

# LES MOINDRES CARRÉS

OLIVIER CASTÉRA

RÉSUMÉ. La méthode des moindres carrés repose sur un fondement probabiliste sérieux.

## TABLE DES MATIÈRES

1. Justification de la méthode des moindres carrés	1
2. Caractéristiques numériques des systèmes de deux variables	2
3. Caractéristiques statistiques d'une répartition	3
4. Application	4

### 1. JUSTIFICATION DE LA MÉTHODE DES MOINDRES CARRÉS

**Théorème 1.1.** *Lorsque les erreurs de mesure suivent une loi normale, pour qu'un ensemble de valeurs observées  $(y_1, y_2, \dots, y_n)$  d'une fonction à déterminer  $y = \varphi(x)$  soit le plus probable, il faut choisir cette fonction de telle sorte que la somme des carrés des écarts des valeurs observées par rapport à  $\varphi(x)$  soit minimale :*

$$d \sum_{i=1}^n [y_i - \varphi(x_i)]^2 = 0$$

*Démonstration.* Soit  $y = \varphi(x)$  l'expression exacte de la relation réelle existant entre  $y$  et  $x$ . Les points expérimentaux s'écartent de cette relation par suite des erreurs inévitables de mesure. Le théorème central limite<sup>1</sup> montre que les erreurs de mesure suivent généralement la loi normale<sup>2</sup>. Supposons qu'il en soit ainsi, et choisissons une valeur  $x_i$  de l'argument pour laquelle nous effectuons une mesure. Le résultat de cette mesure est l'évènement  $Y_i = y_i$  où  $Y_i$  est une variable aléatoire répartie suivant une loi normale d'espérance mathématique  $\varphi(x_i)$ , et d'écart quadratique moyen  $\sigma_i$ . L'espérance mathématique  $\varphi(x_i)$  est le résultat que l'on trouverait s'il n'y avait pas d'erreur de mesure, et  $\sigma_i$  caractérise l'erreur de mesure. Supposons que l'erreur de mesure soit

---

*Date:* 4 mai 2015.

1. Voir Théorème central limite.pdf

2. Voir Loi normale.pdf

la même en tout point :

$$\forall i \sigma_i = \sigma$$

La densité de probabilité de la variable  $Y_i$  s'écrit :

$$f_i(y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-[y_i - \varphi(x_i)]^2 / (2\sigma^2)}$$

Supposons que l'expérience, qui est dans notre exemple une série de mesures, ait réalisé l'évènement consistant en ce que les variables aléatoires  $(Y_1, Y_2, \dots, Y_n)$  aient pris les valeurs  $(y_1, y_2, \dots, y_n)$ . Les variables  $Y_i$  étant continues, la probabilité de chacun des évènements  $Y_i = y_i$  est nulle, c'est pourquoi nous considérerons les éléments de probabilité correspondants :

$$f_i(y_i) dy_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-[y_i - \varphi(x_i)]^2 / (2\sigma^2)} dy_i$$

Les mesures étant indépendantes, la probabilité pour que le système de variables aléatoires  $(Y_1, Y_2, \dots, Y_n)$  prenne l'ensemble des valeurs se trouvant chacune dans l'intervalle  $(y_i, y_i + dy_i)$  est égale au produit des éléments de probabilité :

$$\begin{aligned} \prod_{i=1}^n f_i(y_i) dy_i &= \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-[y_i - \varphi(x_i)]^2 / (2\sigma^2)} dy_i \\ &= \left( \frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-1/(2\sigma^2) \sum_{i=1}^n [y_i - \varphi(x_i)]^2} \prod_{i=1}^n dy_i \end{aligned}$$

Pour que cette probabilité soit maximale, il faut que :

$$\sum_{i=1}^n [y_i - \varphi(x_i)]^2$$

soit minimale. □

## 2. CARACTÉRISTIQUES NUMÉRIQUES DES SYSTÈMES DE DEUX VARIABLES

**Définition 2.1.** On appelle moment initial d'ordre  $k, s$  du système de variables aléatoires  $(X, Y)$ , l'espérance mathématique du produit de  $X^k$  par  $Y^s$

$$\alpha_{k,s}[X^k Y^s] \triangleq M[X^k Y^s]$$

Pour des variables aléatoires discrètes :

$$\alpha_{k,s}[X^k Y^s] \triangleq \sum_{i=1}^n \sum_{j=1}^m x_i^k y_j^s p_{ij}$$

Pour des variables aléatoires continues :

$$\alpha_{k,s}[X^k Y^s] \triangleq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^k y^s f(x, y) dx dy$$

**Définition 2.2.** On appelle moment centré d'ordre  $k, s$  du système de variables aléatoires  $(X, Y)$ , l'espérance mathématique du produit de  $\overset{\circ}{X}^k$  par  $\overset{\circ}{Y}^s$  :

$$\mu_{k,s}[X^k Y^s] \triangleq M[\overset{\circ}{X}^k \overset{\circ}{Y}^s]$$

Pour des variables aléatoires discrètes :

$$\mu_{k,s}[X^k Y^s] \triangleq \sum_{i=1}^n \sum_{j=1}^m (x_i - m_x)^k (y_j - m_y)^s p_{ij}$$

Pour des variables aléatoires continues :

$$\mu_{k,s}[X^k Y^s] \triangleq \iint_{-\infty}^{+\infty} (x - m_x)^k (y - m_y)^s f(x, y) dx dy$$

**Définition 2.3.** On appelle covariance des variables aléatoires  $X, Y$ , le moment central mixte d'ordre deux  $\mu_{1,1}$  :

$$K_{xy} \triangleq M[\overset{\circ}{X} \overset{\circ}{Y}]$$

Pour des variables aléatoires discrètes :

$$K_{xy} \triangleq \sum_{i=1}^n \sum_{j=1}^m (x_i - m_x)(y_j - m_y) p_{ij}$$

Pour des variables aléatoires continues :

$$K_{xy} \triangleq \iint_{-\infty}^{+\infty} (x - m_x)(y - m_y) f(x, y) dx dy$$

### 3. CARACTÉRISTIQUES STATISTIQUES D'UNE RÉPARTITION

A toute caractéristique numérique<sup>3</sup> de la variable aléatoire  $X$  correspond une caractéristique statistique.

