



Chapitre XI – Variables aléatoires, concentration et loi des grands nombres

Bacomathiques — <https://bacomathiqu.es>

TABLE DES MATIÈRES

| | |
|---|----------|
| I – Somme de deux variables aléatoires | 1 |
| 1. Définition | 1 |
| 2. Espérance, variance et écart-type | 1 |
| II – Somme de plusieurs variables aléatoires | 3 |
| 1. Définition et propriétés | 3 |
| 2. Somme / décompositions de certaines variables aléatoires | 3 |
| III – Concentration et loi des grands nombres | 5 |
| 1. Inégalité de Bienaymé-Tchebychev | 5 |
| 2. Inégalité de concentration | 6 |
| 3. Loi des grands nombres | 6 |

I – Somme de deux variables aléatoires

1. Définition

Il arrive que, on ne puisse pas modéliser une situation donnée à l'aide d'une seule variable aléatoire "simple". C'est pourquoi il est parfois utile d'en additionner plusieurs ou bien d'en multiplier par un réel.

À RETENIR

Définition

Soient X et Y deux variables aléatoires définies sur un univers Ω et soit λ un réel. On définit :

- $X + Y$ la variable aléatoire somme de X et Y définie pour tout $\omega \in \Omega$ par $(X + Y)(\omega) = X(\omega) + Y(\omega)$.
- λX la variable aléatoire produit de λ et X définie pour tout $\omega \in \Omega$ par $(\lambda X)(\omega) = \lambda X(\omega)$.

À LIRE

Exemple

Encore une fois, il s'agit d'une définition un peu compliquée. Illustrons ceci par un exemple.

On lance deux dés différents, équilibrés, et numérotés de 1 à 6. On note par X la variable aléatoire donnant le résultat sur lequel tombe le premier dé, et par Y la variable aléatoire donnant le résultat sur lequel tombe le second dé.

Dans cette situation, la variable aléatoire somme $X + Y$ donne la somme obtenue en additionnant le nombre sur lequel le premier dé est tombé avec celui sur lequel le deuxième dé est tombé.

2. Espérance, variance et écart-type

Une somme de variable aléatoire reste une variable aléatoire. Donc il est tout à fait possible de calculer l'espérance, la variance, l'écart-type, ... d'une somme de variables aléatoires.

Voyons dans un premier temps une propriété de l'espérance permettant de calculer plus facilement l'espérance d'une combinaison linéaire de variables aléatoires.

À RETENIR

Linéarité de l'espérance

Soient X et Y deux variables aléatoires définies sur un univers Ω et soit λ un réel. Alors :

- $E(X + Y) = E(X) + E(Y)$.
- $E(\lambda X) = \lambda E(X)$.

À LIRE

Applications

Appliquons la première formule à l'exemple de la partie précédente.

On a $E(X) = E(Y) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3,5$.

Donc $E(X + Y) = E(X) + E(Y) = 3,5 + 3,5 = 7$.

L'interprétation de ces calculs est, qu'en moyenne, sur un grand nombre de lancers, la somme obtenue lorsque l'on additionne le résultat des deux dés vaut 7.

Voyons désormais des formules permettant de calculer la variance et l'écart-type d'une combinaison linéaire de variables aléatoires.

À RETENIR

Variance et écart-type

Soient X et Y deux variables aléatoires définies sur un univers Ω et soit λ un réel. Alors :

- $V(X + Y) = V(X) + V(Y)$ si X et Y sont indépendantes (c'est-à-dire si le résultat de l'une n'a pas d'incidence sur le résultat de l'autre).
- $V(\lambda X) = \lambda^2 V(X)$.
- $\sigma(\lambda X) = \sqrt{\lambda^2} \sigma(X)$.

II – Somme de plusieurs variables aléatoires

1. Définition et propriétés

Nous allons tenter de généraliser un petit peu le concept vu dans la section précédente. En effet, au lieu d'étudier la somme de deux variables aléatoires, on va étudier la somme de n variables aléatoires.

En classe de Terminale, on se limite au cas où ces variables aléatoires suivent une même loi.

À RETENIR

Échantillon aléatoire

Un n -uplet de variables aléatoires (X_1, X_2, \dots, X_n) qui sont toutes indépendantes et qui suivent une même loi de probabilité est appelé **échantillon aléatoire de taille n associé à cette loi**.

À RETENIR

Espérance de variables aléatoires de même loi

Soit (X_1, X_2, \dots, X_n) un échantillon aléatoire de taille n . On pose $S_n = X_1 + X_2 + \dots + X_n$ et $M_n = \frac{S_n}{n}$. Alors :

- $E(S_n) = nE(X_1)$ et $V(S_n) = nV(X_1)$.
- $E(M_n) = E\left(\frac{S_n}{n}\right) = \frac{E(S_n)}{n} = E(X_1)$ et $V(M_n) = \frac{V(S_n)}{n^2} = \frac{V(X_1)}{n}$.

À LIRE

Note

Petite note sur le nom des variables aléatoires précédentes :

- S_n est la **somme** des n variables aléatoires.
- M_n est la **moyenne empirique** des n variables aléatoires.

