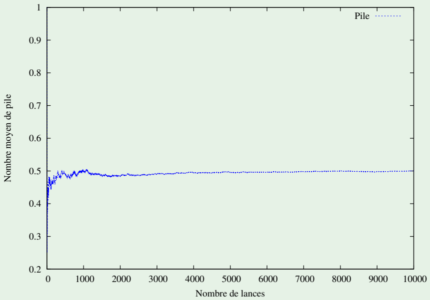
Loi des grands nombres

Énoncé

Soit $(X_n)_{n\geq 1}$ une suite de $v.a.r.$ indépendantes et identiquement distribuées, d'espérance E et de variance v finies. Alors:

$$\forall \epsilon > 0, \lim_{n\to+\infty} P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - E\right| \geq \epsilon\right) = 0$$

Exemple: Pile ou Face avec une pièce non-truquée



Lemme de Borel-Cantelli

Soit une suit A_n 'événements aléatoires indépendants et $\lim\ \sup_n\ A_n = \cap_p \cup_{n \geq p} A_n$, alors

- si la série $\sum_n P(A_n) < \infty$, alors $P(\lim\ \sup_n\ A_n) = 0$, presque surement un nombre fini de A_n sont réalisées;
- si les (A_n) sont indépendants, $\sum_n P(A_n)$ diverge alors $P(\lim\ \sup_n\ A_n) = 1$, une infinité de A_n sont réalisés presque surement.

Lemme de Borel-Cantelli

Soit une suit A_n 'événements aléatoires indépendants et $\lim\ \sup_n\ A_n = \cap_p \cup_{n \geq p} A_n$, alors

- si la série $\sum_n P(A_n) < \infty$, alors $P(\lim\ \sup_n\ A_n) = 0$, presque surement un nombre fini de A_n sont réalisées;
- si les (A_n) sont indépendants, $\sum_n P(A_n)$ diverge alors $P(\lim\ \sup_n\ A_n) = 1$, une infinité de A_n sont réalisés presque surement.

Applications

- Combien de fois k piles arrivent dans une succession infinie de lancés pile/face ?
- Quelle est la probabilité de tirer aléatoirement un nombre rationel ?