Chapitre VII — Variables aléatoires et lois de concentration

1 Rappels

Définition

Une variable aléatoire réelle X est une fonction définie sur un univers Ω muni d'une loi de probabilité P.

Définition

Espérance : Soit X une variable aléatoire discrète prenant les valeurs $\{x_1, x_2, \ldots, x_k\}$.

$$E(X) = \sum_{i=1}^{k} x_i P(X = x_i)$$

Définition

Variance et Écart-type : Soit X une variable aléatoire d'espérance $\mu = E(X)$.

- La **variance** de X, notée $V(X)$, est la moyenne du carré des écarts à la moyenne :

$$V(X) = E\Big((X - \mu)^2\Big)$$

- En pratique, on utilise souvent la formule :

$$V(X) = E(X^2) - \Big(E(X)\Big)^2$$

- L'**écart-type** de X est défini par : $\sigma(X) = \sqrt{V(X)}$.

Propriété

Pour tous réels a et b :

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

Propriété

Si X suit une loi binomiale $\mathcal{B}(n, p)$:

$$E(X) = np \quad \text{et} \quad V(X) = np(1 - p)$$

2 Somme de variables aléatoires

Définition

Deux variables aléatoires X et Y sont **indépendantes** si pour tous réels x_i et y_j de leurs ensembles de valeurs respectifs :

$$P((X = x_i) \cap (Y = y_j)) = P(X = x_i) \times P(Y = y_j)$$

Propriété

Pour toutes variables aléatoires X et Y :

$$E(X + Y) = E(X) + E(Y) \quad \text{(Linéarité, toujours vrai)}$$

Propriété

Si X et Y sont indépendantes :

$$V(X + Y) = V(X) + V(Y)$$

3 Inégalité de Bienaymé-Tchebychev

Propriété

Pour toute variable aléatoire X d'espérance $\mu = E(X)$ et de variance $V(X)$:

$$\forall a > 0, \quad P(|X - \mu| \geq a) \leq \frac{V(X)}{a^2}$$

Remarque : On utilise souvent la forme équivalente (concentration) : $P(|X - \mu| < a) \geq 1 - \frac{V(X)}{a^2}$.

Démonstration. On pose $A = \{\omega \in \Omega \mid |X(\omega) - \mu| \geq a\}$.
Pour tout $\omega \in A$, on a $|X(\omega) - \mu|^2 \geq a^2$.
Par définition de la variance :

$$V(X) = \sum_{\omega \in \Omega} (X(\omega) - \mu)^2 P(\omega)$$

Comme tous les termes de la somme sont positifs, on peut minorer la somme sur Ω par la somme sur les éléments de A :

$$V(X) \geq \sum_{\omega \in A} (X(\omega) - \mu)^2 P(\omega) \geq \sum_{\omega \in A} a^2 P(\omega) = a^2 P(A)$$

D'où $V(X) \geq a^2 P(A)$, ce qui donne $P(A) \leq \frac{V(X)}{a^2}$.

4 Échantillon

Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires **indépendantes et de même loi**.

Définition

Somme : $S_n = \sum_{i=1}^{n} X_i$ et Moyenne : $M_n = \frac{S_n}{n}$

Propriété

L'espérance est conservée, mais la variance décroît avec la taille de l'échantillon :

$$E(M_n) = E(X) \quad \text{et} \quad V(M_n) = \frac{V(X)}{n}$$

5 Loi faible des grands nombres

Propriété

Pour tout $a > 0$, l'inégalité de concentration pour la moyenne s'écrit :

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

Propriété

(Loi faible des grands nombres)

$$\lim_{n \to +\infty} P(|M_n - E(X)| < a) = 1$$

Interprétation : Ce résultat permet de justifier l'estimation d'une probabilité par une fréquence observée sur un échantillon de grande taille.