

Chapitre 11

Somme de variables aléatoires et concentration

I. Variables aléatoires et opérations

Définition 1

Soit X et Y deux variables aléatoires définies sur l'univers Ω et w les éléments de cet univers. On peut alors définir une variable aléatoire Z sur Ω telle que, pour tout élément $w \in \Omega$, $Z(w) = X(w) + Y(w)$. La variable aléatoire Z est appelée **somme de variables aléatoires** X et Y . On note $Z = X + Y$.

Exemple

On lance deux dés cubiques équilibrés dont les faces sont numérotées de 1 à 6. On note X la variable aléatoire donnant les résultats d'un des deux dés et Y celle donnant les résultats de l'autre dé. La variable aléatoire $X + Y$ donne alors la somme des résultats des deux dés.

Remarque

On peut additionner autant de variables aléatoires que nécessaire.

Définition 2

Soit X une variable aléatoire définie sur l'univers Ω , a un nombre réel et w les éléments de l'univers Ω . On peut alors définir une variable aléatoire Y telle que, pour tout élément $\omega \in \Omega$, $Y(w) = aX(w)$. On note $Y = aX$ cette variable aléatoire.

Exemple

On lance un dé équilibré à six faces et on joue au jeu suivant : le nombre de points obtenus est le résultat du dé multiplié par 5.

En notant respectivement X et Y les variables aléatoires correspondant au résultat du dé et aux points obtenus, on a alors $Y = 5X$.

II. Espérance et variance d'une somme de variables aléatoires

II.1 Espérance

Propriété 3

Soit X et Y deux variables aléatoires définies sur le même univers Ω . Alors on a $E(X + Y) = E(X) + E(Y)$.

Démonstration

Cette démonstration est admise car elle est hors programme.

Exemple

Dans l'exemple précédent du lancer de deux dés cubiques équilibrés, on obtient l'espérance suivante pour la variable aléatoire X : $E(X) = \sum_{i=1}^6 p_i x_i = (1+2+3+4+5+6) \times \frac{1}{6} = 3,5$.

De même, on obtient l'espérance de la variable aléatoire Y : $E(Y) = 3,5$.

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

Ainsi, lorsque l'on lance ces deux dés, la somme obtenue est en moyenne 7 sur un très grand nombre de lancers.

Remarques

- On peut déterminer l'espérance d'une addition de plus de deux variables aléatoires. Si Z est aussi une variable aléatoire, alors $E(X+Y+Z) = E(X) + E(Y) + E(Z)$.
- Si Y est constante égale à b , où b est un réel, alors $E(X+b) = E(X) + b$ car $E(b) = b$.

Propriété 4

Soit X une variable aléatoire définie sur l'univers Ω et a un nombre réel.
Alors on a $E(aX) = aE(X)$.

Démonstration

Cette démonstration est admise car elle est hors programme.

Remarques

- Plus généralement, on a $E(aX + bY) = aE(X) + bE(Y)$.
- On appelle les propriétés 12 et 13 la **linéarité de l'espérance**.

II.2 Variance

Propriété 5

Soit $X_1, X_2, \dots, X_n$, n variables aléatoires à valeurs respectivement dans $E_1, E_2, \dots, E_n$.
On dit que $X_1, X_2, \dots, X_n$ sont indépendantes lorsque, pour tous $x_1 \in E_1, x_2 \in E_2, \dots, x_n \in E_n$, on a :

$$P(X_1 = x_1 \cap X_2 = x_2 \cap \dots \cap X_n = x_n) = P(X_1 = x_1) \times P(X_2 = x_2) \times \dots \times P(X_n = x_n)$$

Démonstration

Cette démonstration est admise car elle est hors programme.

Remarque

On dit aussi que les variables aléatoires $X_1, X_2, \dots, X_n$ sont mutuellement indépendantes.

Propriété 6

Soit X et Y deux variables aléatoires **indépendantes** définies sur l'univers Ω .
Alors on a $V(X+Y) = V(X) + V(Y)$.

Démonstration

Cette démonstration est admise car elle est hors programme.

