Descriptive statistics

Enrique Calderon

March 2024

---

1 Dispersion measures

1.1 Range

$$\text{Range } = x_{\max} - x_{\min}$$

1.2 Variance

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \hat{x})}{n-1}$$

1.3 Standard deviation

$$s = \sqrt{s^2}$$

1.4 Variation coefficient

$$\text{C.V. } = \frac{s}{\hat{x}}$$

---

2 Frequency tables

2.1 Number of classes, Sturges's rule

$$k = 1 + log_2(n)$$

2.2 Width of classes

$$W = \frac{\text{Range}}{k} = \frac{|x_{\max - x_{\min}}}{n}$$

---

3 Normal distribution

3.1 Empiric rule

Value then	Population quantity
$\hat{x} \pm 1 \cdot s$	68%
$\hat{x} \pm 2 \cdot s$	95%
$\hat{x} \pm 3 \cdot s$	99.9%

3.2 Percentiles

$$Q_1 = \frac{n+1}{4}$$

$$Q_2 = \frac{2(n+1)}{4}$$

$$Q_3 = \frac{3(n+1)}{4}$$

3.3 Boxplot

$$\text{RIC} = Q_3 - Q_1$$

$$\text{L.I.} = Q_1 - 1.5 \cdot \text{RIC}$$

$$\text{L.S.} = Q_3 + 1.5 \cdot \text{RIC}$$

3.4 Bias

$$\alpha_3 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x})^3}{s^3}$$

3.5 Kurtosis

$$\alpha_4 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x})^4}{s^4}$$

3.6 Equation

From $-\infty$ to ∞:

$$f(x) = \frac{1}{\sigma^2\sqrt{2\pi}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma^2}\right)^2}}$$

---

4 Distributions

4.1 Z

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{\hat{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

4.1.1 Standard error

$$s_{\hat{x}} = \frac{s}{\sqrt{n}}$$

4.1.2 Table

- $z \leq$: From table
- $z \geq$: $1-$ from table