

Relazione Esercizio sulla Statistica Descrittiva

- Esercitazione 2 -

Andrea Lops
Paolo Rotolo
Laura Loperfido
Teresa Pantone

16/12/2018

1 Esercizio:

Utilizzando i metodi dell'*analisi statistica descrittiva* si vogliono ottenere informazioni utili sulle proprietà di alcuni dati (6 casi differenti) estratti dall'analisi di un processo.

1.1 Azioni preliminari:

Prima di andare ad analizzare i dati effettuiamo il parse dei .txt grezzi forniti dal Prof. Attivissimo e li andiamo ad ordinare.

2 Caso 1

Ci viene fornito un set di dati che afferiscono dalle misure delle resistenze di superleghe per razzi.

2.1 Istogramma (Figure 1)

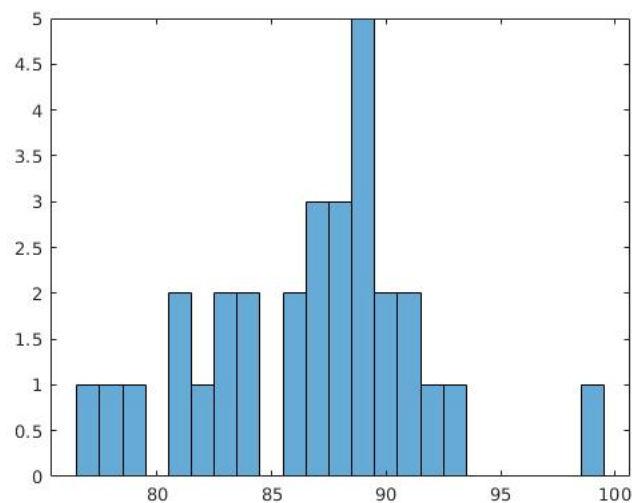


Figure 1: Istogramma caso 1

2.2 Mediana e dispersione

Il valore centrale dell'array è la mediana:

$$\text{mediana} = 87.5$$

Dispersione :

$$d = M_{ax} - M_{in} = 22 \quad (1)$$

2.3 Primo, Secondo e Terzo Quartile

$$30 * 0.25 = 8 \quad Q_1 = x_8 = 83 \quad (2)$$

$$30 * 0.50 = 15 \quad Q_2 = (x_{15} + x_{16})/2 = 87.5 \quad (3)$$

$$30 * 0.75 = 23 \quad Q_3 = x_{23} = 89 \quad (4)$$

2.4 Interquartile e Outliers

Interquartile: è l'ampiezza della fascia centrale (compresa tra primo e terzo quartile) dei valori analizzati:

$$Q_3 - Q_1 = 6 \quad (5)$$

Outliers: sono i valori che si discostano dagli altri valori osservati

$$Q_3 + (1.5 * \text{Interquartile}) = 98 \quad (6)$$

2.5 Boxplot (Figure 2)

Metodo grafico per rappresentare la distribuzione dei valori usando gli Outliers

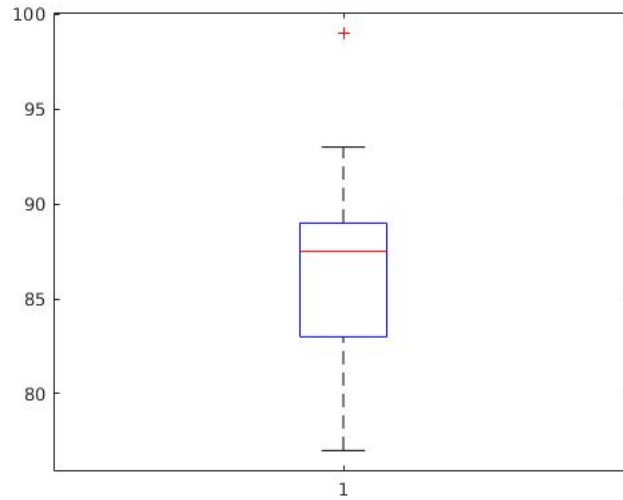


Figure 2: Boxplot caso 1

2.6 Media e Varianza

Media:

