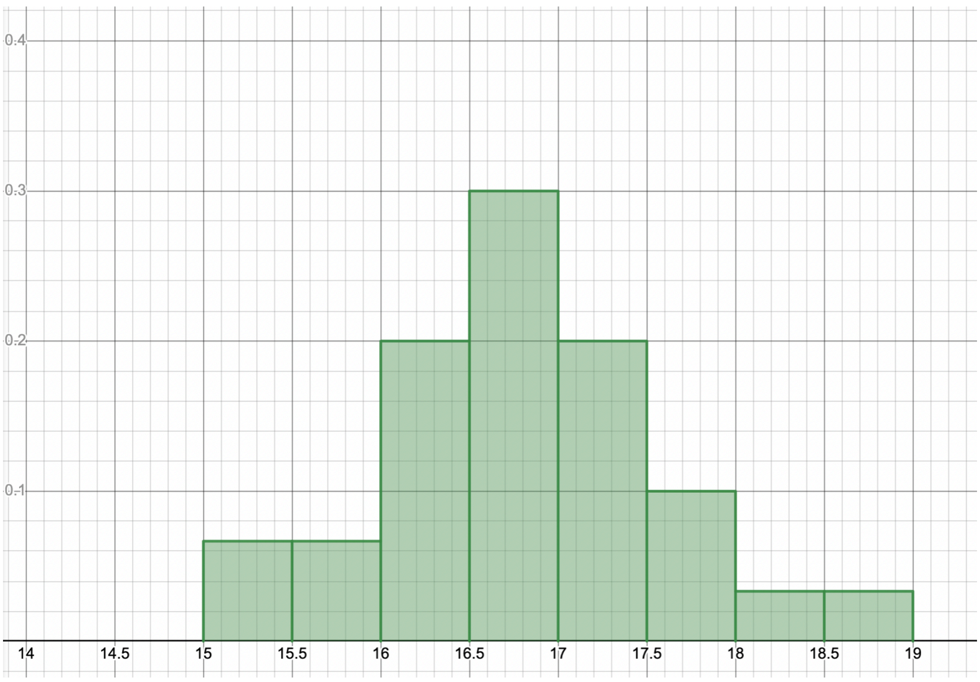
Esercizio 3 - Soluzione

1. La tabella delle frequenze assolute e relative è la seguente

Classe	Frequenza	Frequenza relativa
15,0 <= x < 15,5	2	0,07
15,5 <= x < 16,0	2	0,07
16,0 <= x < 16,5	6	0,20
16,5 <= x < 17,0	9	0,30
17,0 <= x < 17,5	6	0,20
17,5 <= x < 18,0	3	0,10
18,0 <= x < 18,5	1	0,03
18,5 <= x < 19,0	1	0,03

2. L'istogramma delle frequenze relative è il seguente, dove in ascissa va riportata l'altezza (in centimetri), mentre in ordinata la frequenza relativa (adimensionale)



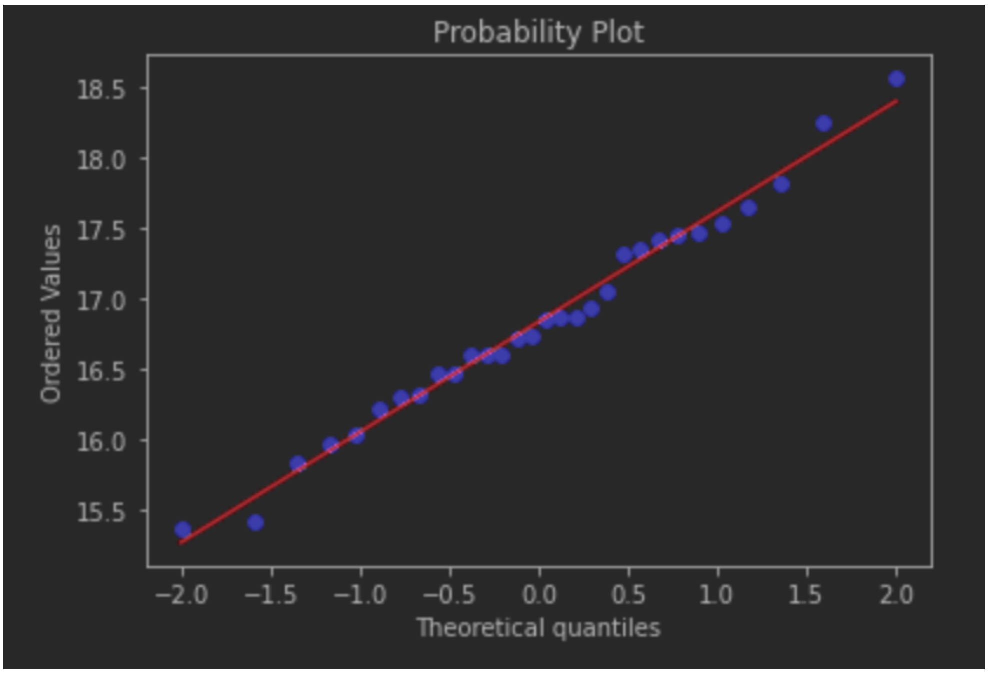
3. Un buon stimatore per l'attesa della distribuzione è la media campionaria