Exemple

Dans l'exemple précédent du lancer de deux dés cubiques équilibrés, on obtient la variance suivante pour la variable aléatoire X :

$$V(X) = \sum_{i=1}^6 p_i (x_i - E(X))^2 = \frac{1}{6} \times (6,25 + 2,25 + 0,25 + 0,25 + 2,25 + 6,25) = \frac{1}{6} \times 17,5 = \frac{35}{12} \approx 2,92$$

De même, on obtient la variance de la variable aléatoire Y : $V(Y) = \frac{35}{12}$.

$$\text{Donc } V(X+Y) = V(X) + V(Y) = \frac{35}{12} + \frac{35}{12} = \frac{70}{12} = \frac{35}{6} \approx 5,83.$$

Remarques

- De manière intuitive, deux variables aléatoires sont indépendantes si les résultats de l'une n'ont pas influence sur les résultats de l'autre.
- Si $X_1, X_2, \dots, X_n$ sont des variables aléatoires indépendantes, alors on a $V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$.

Propriété 7

Soit X une variable aléatoire définie sur l'univers Ω dont on note $V(X)$ sa variance. Alors on a $V(aX) = a^2 V(X)$ et $\sigma(aX) = |a| \sigma(X)$.

Démonstration

Cette démonstration est admise car elle est hors programme.

Exemple

Soit X une variable aléatoire telle que $V(X) = 7$. Alors $V(3X) = 3^2 V(X) = 9 \times 7 = 63$.
De plus, $\sigma(3X) = \sqrt{V(3X)} = \sqrt{63} = 3\sqrt{7} = 3\sqrt{V(X)} = 3\sigma(X)$.

III. Applications

III.1 Loi binomiale

Définition 8

Deux variables aléatoires sont **identiquement distribuées** lorsqu'elles ont la même loi de probabilité.

Remarque

Deux variables aléatoires identiquement distribuées peuvent être ou ne pas être indépendantes.

Propriété 9

Soit $X_1, X_2, \dots, X_n$, n variables aléatoires indépendantes suivant toutes une même loi de Bernoulli de paramètre p .

Alors la variable aléatoire $X_1 + X_2 + \dots + X_n$ suit la loi binomiale de paramètres n et p .

Démonstration

Pour tout X_i , si l'on considère que l'événement « Obtenir 1 » de probabilité p comme un succès, alors $X_1 + X_2 + \dots + X_n$ donne le nombre de succès lorsque l'on réalise n fois de manière identique une même épreuve de Bernoulli. Donc $X_1 + X_2 + \dots + X_n$ suit une loi binomiale de paramètres n et p .

Exemple

Si X_i suit une loi de Bernoulli de paramètre $p=0,17$ pour tout entier $i \in [1; 12]$, alors $X = X_1 + X_2 + \dots + X_{12}$ suit la loi binomiale de paramètres $n=12$ et $p=0,17$.

Propriété 10

Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p .

Alors on a $X = X_1 + X_2 + \dots + X_n$ où les variables $X_1, X_2, \dots, X_n$ sont des variables aléatoires de Bernoulli indépendantes et identiquement distribuées.

Démonstration

Cette démonstration est admise car elle est hors programme.

Propriété 11

Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p .

- L'espérance de X est $E(X) = np$.
- La variance de X est $V(X) = np(1-p)$.
- L'écart-type de X est $\sigma(X) = \sqrt{np(1-p)}$.

Démonstration (exigible)

• Il existe n variables aléatoires de Bernoulli de paramètre p telles que $X = X_1 + X_2 + \dots + X_n$.

D'où, pour tout $i \in [1; n]$, $E(X_i) = p$.

Ainsi $E(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = p + p + \dots + p = np$.

• Les variables aléatoires $X_1, X_2, \dots, X_n$ sont indépendantes, donc par définition du schéma de Bernoulli, on a $V(X) = V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$.

Or, pour tout $i \in [1; n]$, $V(X_i) = p(1-p)$.

Ainsi $V(X) = p(1-p) + p(1-p) + \dots + p(1-p) = np(1-p)$.

• $\sigma(X) = \sqrt{V(X)} = \sqrt{np(1-p)}$.

Exemple

Soit X une variable aléatoire suivant la loi binomiale de paramètres $n=40$ et $p=0,3$.

$E(X) = np = 40 \times 0,3 = 12$, $V(X) = np(1-p) = 12(1-0,3) = 8,4$ et $\sigma(X) = \sqrt{V(X)} = \sqrt{8,4} \approx 2,9$.