$$M_a = \frac{1}{n} \sum_{i=1}^n x_i \quad (7)$$

Varianza:

$$\sigma_X^2 = \mathbb{E}[(X - \mathbb{E}[X])^2] \quad (8)$$

2.7 Diagrama quantile-quantile (Figure 3)

Metodo grafico per rappresentare la distribuzione dei valori usando gli Outliers

2.8 Verifica del tipo di distribuzione col *Chi Quadro test*

La funzione *Chi Quadro test* ci restituisce un valore utile per valutare la natura della curva. Infatti se il suo valore supera 0.5 essa sicuramente non sarà Gaussiana. In questo caso specifico la curva può essere Gaussiana.

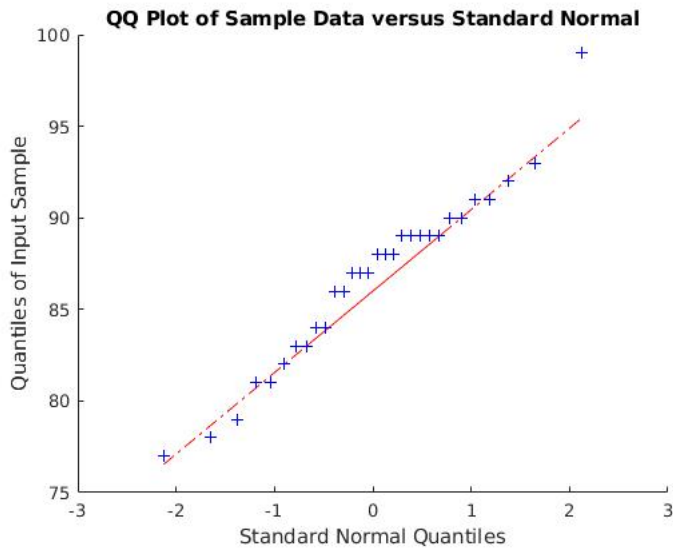


Figure 3: QQPlot caso 1

3 Caso 2

Ci viene fornito un set di dati in numero pari che afferiscono all'ambito medico, nello specifico alla pressione sistolica di 15 pazienti donna di età 20-22.

3.1 Istogramma (Figure 4)

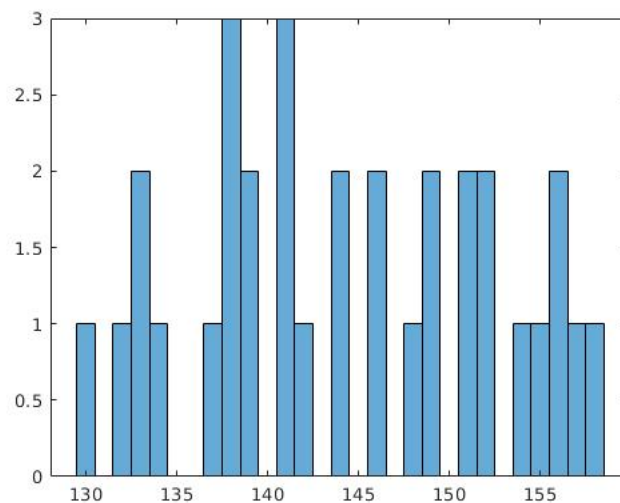


Figure 4: Istogramma caso 2

3.2 Mediana e dispersione

mediana = 144

dispersione = 28

3.3 Primo, Secondo e Terzo Quartile

Primo Quartile:

$$Q_1 = 138$$

Secondo Quartile:

$$Q_2 = 144$$

Terzo Quartile:

