

Chapitre XVIII

Loi des Grands Nombres

I. INÉGALITÉS PROBABILITISTES

A. INÉGALITÉ DE MARKOV

Pour toute variable aléatoire X à valeurs positives, et pour tout nombre réel δ strictement positif :

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

1. DÉMONSTRATION

$$\begin{aligned} E(X) &= \sum_{i=1}^n x_i P(X = x_i) \\ &= \overbrace{\sum_{x_i < \delta} x_i P(X = x_i)}^{\geq 0} + \sum_{x_i \geq \delta} x_i P(X = x_i) \\ &\geq \sum_{x_i \geq \delta} x_i P(X = x_i) \quad \text{on minore} \\ &\geq \sum_{x_i \geq \delta} \delta P(X = x_i) = \delta \overbrace{\sum_{x_i \geq \delta} P(X = x_i)}^{P(X \geq \delta)} \quad \text{car } x_i \geq \delta \text{ on minore encore} \\ E(X) &\geq \delta P(X \geq \delta) \\ \Leftrightarrow P(X \geq \delta) &\leq \frac{E(X)}{\delta} \end{aligned}$$

B. INÉGALITÉ DE BIENAYMÉ-TCHEBYCHEV

Pour toute variable aléatoire X , et pour tout nombre réel δ strictement positif :

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

1. DÉMONSTRATION

$$P(|X - E(X)| \geq \delta) = P((X - E(X))^2 \geq \delta^2)$$

D'après l'inégalité de MARKOV :

$$P((X - E(X))^2 \geq \delta^2) \leq \frac{\overbrace{E((X - E(X))^2)}^{V(X)}}{\delta^2} \quad \square$$

C. REMARQUE

Souvent, on prend $\delta = \sigma$ ou $k\sigma$ car l'écart type σ d'une variable aléatoire X est l'unité naturelle pour étudier la dispersion de X autour de son espérance.

L'inégalité de BIENAYMÉ-TCHEBYCHEV montre notamment que des écarts de X à $E(X)$ de quelques σ deviennent improbables.

D. REMARQUE

L'Inégalité de Bienaymé-Tchebychev est loin de donner le meilleur majorant.

Si X est une variable aléatoire suivant la loi binomiale $\mathcal{B}(n; p)$, on a

$P(|X - E(X)| \geq 2\sigma) \approx 0,05$ ce qui est bien meilleur que le 0,25 obtenu avec l'inégalité.

En effet, si $\delta = 2\sigma$ et $X \sim \mathcal{B}(n; p)$:

$$\frac{V(X)}{\delta^2} = \frac{np(1-p)}{(4\sqrt{np(1-p)})^2} = \frac{1}{4}$$

II. LOI DES GRANDS NOMBRES

A. PROPRIÉTÉ

Soit une variable aléatoire X associée à un échantillon (X_1, X_2, \dots, X_n) , c'est-à-dire que (X_1, X_2, \dots, X_n) est un échantillon de n variables aléatoires identiques et indépendantes qui suivent toutes la loi de probabilité suivie par X . On note M_n la moyenne de cet échantillon.

Alors, pour tout réel $\delta > 0$:

$$P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2} \quad (\text{c'est l'inégalité de concentration})$$

$$\lim_{n \rightarrow +\infty} P(|M_n - E(X)| \geq \delta) = 0 \quad (\text{c'est la loi des grands nombres})$$

Autrement dit, plus la taille n d'un échantillon d'une variable aléatoire X est grande, plus l'écart entre la moyenne de cet échantillon et l'espérance de X est faible.

1. DÉMONSTRATION

A. RAPPELS

Si X et Y sont deux variables aléatoires avec Y telle que $Y = aX + b$ alors :

$$— E(Y) = aE(X) + b \quad \text{et} \quad V(Y) = a^2V(X)$$

$$— E(X + Y) = E(X) + E(Y) \quad \text{et, si } X \text{ et } Y \text{ indépendantes} \quad V(X+Y) = V(X) + V(Y)$$

Pour un échantillon $\{X_1; X_2; \dots; X_n\}$ où X_i sont indépendantes, et suivent une loi X :

$$— S_n = X_1 + X_2 + \dots + X_n \quad E(S_n) = nE(X) \quad V(S_n) = nV(X)$$

$$— M_n = \frac{S_n}{n} \quad E(M_n) = E(X) \quad V(M_n) = \frac{V(X)}{n}$$

Suite Démonstration :

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

$$\Leftrightarrow P(|M_n - E(M_n)| \geq \delta) \leq \frac{V(M_n)}{\delta^2} \quad \text{Inégalité de BIENAYMÉ-TCHÉBYCHEV}$$

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

On a donc, pour un δ fixe :

$$0 \leq P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2}$$

Et, $\lim_{n \rightarrow +\infty} \frac{V(X)}{n\delta^2} = 0$, donc, d'après le théorème des Gendarmes :

$$\lim_{n \rightarrow +\infty} P(|M_n - E(X)| \geq \delta) \leq \frac{V(X)}{n\delta^2} = 0$$