

01 Statistique Descriptive

mardi 6 septembre 2022

1. Cours sur [Internet](#)

2. Introduction

3. Graphique : une variable

-1- Tableau de fréquence

-2- Diagramme de bandes

-3- Différentes représentations

1- Quantitative ordinale

1> La fréquence relative cumulée

$$F_k = \sum_{j=1}^k f_j - f_k = \frac{n_k}{n}$$

2- Histogramme 直方图 : pour la variable continue

1> Obtient K intervalles $[a_k, a_{k+1}]$, $k = 1, \dots, K$

$$2> \sum_{i=1}^K h_i(\alpha_i - \alpha_{i-1}) = 1, K = 1 + \frac{10}{3} \log_{10} n$$

3- Diagramme en tigre et feuilles

-4- Fonction de répartition empirique 经验分布函数 (阶跃函数)

$$1- F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i < x) - I_A(y) = \begin{cases} 1, y \in A \\ 0, y \notin A \end{cases}$$

$$2- \|F_n - F\|_{\infty} \rightarrow 0$$

4. Graphique : deux variable

-1- Table de contingence 列联表

1- Application aux probabilités conditionnelles

-2- Diagramme en bandes superposées

5. Résumés numérique

-1- Indicateurs de tendance centrale

1- La Moyenne Empirique \bar{x}

1> Inconvénient : sensible à la valeurs « aberrantes ». 异常值敏感

2> La Moyenne tronquée d'ordre k (plus robust) : supprimer les \square plus petites et les \square plus grandes valeurs

2- La Médiane $M = x_{\lfloor \frac{n}{2} \rfloor}$

1> Premier Fractile: $\hat{f}_{0.25}$ 分位数

2> 2 Fractile: $\hat{f}_{0.5}$

3> 3 Fractile: $\hat{f}_{0.75}$

4> Fractile empirique d'ordre α : $\hat{f}_\alpha = \hat{F}^{-1}(\alpha) = x([n\alpha])$

3- Le Mode : le plus grand effectif ou la plus grande fréquence

-2- Indicateurs dispersion

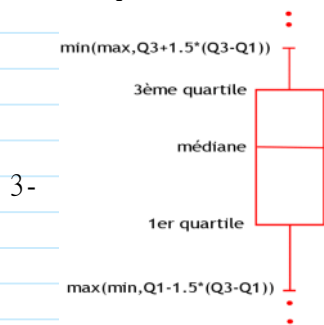
1- Variance empirique $s^2 = \frac{1}{n} \sum (x_i - \bar{x})^2 = \frac{1}{n} \sum x_i^2 - \bar{x}^2$ (Etat-type empirique : s)

2- Variance empirique corrigée $s^{*2} = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{n}{n-1} s^2$

-3- Diagramme en boîte/boîte à moustaches

1- Les moustaches sont des segments à une distance inférieure à 1.5H

2- Les points des extrémités des moustaches sont représentés individuellement.



-4- Coefficients de corrélation

1-
$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n \cdot s_x s_y} = \frac{\sum_{i=1}^n x_i y_i - \bar{x} \bar{y}}{n \cdot s_x s_y} \in [-1, 1]$$

2- Coefficient de corrélation de Spearman : il mesure le degré de dépendance monotone entre deux variables

1> Soit le rang de, le rang de

$$r_s = \frac{\left(\sum_{i=1}^n r_i s_i - \frac{(n+1)^2}{4} \right)}{\frac{n(n^2 - 1)}{12}} = 1 - \frac{6 \sum_{i=1}^n (r_i - s_i)^2}{n(n^2 - 1)}$$

-5- Indicateurs quantitatifs