$$Q_3 = 152$$

3.4 Interquartile e Outliers

$$\text{interquartile} = 14$$

$$\text{outliers} = 173$$

3.5 Boxplot (Figure 5)

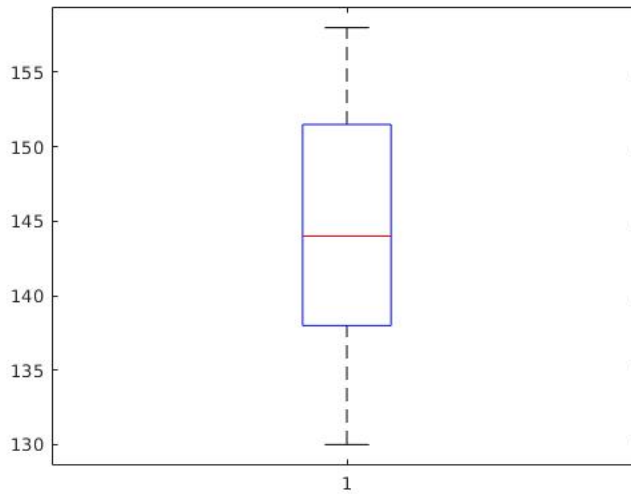


Figure 5: Boxplot caso 2

3.6 Media e Varianza

Media:

$$M_a = 144.5 \quad (9)$$

Varianza:

$$\sigma_X^2 = 67.29 \quad (10)$$

3.7 Diagrama quantile-quantile (Figure 6)

3.8 Verifica del tipo di distribuzione col *Chi Quadro test*

La curva può essere Gaussiana.

4 Caso 3

Come nel caso 2, abbiamo una serie di dati inerenti alla pressione sistolica, ma dispari. In più il caso 3 si differenzia dal 2 solo per un dato, infatti dai calcoli si può notare come statisticamente parlando non discordano molto tra loro.

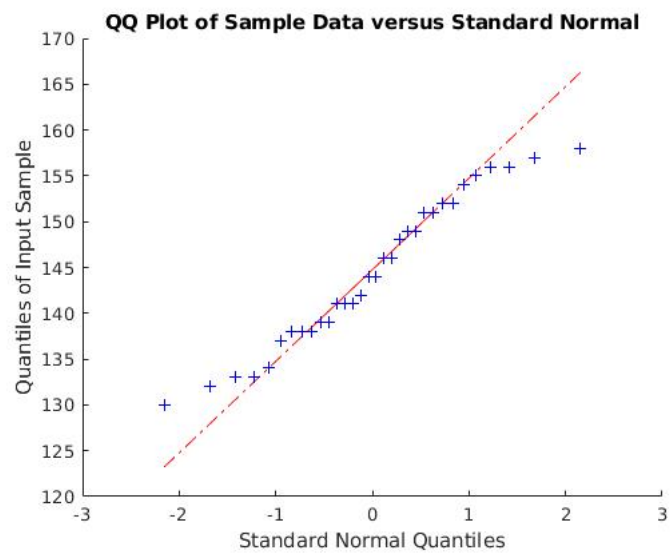


Figure 6: QQPlot caso 2

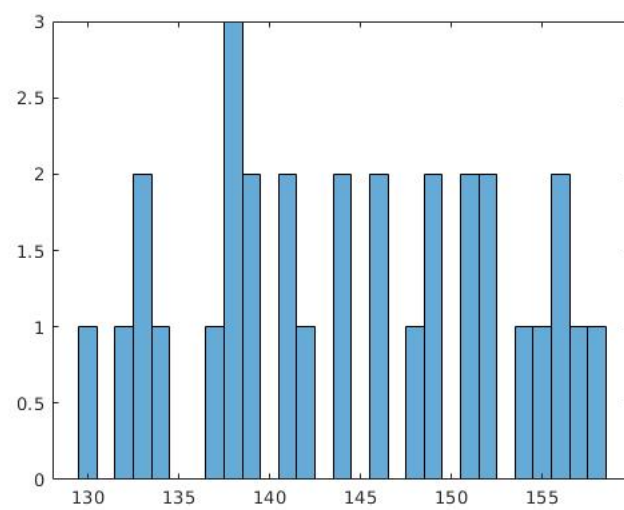


Figure 7: Istogramma caso 3

4.1 Istogramma (Figure 7)

4.2 Mediana e dispersione

La mediana ha l'influenza del caso dispari, infatti prende l'elemento mediano che è singolo

$$\text{mediana} = 144$$

$$\text{dispersione} = 28$$

4.3 Primo, Secondo e Terzo Quartile

Primo Quartile:

$$Q_1 = 138$$

Secondo Quartile:

