

24-06-2019

2022-10-27

1)

Età	No. laureati Ingegneria El.	No. laureati Lettere
23 ┤ 25	9	282
25 ┤ 27	405	1432
27 ┤ 29	710	1065
29 ┤ 31	252	285
31 ┤ 40	237	591
Totale	1613	3655

Y = Età degli studenti laureati in Ingegneria Elettronica e Lettere

INGEGNERIA ELETTRONICA

```
# frequenze assolute
(Ing_El_Table = data.frame(
  ETA_MIN = c(23,25,27,29,31),
  ETA_MAX = c(25,27,29,31,40),
  fi = c(9,405,710,252,237)
) )

##   ETA_MIN ETA_MAX  fi
## 1     23     25    9
## 2     25     27 405
## 3     27     29 710
## 4     29     31 252
## 5     31     40 237

fCumulate = cumsum(Ing_El_Table$fi)

(Ing_El_Table = cbind( Ing_El_Table,Fi=fCumulate))

##   ETA_MIN ETA_MAX  fi  Fi
## 1     23     25    9    9
## 2     25     27 405  414
## 3     27     29 710 1124
## 4     29     31 252 1376
## 5     31     40 237 1613
```

```
# aggiunge la colonna frequenze relative pi
(Ing_El_Table = Ing_El_Table %>% mutate(pi = fi/sum(fi)))
```

```
##   ETA_MIN ETA_MAX  fi   Fi      pi
## 1      23      25    9    9 0.005579665
## 2      25      27  405  414 0.251084935
## 3      27      29  710 1124 0.440173590
## 4      29      31  252 1376 0.156230626
## 5      31      40  237 1613 0.146931184
```

```
(fRelCum = cumsum(Ing_El_Table$pi))
```

```
## [1] 0.005579665 0.256664600 0.696838190 0.853068816 1.000000000
```

```
# TABELLA DI FREQUENZE ASSOLUTE E RELATIVE, CUMULATE E NON
(Ing_El_Table = cbind(Ing_El_Table,Pi = fRelCum))
```

```
##   ETA_MIN ETA_MAX  fi   Fi      pi      Pi
## 1      23      25    9    9 0.005579665 0.005579665
## 2      25      27  405  414 0.251084935 0.256664600
## 3      27      29  710 1124 0.440173590 0.696838190
## 4      29      31  252 1376 0.156230626 0.853068816
## 5      31      40  237 1613 0.146931184 1.000000000
```

CLASSE MEDIANA

$$y_{0.5} = y_i, i = 0.5 * 100 * n$$

```
(ING_EL_Classe_Mediana = Ing_El_Table[Ing_El_Table$Pi >= 0.5,c("ETA_MIN","ETA_MAX")][1,])
```

```
##   ETA_MIN ETA_MAX
## 3      27      29
```

```
(Mediana_ING_EL = (ING_EL_Classe_Mediana$ETA_MAX+ING_EL_Classe_Mediana$ETA_MIN)/2)
```

```
## [1] 28
```

MEDIA ARITMETICA

$$E(Y) = \sum_{i=1}^J y_i^c * p_i$$

```
# Valori Centrali della classe
```

```
(Yc = (Ing_El_Table$ETA_MIN + Ing_El_Table$ETA_MAX)/2)
```

```
## [1] 24.0 26.0 28.0 30.0 35.5
```

```
(Ing_El_Table = cbind(Yci = Yc,Ing_El_Table))
```

```
##   Yci ETA_MIN ETA_MAX  fi   Fi      pi      Pi
## 1 24.0      23      25    9    9 0.005579665 0.005579665
## 2 26.0      25      27  405  414 0.251084935 0.256664600
## 3 28.0      27      29  710 1124 0.440173590 0.696838190
## 4 30.0      29      31  252 1376 0.156230626 0.853068816
## 5 35.5      31      40  237 1613 0.146931184 1.000000000
```

```
(Media_ING_EL = sum(Yc*Ing_El_Table$pi))
```

```
## [1] 28.88996
```

VARIANZA

$$V(Y) = E[(Y - E(Y))^2] = E(Y^2) - (E(Y))^2$$

$$V(Y) = \sum_{i=1}^J (y_i^c - E(Y))^2 * p_i$$

```
(Varianza_ING_EL = sum(((Ing_El_Table$Yci-Media_ING_EL)**2)*Ing_El_Table$pi))
```