V.2 Échantillonnage de n variables aléatoires identiques et indépendantes

Définition 12

Une liste $(X_1; X_2; \dots; X_n)$ de variables aléatoires indépendantes définies sur l'univers Ω suivant toutes la même loi est appelée **échantillon de taille n** associée à cette loi (ou associée à une variable aléatoire X suivant cette loi).

Exemple

On lance un dé équilibré à six faces. Une face porte le numéro 1, deux faces le numéro 2 et trois faces le numéro 3. Soit X la variable aléatoire donnant le nombre obtenu. Un échantillon de taille n de la loi suivie par X est la liste $(X_1; X_2; X_3; X_4)$ où chacun des X_i suit la loi de X . Cela correspond concrètement à la liste de quatre résultats de quatre lancers du dé.

Définition 13

On considère n expériences aléatoires identiques et indépendantes. On note $X_1, X_2, \dots, X_n$ les variables aléatoires associées à ces expériences, toutes de même loi.

On appelle respectivement S_n et M_n les **variables aléatoires somme et moyenne** d'un

échantillon de taille n des X_i . On a $S_n = \sum_{i=1}^n X_i$ et $M_n = \frac{1}{n} \sum_{i=1}^n X_i = \frac{S_n}{n}$.

Propriété 14

Soit $(X_1; X_2; \dots; X_n)$ un échantillon de taille n d'une variable aléatoire X .

- Les espérances de S_n et M_n sont $E(S_n) = n E(X_i)$ et $E(M_n) = E(X_i)$.
- Les variances de S_n et M_n sont $V(S_n) = n V(X_i)$ et $V(M_n) = \frac{V(X_i)}{n}$.
- Les écarts-types de S_n et M_n sont $\sigma(S_n) = \sqrt{n} \sigma(X_i)$ et $\sigma(M_n) = \frac{\sigma(X_i)}{\sqrt{n}}$.

IV. Inégalités de concentration

IV.1 Inégalité de Markov

Définition 15

Une variable aléatoire est dite positive ou nulle dans un univers Ω lorsque toutes les valeurs prises par celle-ci sont des réels positifs ou nuls.

Exemple

La variable aléatoire donnant le nombre de faces numérotées 1 obtenues sur dix lancers d'un dé est positive ou nulle.

Théorème 16 (inégalité de Markov)

Soit X une variable aléatoire réelle à valeurs positives ou nulles d'espérance $E(X)$.

Pour tout réel δ strictement positif, on a $P(X \geq \delta) \leq \frac{E(X)}{\delta}$.

Démonstration

Soit X une variable aléatoire réelle positive ou nulle dont on note x_i les n valeurs pour l'entier i allant de 1 à n .

Par définition de l'espérance, on a $E(X) = \sum_{i=1}^n x_i P(X=x_i)$.

Séparons cette somme en deux blocs en considérant les valeurs supérieures ou égales à δ et celles strictement inférieures à δ .

On obtient $E(X) = \sum_{x_i \geq \delta} x_i P(X=x_i) + \sum_{x_i < \delta} x_i P(X=x_i)$.

Pour tout entier i allant de 1 à n , on sait que $x_i \geq 0$ (car X est positive ou nulle) et $P(X=x_i) \geq 0$ (par définition d'une probabilité), donc $\sum_{x_i < \delta} x_i P(X=x_i) \geq 0$.

On en déduit que $E(X) \geq \sum_{x_i \geq \delta} x_i P(X=x_i)$.

Par ailleurs, dans cette partie de la somme, pour tout entier i allant de 1 à n , $x_i \geq \delta$.

Donc $E(X) \geq \sum_{x_i \geq \delta} x_i P(X=x_i) \geq \sum_{x_i \geq \delta} \delta P(X=x_i)$.

Or $\sum_{x_i \geq \delta} \delta P(X=x_i) = \delta \sum_{x_i \geq \delta} P(X=x_i) = \delta P(X \geq \delta)$.

Par conséquent, $E(X) \geq \delta \sum_{x_i \geq \delta} P(X=x_i)$ ou encore $E(X) \geq \delta P(X \geq \delta)$.

Or $\delta > 0$. Ainsi $P(X \geq \delta) \leq \frac{E(X)}{\delta}$.