$$Q_2 = 144$$

Terzo Quartile:

$$Q_3 = 152$$

4.4 Interquartile e Outliers

$$\text{interquartile} = 14$$

$$\text{outliers} = 173$$

4.5 Boxplot (Figure 8)

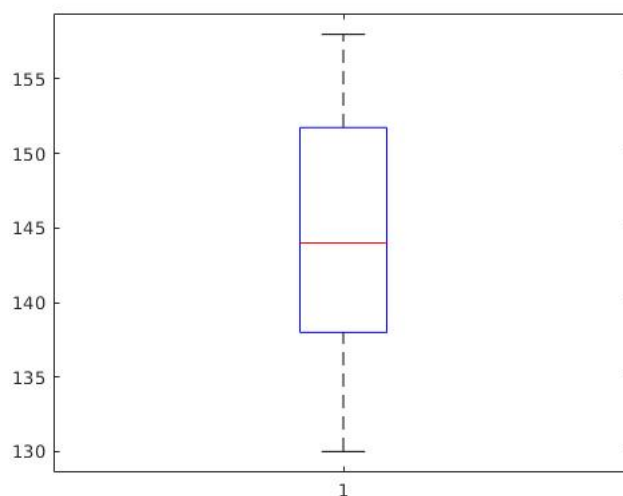


Figure 8: Boxplot caso 3

4.6 Media e Varianza

Media:

$$M_a = 144.613 \quad (11)$$

Varianza:

$$\sigma_X^2 = 69.119 \quad (12)$$

4.7 Diagrama quantile-quantile (Figure 9)

4.8 Verifica del tipo di distribuzione col *Chi Quadro test*

La curva può essere Gaussiana.

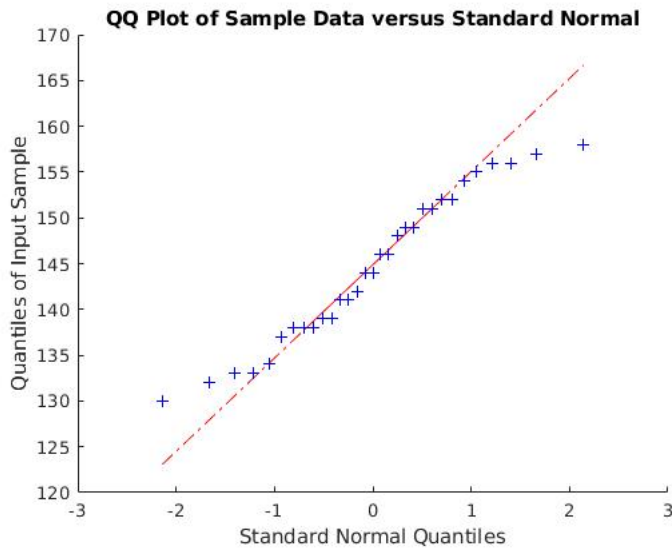


Figure 9: QQPlot caso 3

5 Caso 4

In questo caso abbiamo 100 valori provenienti da delle analisi della resistenza alla trazione di cilindri di cemento.

5.1 Istogramma (Figure 10)

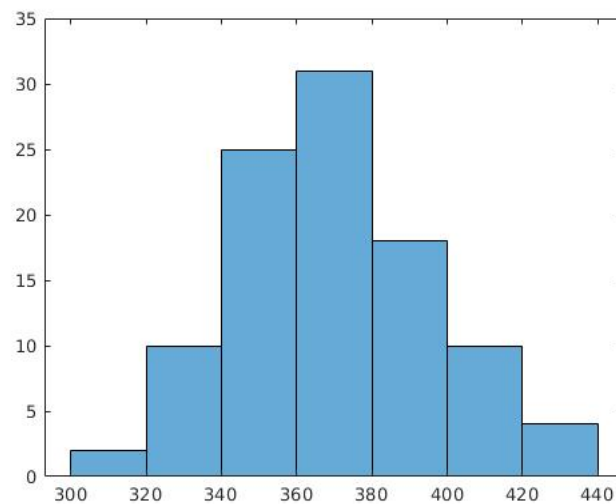


Figure 10: Istogramma caso 4

5.2 Mediana e dispersione

La mediana ha l'influenza del caso dispari, infatti prende l'elemento mediano che è singolo