**Définition 3.1.** On appelle moyenne arithmétique ou moyenne statistique de la variable aléatoire discrète  $X$ , la moyenne de toutes les valeurs  $x_i$  observées de cette variable. On la note  $M^*[X]$  ou  $m_x^*$ .

$$M^*[X] \triangleq \frac{1}{n} \sum_{i=1}^n x_i$$

**Définition 3.2.** On appelle variance statistique de la variable aléatoire discrète  $X$ , la moyenne statistique des carrés des écarts des valeurs observées à la moyenne statistique. On la note  $V^*[X]$  ou  $V_x^*$  ou  $D^*[X]$  ou  $D_x^*$ .

$$V^*[X] \triangleq \frac{1}{n} \sum_{i=1}^n (x_i - m_x^*)^2$$

---

3. Voir Loi normale.pdf

**Définition 3.3.** On appelle moment statistique initial d'ordre  $s$  de la variable aléatoire discrète  $X$ , la fonction  $\alpha_s^*[X]$  définie par :

$$\alpha_s^*[X] \triangleq \frac{1}{n} \sum_{i=1}^n x_i^s$$

**Définition 3.4.** On appelle moment statistique centré d'ordre  $s$  de la variable aléatoire discrète  $X$ , la fonction  $\mu_s^*[X]$  définie par :

$$\mu_s^*[X] \triangleq \frac{1}{n} \sum_{i=1}^n (x_i - m_x^*)^s$$

**Définition 3.5.** On appelle covariance statistique des variables aléatoires discrètes  $X$  et  $Y$ , la fonction  $K_{xy}^*$  définie par :

$$K_{xy}^* \triangleq \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m (x_i - m_x^*)(y_j - m_y^*)$$

#### 4. APPLICATION

Une expérience a fourni un ensemble de valeurs  $(x_i, y_i)_{i=1, \dots, n}$ . Nous supposons que les erreurs de mesure suivent une loi normale, d'écart quadratique constant, et que la relation observée est linéaire. Nous cherchons les paramètres  $a$  et  $b$  de la droite

$$y = ax + b$$

qui représentent le mieux la relation expérimentale :

$$\begin{aligned} d \sum_{i=1}^n [y_i - \varphi(x_i)]^2 &= \sum_{i=1}^n d[y_i - \varphi(x_i)]^2 \\ &= \sum_{i=1}^n d[y_i^2 - 2y_i \varphi(x_i) + \varphi^2(x_i)] \\ &= \sum_{i=1}^n -2y_i d\varphi(x_i) + 2\varphi(x_i) d\varphi(x_i) \\ &= \sum_{i=1}^n 2[\varphi(x_i) - y_i] d\varphi(x_i) \\ &= \sum_{i=1}^n 2[\varphi(x_i) - y_i] [\partial_a \varphi(x_i) da + \partial_b \varphi(x_i) db] \\ &= 0 \end{aligned}$$

d'où le système

$$\begin{cases} \sum_{i=1}^n 2[\varphi(x_i) - y_i] \partial_a \varphi(x_i) da = 0 \\ \sum_{i=1}^n 2[\varphi(x_i) - y_i] \partial_b \varphi(x_i) db = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^n [(ax_i + b) - y_i] x_i = 0 \\ \sum_{i=1}^n [(ax_i + b) - y_i] = 0 \end{cases}$$

$$\begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i - \sum_{i=1}^n y_i x_i = 0 \\ a \sum_{i=1}^n x_i + nb - \sum_{i=1}^n y_i = 0 \end{cases}$$

En divisant par  $n$  et en utilisant les définitions du chapitre précédent,

$$\begin{cases} a\alpha_2^*[X] + bm_x^* - \alpha_{1,1}^*[X, Y] = 0 \\ am_x^* + b - m_y^* = 0 \end{cases}$$

$$\begin{cases} b = m_y^* - am_x^* \\ a\alpha_2^*[X] + (m_y^* - am_x^*)m_x^* - \alpha_{1,1}^*[X, Y] = 0 \end{cases}$$

$$\begin{cases} a \{ \alpha_2^*[X] - (m_x^*)^2 \} = \alpha_{1,1}^*[X, Y] - m_x^* m_y^* \\ b = m_y^* - am_x^* \end{cases}$$

$$\begin{cases} a = \frac{\alpha_{1,1}^*[X, Y] - m_x^* m_y^*}{\alpha_2^*[X] - (m_x^*)^2} \\ b = m_y^* - am_x^* \end{cases}$$

d'où, en utilisant les moments centrés :

$$\begin{cases} a = \frac{K_{xy}^*}{V_x^*} \\ b = m_y^* - am_x^* \end{cases}$$

*E-mail address:* o.castera@free.fr

*URL:* <http://o.castera.free.fr/>