APV - Résumé Olivier D'Ancona

Espérance

Soit X une variable aléatoire à valeurs dans $\mathbb{R}$, on définit l'espérance de X par :

Cas discret : $\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i \cdot p_i$

Cas continu : $\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$ où $f(x)$ est la densité de probabilité de X.

Propriétés

- $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- $\mathbb{E}[aX] = a\mathbb{E}[X]$
- $\mathbb{E}[X^2] = \mathbb{E}[X]^2 + \mathbb{V}[X]$
- $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$ if X and Y are independent.
- $\mathbb{E}[c] = c$ if c is a constant.

Variance

Soit X une variable aléatoire à valeurs dans $\mathbb{R}$.Cas discret :

$$\mathbb{V}[X] = \sum_{i=1}^{\infty} (x_i - \mathbb{E}[X])^2 \cdot p_i$$
$$= \mathbb{E}[(X_i - \mu)^2] = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2$$

Cas continu : Cas discret :

$$\mathbb{V}[X] = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 \cdot f(x)\, dx$$

Lien entre écart-type et variance : $\mathbb{V}[X] = \sigma^2$

Propriétés

- $\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
- $\mathbb{V}[X + Y] = \mathbb{V}[X] + \mathbb{V}[Y]$ if X and Y are independent.
- $\mathbb{V}[aX] = a^2\mathbb{V}[X]$
- $\mathbb{V}[XY] = \mathbb{V}[X] \cdot \mathbb{V}[Y] + \mathbb{V}[X] \cdot \mathbb{E}[Y]^2 + \mathbb{V}[Y] \cdot \mathbb{E}[X]^2$ if X and Y are independent.
- $\mathbb{V}[c] = 0$ if c is a constant.

Loi de probabilité

Loi	Param	$\mathbb{E}$	$\mathbb{V}$
Bernoulli	p	p	$p(1-p)$
Binomiale	n, p	np	$np(1-p)$
Uniforme	a, b	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normale	μ, σ	μ	σ^2
Exponentielle	λ	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Poisson	λ	λ	λ
Géométrique	p	$\frac{1}{p}$	$\frac{1}{p^2}$

Likelihood

Soit X une variable aléatoire à valeurs dans $\mathbb{R}$, on définit la likelihood de θ par :

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

où $f(x)$ est la densité de probabilité de X. Souvent, on utilise le logarithme de la likelihood :

$$\log L(\theta) = \sum_{i=1}^{n} \log f(x_i|\theta)$$

MLE

$\hat{\theta}_{MLE}$ est de θ est le paramètre qui maximise la likelihood de θ.
Soit θ un paramètre d'une loi de probabilité, on définit le MLE de θ par :

$$\hat{\theta} = \arg\max_{\theta} L(\theta)$$

où $L(\theta)$ est la likelihood de θ.

Procédure

1. Définir la likelihood de θ $L(\theta)$.
2. On passe au logarithme de la likelihood. $l(\theta) = \log L(\theta)$
3. On dérive la log-likelihood par rapport à θ
4. On cherche $\hat{\theta}$ tel que $\frac{\partial}{\partial\theta} l(\hat{\theta}) = 0$.
5. On vérifie que $\frac{\partial^2}{\partial\theta^2} l(\hat{\theta}) < 0$

Régression Linéaire

Soit un tableau de données :
$x = \text{Soap(g)}$, $y = \text{Height(cm)}$, $x \cdot y$, x^2

$$X = [1, Soap]$$

$$X^T X = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} = \begin{bmatrix} 7 & 38.5 \\ 38.5 & 218.95 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} = \begin{bmatrix} 348 \\ 1975 \end{bmatrix}$$

$$\hat{\theta} = (X^T X)^{-1} X^T y = \begin{bmatrix} -2.67 \\ 9.51 \end{bmatrix}$$

Inverse d'une matrice 2x2 :

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Normale

1. $P(X \leq x) = \Phi(x) = \int_{-\infty}^{x} N(0,1)dx$
2. $\Phi(x) = 1 - \Phi(-x)$
3. $F_x(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ est une loi normale centrée réduite