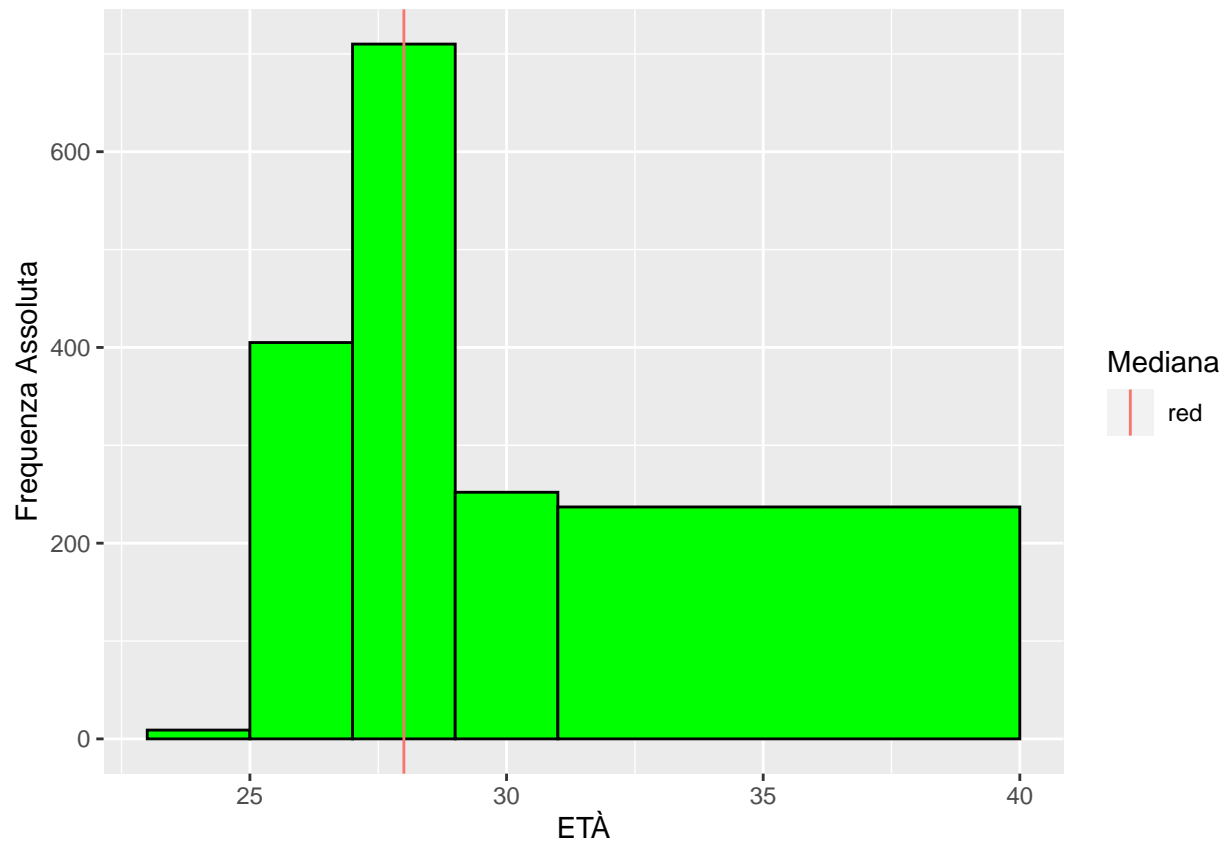
```
## [1] 9.191393
```

ISTOGRAMMA

Ing_El_Table

##	Yci	ETA_MIN	ETA_MAX	fi	Fi	pi	Pi
## 1	24.0	23	25	9	9	0.005579665	0.005579665
## 2	26.0	25	27	405	414	0.251084935	0.256664600
## 3	28.0	27	29	710	1124	0.440173590	0.696838190
## 4	30.0	29	31	252	1376	0.156230626	0.853068816
## 5	35.5	31	40	237	1613	0.146931184	1.000000000

```
ggplot(Ing_El_Table,aes(x=Yci,y=fi)) + geom_col( color="black", fill="green",width = Ing_El_Table$ETA_L
```



SIMMETRIA

$$\gamma_Y = \frac{E[(Y - E(Y))^3]}{\sigma_Y^3}$$

```
(sigma_ING_EL = Varianza_ING_EL**0.5) #radice quadrata della varianza
```

```
## [1] 3.031731
```

```
(gamma_ING_EL = mean((Ing_El_Table$Yci-Media_ING_EL)**3)/(sigma_ING_EL**3))
```

```
## [1] 1.065174
```

Avendo un valore POSITIVO c'è asimmetria positiva, quindi a destra della mediana, come si evince dall'istogramma

CURTOSI

$$\beta_Y = \frac{E[(Y - E(Y))^4]}{\sigma_Y^4}$$

```
(curtosi_ING_EL = mean((Ing_El_Table$Yci-Media_ING_EL)**4)/(sigma_ING_EL**4))
```

```
## [1] 6.043242
```

Avendo un valore > 3 vi è una situazione LEPTOCURTICA, quindi le frequenze possiedono un valore elevato ai margini del Range di Y.