$$\text{mediana} = 360$$

$$\text{dispersione} = 140$$

5.3 Primo, Secondo e Terzo Quartile

Primo Quartile:

$$Q_1 = 340$$

Secondo Quartile:

$$Q_2 = 360$$

Terzo Quartile:

$$Q_3 = 390$$

5.4 Interquartile e Outliers

interquartile = 50

outliers = 465

5.5 Boxplot (Figure 11)

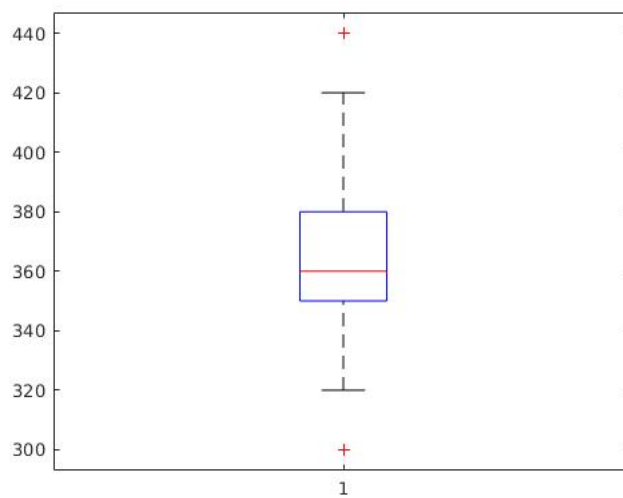


Figure 11: Boxplot caso 4

5.6 Media e Varianza

Media:

$$M_a = 364.7 \quad (13)$$

Varianza:

$$\sigma_X^2 = 720.111 \quad (14)$$

5.7 Diagrama quantile-quantile (Figure 12)

5.8 Verifica del tipo di distribuzione col *Chi Quadro test*

La curva non può essere Gaussiana.

6 Caso 5

In questo caso abbiamo 5000 valori generati casualmente con distribuzione Gaussiana. A differenza dei casi precedenti è inutile riportare qualsivoglia dato, visto che sono generati casualmente, però l'unica analisi sicuramente giusta per tutti i "Casi 5" è che il Chi Quadro test restituirà sempre che la curva può essere Gaussiana.

6.1 Verifica del tipo di distribuzione col *Chi Quadro test*

La curva può essere Gaussiana.

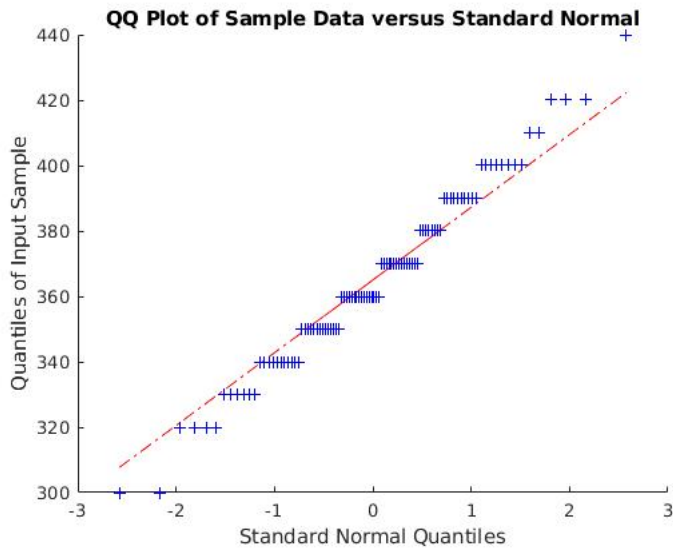


Figure 12: QQPlot caso 4

7 Caso 6

Come per il caso 5 abbiamo 5000 valori generati casualmente, ma con distribuzione Uniforme. Questo significa che sicuramente il *Chi Quadro test* restituirà che la curva non può essere Gaussiana.

7.1 Verifica del tipo di distribuzione col *Chi Quadro test*

La curva non può essere Gaussiana.