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 16.84\,\text{cm}.$$

Mentre un buon stimatore per la varianza è la varianza campionaria

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2 = 0.58\,\text{cm}^2.$$

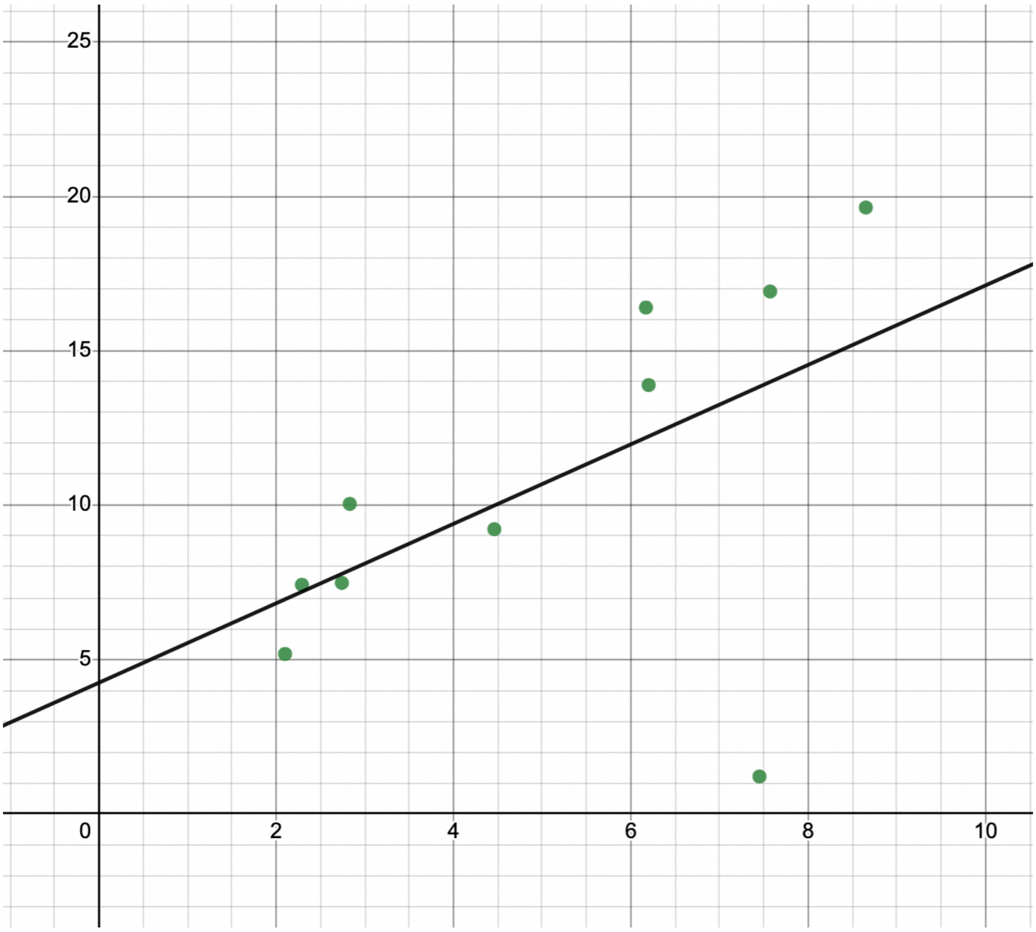
4. Il normal plot dei dati è il seguente



e da esso si evince che un'ipotesi di normalità dei dati è ragionevole, in quanto il plot dei quantili dei dati versus i quantili Gaussiani si ben dispone lungo una retta.

