Somme de variables aléatoires, concentration et loi des grands nombres

L'essentiel : Tle Spé

Mathématiques maths-mde.fr

Linéarité de l'espérance

Soient X et Y deux variables aléatoires. L'espérance de la variable aléatoire $X+Y$ est :

$$E(X+Y) = E(X) + E(Y).$$

De plus, pour tout réel a, l'espérance de la variable aléatoire aX est :

$$E(aX) = aE(X).$$

Variance d'une variables aléatoire

✧ Soit X une variable aléatoire.
Pour tout réel a, la variance de la variable aléatoire aX est : $V(aX) = a^2V(X)$.
Son écart-type est donc : $\sigma(aX) = |a|\sigma(X)$.

✧ Soient X et Y deux variables aléatoires *indépendantes*.
La variance de la variable aléatoire $X+Y$ est : $V(X+Y) = V(X) + V(Y)$.

Interprétation

Maria et Sara vont au restaurant. Maria hésite de façon équiprobable entre les menus à 18 €, 24 € et 36 €. Sara hésite, quant à elle, de façon équiprobable entre les menus à 24 € et 36 €.
On note X et Y les variables aléatoires représentant le prix du menu choisi respectivement par Maria et Sara. Alors, $X = \{18; 24; 36\}$ et $Y = \{24; 36\}$.
Ainsi, $E(X) = \frac{1}{3} \times 18 + \frac{1}{3} \times 24 + \frac{1}{3} \times 36 = 26$ et $E(Y) = \frac{1}{2} \times 24 + \frac{1}{2} \times 36 = 30$.
En conséquence, $E(X+Y) = E(X) + E(Y) = 26 + 30 = 56$. On peut ainsi dire que le prix moyen de la facture de ce repas est 56 €.

Si Maria va seul au restaurant 5 fois, et si elle choisit à chaque fois au hasard un menu parmi les 3 cités précédemment, la facture moyenne totale correspondra à $E(5X)$, soit à $5E(X)$, et donc à $5 \times 26 = 130$ €.
Par ailleurs, $V(X) = \frac{1}{3}\left[(18-26)^2 + (24-26)^2 + (36-30)^2\right] = 56$ et donc, $V(5X) = 5^2V(X) = 25 \times 56 = 1400$. Soit, $\sigma(5X) = 5\sigma(X) \approx 37,42$.
Ce dernier résultat peut-être interprété comme une marge d'erreur par rapport à l'espérance : on peut alors dire que pour 5 repas, Maria paiera 130 € avec une marge d'erreur de 37,42 €. Cet intervalle $[130 - 37,42\ ;\ 130 + 37,42]$ est très approximatif !

Somme et moyenne de variables indépendantes

Soient X_1, ..., X_n n variables aléatoires indépendantes. On pose : $S_n = \sum_{k=1}^{n} X_k = X_1 + \cdots + X_n$ et $M_n = \frac{S_n}{n}$. On a alors :

✧ $E(S_n) = \sum_{k=1}^{n} E(X_k)$ et $V(S_n) = \sum_{k=1}^{n} V(X_k)$.

✧ $E(M_n) = \frac{1}{n}\sum_{k=1}^{n} E(X_k)$ et $V(M_n) = \frac{1}{n^2}\sum_{k=1}^{n} V(X_k)$.

Inégalité de Markov

Soit X une variable aléatoire discrète d'espérance finie, et à valeurs positives. Alors :

$$\forall a > 0, \quad P(X \geqslant a) \leqslant \frac{E(X)}{a}.$$

Inégalité de concentration

Soient $X_1, X_2, \ldots, X_n$ n variables aléatoires discrètes réelles indépendantes ayant toutes la même loi d'espérance μ et de variance σ^2. Alors,

$$\forall \delta > 0, \quad P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geqslant \delta\right) \leqslant \frac{\sigma^2}{n\delta^2}.$$

Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire d'espérance μ et de variance $V(X)$. Quel que soit le réel $\delta > 0$,

$$P\big(|X - \mu| \geqslant \delta\big) \leqslant \frac{V(X)}{\delta^2}.$$

De cette inégalité, on peut notamment conclure que pour une variable aléatoire quelconque X d'écart-type σ et d'espérance μ :

$$P\big(|X - \mu| \geqslant 2\sigma\big) \leqslant \frac{\sigma^2}{(2\sigma)^2}$$

soit,

$$P\big(|X - \mu| \geqslant 2\sigma\big) \leqslant \frac{1}{4}.$$

Cela signifie alors que la probabilité que les valeurs prises par X diffèrent de 2σ de sa moyenne est inférieure à 0,25.

Loi des grands nombres

Soient X_1, ..., X_n n variables aléatoires indépendantes de même loi, d'espérance μ. Alors, pour tout réel $\delta > 0$,

$$\lim_{n \to +\infty} P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geqslant \delta\right) = 0.$$