Remarques

- Il est nécessaire que la variable aléatoire X soit positive ou nulle.
- Si $\delta \leq E(X)$, l'inégalité de Markov n'a pas d'intérêt (on dit qu'elle est triviale).

En effet, la borne $\frac{E(X)}{\delta}$ est alors supérieure à 1 et donc nécessairement à la probabilité $P(X \geq \delta)$.

- Cette inégalité permet de trouver un majorant mais pas forcément le plus petit possible.

Interprétation

Ce résultat signifie que la probabilité que X prenne des valeurs plus grandes que δ est d'autant plus petite que δ est grand.

Exemples

- Soit X une variable aléatoire positive d'espérance 1.
D'après l'inégalité de Markov, on a $P(X \geq 100) \leq 0,01$. Autrement dit, une variable aléatoire positive dont l'espérance est 1 a au plus une chance sur cent de dépasser 100.
- En 2015, le salaire brut mensuel moyen en France était de 2442€. On choisit un salarié au hasard et on note X la variable aléatoire donnant son salaire. Les salaires étant positifs ou nuls, on sait que X est une variable aléatoire réelle positive ou nulle.
On peut donc appliquer l'inégalité de Markov sur un exemple :
On a $P(X \geq 7326) \leq \frac{2442}{7326}$, soit $P(X \geq 7326) \leq \frac{1}{3}$.

IV.2 Inégalité de Bienaymé-Tchebychev

Théorème 17 (inégalité de Bienaymé-Tchebychev)

Soit X une variable aléatoire réelle d'espérance $E(X) = \mu$ et de variance $V(X)$.

Pour tout réel δ strictement positif, on a $P(|X - \mu| \geq \delta) \leq \frac{V(X)}{\delta^2}$.

Démonstration

Comme $\delta > 0$, les inégalités $|X - \mu| \geq \delta$ et $(X - \mu)^2 \geq \delta^2$ sont équivalentes.

De plus, la variable $(X - \mu)^2$ est positive ou nulle.

On applique donc l'inégalité de Markov à la variable $(X - \mu)^2$ et au réel δ^2 .

On a alors $P((X - \mu)^2 \geq \delta^2) \leq \frac{E((X - \mu)^2)}{\delta^2}$. Or $E((X - \mu)^2) = V(X)$. D'où $P((X - \mu)^2 \geq \delta^2) \leq \frac{V(X)}{\delta^2}$.

Ainsi, on en déduit que $P(|X - \mu| \geq \delta) \leq \frac{V(X)}{\delta^2}$.

Remarques

- La variable aléatoire $|X - \mu|$ est positive ou nulle.
- On a les équivalences $P(|X - \mu| < \delta) \geq 1 - \frac{V(X)}{\delta^2}$ ou encore $P(X \in]\mu - \delta; \mu + \delta]) \geq 1 - \frac{V(X)}{\delta^2}$.
- On dit que $[\mu - \delta; \mu + \delta]$ est un intervalle de fluctuation de X .
- Cette inégalité est loin d'être optimale. En réalité, il est fort possible que la probabilité soit bien inférieure au majorant obtenu.

Interprétation

La probabilité que les valeurs prises par X s'écartent d'au moins δ de l'espérance μ est d'autant plus petite que δ est grand.

Exemples

- Si $\delta = 2\sigma(X)$ où $\sigma(X)$ est l'écart-type de la variable X , alors :

$$P(|X - \mu| \geq 2\sigma(X)) \leq \frac{V(X)}{(2\sigma(X))^2} = \frac{1}{4} \quad (\text{car } \sigma(X) = \sqrt{V(X)})$$

Autrement dit, la probabilité qu'une variable aléatoire prenne des valeurs éloignées de son espérance d'au moins le double de son écart-type est inférieure à 0,25.

- Dans une usine, la variable aléatoire L donnant la largeur en millimètres d'une puce électronique prise au hasard a pour espérance $\mu = 12$ et pour variance $V(L) = 0,01$.