Esercizio 4- Soluzione

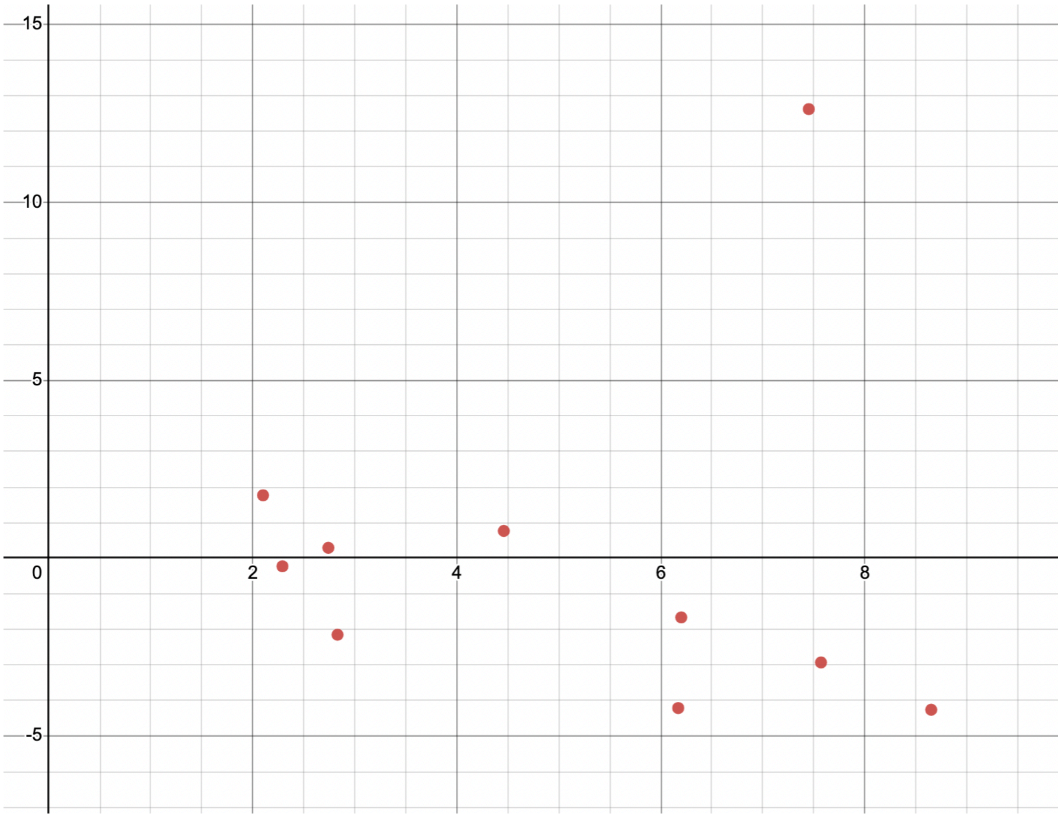
1. Il grafico di dispersione e la retta di regressione sono rappresentati di seguito



E la retta di regressione è (non mi aspetto che scrivano tutte le unità di misura, anche se è un qualcosa di desiderabile)

$$y = 4.25\,\text{cm} + \left(1.29\frac{\text{cm}}{\text{h}}\right)x$$

2. Dal grafico dei residui si evince che il modello lineare **non** è un buon modello per rappresentare i dati, poiché i residui non sono disposti in modo casuale attorno allo zero, bensì esibiscono un trend (crescente o decrescente a seconda del segno con cui è stato definito il residuo).



3. Il punto con residuo maggiore è $(7.45, 1.21)$ e il valore del residuo è, in valore assoluto, 12.6. Se si rimuove il suddetto punto dai dati, i residui di distribuiscono in modo casuale attorno allo zero, ed il modello lineare diviene ragionevole per spiegare i dati.