

Sommes de variables aléatoires et loi des grands nombres

La **moyenne d'un échantillon** a la même espérance que X mais une variance divisée par n : c'est le cœur de la loi des grands nombres.

Propriété — espérance et variance d'une somme

$E(X+Y) = E(X) + E(Y)$ et $E(aX) = aE(X)$ (linéarité, toujours vraie). Si X et Y sont **indépendantes** : $V(X+Y) = V(X) + V(Y)$; $V(aX) = a^2V(X)$.

Définition — échantillon, moyenne

Un échantillon de taille n de X est une liste $(X_1, \dots, X_n)$ de variables indépendantes de même loi que X . Somme $S_n = X_1 + \dots + X_n$ et moyenne $M_n = \frac{S_n}{n}$:

$$E(S_n) = nE(X), \quad V(S_n) = nV(X) \quad E(M_n) = E(X), \quad V(M_n) = \frac{V(X)}{n}, \quad \sigma(M_n) = \frac{\sigma(X)}{\sqrt{n}}.$$

Propriété — loi binomiale comme somme

Une variable $\mathcal{B}(n; p)$ est la somme de n variables de Bernoulli indépendantes de paramètre p , d'où $E = np$ et $V = np(1-p)$.

Inégalité de Bienaymé-Tchebychev

Pour tout $\delta > 0$:

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

(Inégalité de Markov pour $X \geq 0$: $P(X \geq a) \leq \frac{E(X)}{a}$.)

Inégalité de concentration et loi des grands nombres

$P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2}$. Par conséquent, pour tout $\delta > 0$,

$\lim_{n \rightarrow +\infty} P(|M_n - E(X)| \geq \delta) = 0$: la moyenne d'un grand échantillon est proche de l'espérance (**loi faible des grands nombres**).

Exemple

Dé équilibré, $E(X) = 3,5$, $V(X) \approx 2,92$. Sur $n = 1000$ lancers,

$$P(|M_n - 3,5| \geq 0,2) \leq \frac{2,92}{1000 \times 0,04} \approx 0,073.$$

Questions fréquentes sur ce chapitre

Quelle est l'espérance d'une somme de variables aléatoires ?

$E(X+Y) = E(X) + E(Y)$ toujours ; $V(X+Y) = V(X) + V(Y)$ seulement si X et Y sont indépendantes.

Que dit l'inégalité de Bienaymé-Tchebychev ?

$P(|X - E(X)| \geq \delta) \leq \frac{V(X)}{\delta^2}$: la probabilité de s'éloigner de l'espérance est contrôlée par la variance.

Qu'est-ce que la loi des grands nombres ?

La moyenne M_n d'un échantillon de taille n se rapproche de l'espérance quand n grandit : $P(|M_n - \mu| \geq \delta) \rightarrow 0$.

Comment évolue l'écart-type d'une moyenne d'échantillon ?

$\sigma(M_n) = \frac{\sigma}{\sqrt{n}}$: pour diviser l'écart-type par 2, il faut multiplier la taille de l'échantillon par 4.