Si la largeur d'une puce n'appartient pas à $]11; 13[$, c'est-à-dire si $|L - 12| \geq 1$, la puce n'est pas commercialisable. La probabilité qu'une puce ne soit pas commercialisable est donc :

$$P(|L - 12| \geq 1) \quad (\text{c'est-à-dire } P(|L - \mu| \geq \delta) \text{ avec } \mu = 12 \text{ et } \delta = 1)$$

Comme $V(L) = 0,01$, d'après l'inégalité de Bienaymé-Tchebychev, on a :

$$P(|L - 12| \geq 1) \leq \frac{0,01}{1^2} \quad \text{soit } P(|L - 12| \geq 1) \leq 0,01$$

IV.3 Inégalité de concentration**Théorème 18 (inégalité de concentration)**

Soit $X_1, X_2, \dots, X_n$ un échantillon de variables aléatoires d'espérance μ et de variance V et M_n la variable aléatoire moyenne de cet échantillon.

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

Démonstration

D'après les propriétés sur l'espérance et la variance de la variable aléatoire moyenne d'un échantillon, on a $E(M_n) = \mu$ et $V(M_n) = \frac{V}{n}$.

En appliquant l'inégalité de Bienaymé-Tchebychev à M_n , on obtient $P(|M_n - \mu| \geq \delta) \leq \frac{V(M_n)}{\delta^2}$.

Or $\frac{V(M_n)}{\delta^2} = \frac{\frac{V}{n}}{\delta^2} = \frac{V}{n\delta^2}$. Ainsi, on en déduit que $P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$.

Remarque

On a les équivalences $P(|M_n - \mu| < \delta) \geq 1 - \frac{V(X)}{n\delta^2}$ ou encore $P(M_n \in]\mu - \delta; \mu + \delta[) \geq 1 - \frac{V(X)}{n\delta^2}$.

Exemple

On lance n fois un dé équilibré à 8 faces et on nomme X_i la variable aléatoire donnant le résultat du i -ème lancer. On admet que $E(X_i) = 4,5$ et $V(X_i) = 5,25$ pour tout entier $i \in [1; n]$.

Les lancers étant indépendants, les $X_1, X_2, \dots, X_n$ sont un échantillon de variables aléatoires

d'espérance $\mu = 4,5$, de variance $V = 5,25$ et de moyenne $M_n = \frac{1}{n} \sum_{i=1}^n X_i$.

D'après l'inégalité de concentration pour $n=100$ et $\delta=0,5$: $P(|M_{100}-4,5|\geq 0,5) \leq \frac{5,25}{100 \times 0,5^2}$

Or $\frac{5,25}{100 \times 0,5^2} = \frac{5,25}{25} = 0,21$. Donc $P(|M_{100}-4,5|\geq 0,5) \leq 0,21$.

Ainsi, la probabilité que l'écart entre M_{100} (la moyenne des 100 premiers résultats) et 4,5 soit supérieur ou égal à 0,5 est inférieur ou égale à 0,21.

V. Loi des grands nombres

Théorème 19 (loi faible des grands nombres)

Soit $X_1, X_2, \dots, X_n$ un échantillon de variables aléatoires d'espérance μ et de variance $V(X)$ et M_n la variable aléatoire moyenne de cet échantillon.

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

On dit que M_n **converge en probabilité** vers μ lorsque n tend vers $+\infty$.

Démonstration

On applique l'inégalité de concentration à la variable aléatoire M_n .

On a $0 \leq P(|M_n - \mu| \geq \delta) \leq \frac{V(X)}{n \delta^2}$. Puisque $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$, par somme, on a $\lim_{n \rightarrow +\infty} \frac{V(X)}{n \delta^2} = 0$.

Ainsi, d'après le théorème d'encadrement, on en déduit que $\lim_{n \rightarrow +\infty} P(|M_n - \mu| \geq \delta) = 0$.

Remarque

Il existe un théorème similaire dit loi forte des grands nombres.

Exemple

On reprend l'exemple précédent et on considère $\delta=0,1$.

D'après la loi des grands nombres, $\lim_{n \rightarrow +\infty} P(|M_n - 4,5| \geq 0,1) = 0$.

On en déduit que $\lim_{n \rightarrow +\infty} P(|M_n - 4,5| < 0,1) = 1$ (ou encore $\lim_{n \rightarrow +\infty} P(M_n \in]4,4; 4,6[) = 1$).

Autrement dit, si l'on fait un nombre suffisamment grand de lancers, on peut rendre l'événement « la moyenne de l'échantillon est dans $]4,4; 4,6[$ » aussi probable qu'on le souhaite en prenant n suffisamment grand.