Chapitre 18

Loi des Grands Nombres

I. Inégalités Probabilitstes

A. Inégalité de Markov

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

$$P\left(X \geqslant \delta\right) \leqslant \frac{E(X)}{\delta}$$

1. Démonstration

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

B. Inégalité de Bienaymé-Tchebychev

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

$$\boxed{\mathrm{P}\left(|\mathrm{X}-\mathrm{E(X)}| \geqslant \delta\right) \leqslant \frac{\mathrm{V(X)}}{\delta^2}}$$

1. Démonstration

$\mathrm{P}\left(|\mathrm{X}-\mathrm{E(X)}| \geqslant \delta\right) = \mathrm{P}\left((\mathrm{X}-\mathrm{E(X)})^2 \geqslant \delta^2\right)$

D'après l'inégalité de Markov :

$$\mathrm{P}\left((\mathrm{X}-\mathrm{E(X)})^2 \geqslant \delta^2\right) \leqslant \frac{\overbrace{\mathrm{E}\left((\mathrm{X}-\mathrm{E(X)})^2\right)}^{\mathrm{V(X)}}}{\delta^2} \quad \square$$

C. Remarque

Souvent, on prendre $\delta = \sigma$ 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 sons 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 $\mathscr{B}\left(n\,;p\right)$, on a $\mathrm{P}\left(|\mathrm{X}-\mathrm{E(X)}| \geqslant 2\sigma\right) \approx 0{,}05$ ce qui est bien meilleur que le 0,25 obtenu avec l'inégalité.

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

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

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 $\delta > 0$:

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

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

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 :

— $E(Y) = aE(X) + b$ et $V(Y) = a^2V(X)$

— $E(X+Y) = E(X) + E(Y)$ et, si X et Y indépendantes $V(X+Y)=V(X)+V(Y)$

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

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

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

Suite Démonstration :

$$P\left(|X - E(X)| \geqslant \delta\right) \leqslant \frac{V(X)}{n\delta^2}$$

$$\iff P\left(|M_n - E(M_n)| \geqslant \delta\right) \leqslant \frac{V(M_n)}{\delta} \quad \text{Inégalité de Bienaymé-Tchébychev}$$

$$\iff P\left(|M_n - E(X)| \geqslant \delta\right) \leqslant \frac{V(X)}{n\delta^2}$$

On a donc, pour un δ fixe :

$$0 \leqslant P\left(|M_n - E(X)| \geqslant \delta\right) \leqslant \frac{V(X)}{n\delta^2}$$

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

$$\lim_{n\to+\infty} P\big(|M_n - E(X)| \geqslant \delta\big) \leqslant \frac{V(X)}{n\delta^2} = 0$$