2. Somme / décompositions de certaines variables aléatoires

Nous allons maintenant énoncer une propriété utile qui permet de trouver la loi suivie par une somme de variables aléatoires indépendantes suivant une loi de Bernoulli.

À RETENIR

Somme de variables aléatoires indépendantes suivant une même loi de Bernoulli

Soit (X_1, X_2, \dots, X_n) un échantillon aléatoire de taille n associé à une loi de Bernoulli de paramètre p .

Alors $X_1 + X_2 + \dots + X_n$ suit une loi binomiale $\mathcal{B}(n; p)$.

À LIRE

Exemple

On lance en même temps deux pièces équilibrées en l'air. On suppose qu'un succès est représenté par Pile.

On modélise le résultat de la première par une variable aléatoire X qui suit une loi de Bernoulli $\mathcal{B}(0, 5)$. De même, on modélise le résultat de la seconde par une variable aléatoire Y suivant la même loi que X .

Alors $X + Y$ (qui modélise le nombre de Pile obtenus au total par les deux pièces) suit une loi binomiale $\mathcal{B}(2; 0, 5)$.

Enfin, signalons qu'il existe une réciproque de la première propriété qui permet de "transformer" une variable aléatoire suivant une loi binomiale en somme de variables aléatoires suivant une loi de Bernoulli.

À RETENIR

Décomposition d'une variable aléatoire suivant une loi binomiale

Soit X une variable aléatoire suivant une loi binomiale $\mathcal{B}(n; p)$.

Alors il existe n variables aléatoires indépendantes X_1, X_2, \dots, X_n suivant toutes une loi de Bernoulli $\mathcal{B}(p)$, et telles que $X = X_1 + X_2 + \dots + X_n$.

À LIRE

Exemple

Soit X suivant une loi binomiale $\mathcal{B}(3; \frac{1}{6})$. Alors par la propriété précédente, il existe X_1, X_2 et X_3 , indépendantes et suivant une loi $\mathcal{B}(\frac{1}{6})$ telles que $X = X_1 + X_2 + X_3$.

III – Concentration et loi des grands nombres

1. Inégalité de Bienaymé-Tchebychev

À RETENIR

Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire d'espérance $E(X) = \mu$ et de variance $V(X) = V$. Alors pour tout réel strictement positif δ , $P(|X - \mu| \geq \delta) \leq \frac{V}{\delta^2}$.

À LIRE

Autre formulation

Une autre formulation de cette inégalité est la suivante : $P(X \notin]\mu - \delta; \mu + \delta[) \leq \frac{V(X)}{\delta^2}$.

Cette inégalité est un peu abstraite, donnons tout de suite un exemple.

À LIRE

Exemple

Le poids moyen d'un bébé en kilogrammes à la naissance peut être modélisé par une variable aléatoire X d'espérance $\mu = 3,3$ et de variance $V = 0,25$.

Un bébé est considéré "de poids normal" si son poids est compris entre 2,4 et 4,2 kilogrammes. Nous allons calculer une majoration pour la probabilité qu'un bébé ne soit pas de poids normal à la naissance (c'est-à-dire, si $X \notin]3,3 - 0,9; 3,3 + 0,9[$ ou encore si $|X - 3,3| \geq 0,9$).

On a $P(|X - 3,3| \geq 0,9) \leq \frac{0,25}{0,9^2} = 0,2025$ par l'inégalité de Bienaymé-Tchebychev.

La probabilité qu'un bébé ne soit pas de poids normal à la naissance ne dépasse pas 0,2025.

C'est majoration n'est pas très satisfaisante, mais elle vient principalement du fait que la variance est trop élevée.

2. Inégalité de concentration

À RETENIR

Inégalité de concentration

Soit (X_1, X_2, \dots, X_n) un échantillon aléatoire de taille n associé à une loi d'espérance μ et de variance V . On pose $M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$, la moyenne empirique de cet échantillon.

Alors pour tout réel strictement positif δ , $P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$.

3. Loi des grands nombres

À RETENIR

Loi faible des grands nombres

Soit (X_1, X_2, \dots, X_n) un échantillon aléatoire de taille n associé à une loi d'espérance μ et de variance V . On pose $M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$, la moyenne empirique de cet échantillon.

Alors pour tout réel strictement positif δ , $\lim_{n \rightarrow +\infty} P(|M_n - \mu| \geq \delta) = 0$.

À LIRE

Ce théorème signifie que la moyenne de l'échantillon se rapproche des moyennes des variables aléatoires quand la taille de l'échantillon augmente.

Prenons l'exemple d'une maternité au 1^{er} janvier et supposons que le premier-né soit un garçon. Il est tout à fait possible que le deuxième bébé soit également un garçon, alors que, statistiquement, on aurait pu s'attendre à une fille.

Mais l'année peut très bien commencer par une dizaine de naissances de garçons à la suite!

Cependant, si on fait un nouveau point au 31 décembre, on va se rendre compte, qu'effectivement, il y a eu environ 50 % de naissances de garçons et 50 % de naissances de filles. Il s'agit là d'un cas d'application de la loi des grands nombres.