Ciò conferma anche la situazione di ASIMMETRIA POSITIVA

LETTERE

```
# frequenze assolute
(Lettere_Table = data.frame(
  ETA_MIN = c(23,25,27,29,31),
  ETA_MAX = c(25,27,29,31,40),
  fi = c(282,1432,1065,285,591)
) )
```

```
##   ETA_MIN ETA_MAX   fi
## 1      23      25  282
## 2      25      27 1432
## 3      27      29 1065
## 4      29      31  285
## 5      31      40  591
```

```
(fCumulate = cumsum(Lettere_Table$fi))
```

```
## [1] 282 1714 2779 3064 3655
```

```
(Lettere_Table = cbind( Lettere_Table,Fi=fCumulate))
```

```
##   ETA_MIN ETA_MAX   fi   Fi
## 1      23      25  282  282
## 2      25      27 1432 1714
## 3      27      29 1065 2779
## 4      29      31  285 3064
## 5      31      40  591 3655
```

```
# aggiunge la colonna frequenze relative pi
(Lettere_Table = Lettere_Table %>% mutate(pi = fi/sum(fi)))
```

```
##   ETA_MIN ETA_MAX   fi   Fi      pi
## 1      23      25  282  282 0.07715458
## 2      25      27 1432 1714 0.39179207
```

```
## 3      27      29 1065 2779 0.29138167
## 4      29      31  285 3064 0.07797538
## 5      31      40  591 3655 0.16169631
```

```
(fRelCum = cumsum(Lettere_Table$pi))
```

```
## [1] 0.07715458 0.46894665 0.76032832 0.83830369 1.00000000
```

```
# TABELLA DI FREQUENZE ASSOLUTE E RELATIVE, CUMULATE E NON
(Lettere_Table = cbind(Lettere_Table, Pi = fRelCum))
```

```
##   ETA_MIN ETA_MAX   fi   Fi       pi       Pi
## 1      23      25  282  282 0.07715458 0.07715458
## 2      25      27 1432 1714 0.39179207 0.46894665
## 3      27      29 1065 2779 0.29138167 0.76032832
## 4      29      31  285 3064 0.07797538 0.83830369
## 5      31      40  591 3655 0.16169631 1.00000000
```

CLASSE MEDIANA

$$y_{0.5} = y_i, i = 0.5 * 100 * n$$

```
(ING_EL_Classe_Mediana = Lettere_Table[Lettere_Table$Pi >= 0.5, c("ETA_MIN", "ETA_MAX")][1,])
```

```
##   ETA_MIN ETA_MAX
## 3      27      29
```

```
(Mediana_ING_EL = (ING_EL_Classe_Mediana$ETA_MAX + ING_EL_Classe_Mediana$ETA_MIN)/2)
```

```
## [1] 28
```

MEDIA ARITMETICA

$$E(Y) = \sum_{i=1}^J y_i^c * p_i$$

```
# Valori Centrali della classe
```

```
(Yc = (Lettere_Table$ETA_MIN + Lettere_Table$ETA_MAX)/2)
```

```
## [1] 24.0 26.0 28.0 30.0 35.5
```

```
(Lettere_Table = cbind(Yci = Yc, Lettere_Table))
```

```
##   Yci ETA_MIN ETA_MAX   fi   Fi       pi       Pi
## 1 24.0      23      25  282  282 0.07715458 0.07715458
## 2 26.0      25      27 1432 1714 0.39179207 0.46894665
## 3 28.0      27      29 1065 2779 0.29138167 0.76032832
## 4 30.0      29      31  285 3064 0.07797538 0.83830369
## 5 35.5      31      40  591 3655 0.16169631 1.00000000
```

```
(Media_LET = sum(Yc*Lettere_Table$pi))
```

```
## [1] 28.27647
```

VARIANZA

$$V(Y) = E[(Y - E(Y))^2] = E(Y^2) - (E(Y))^2$$

$$V(Y) = \sum_{i=1}^J (y_i^c - E(Y))^2 * p_i$$

```
(Varianza_LET = sum(((Lettere_Table$Yci-Media_LET)**2)*Lettere_Table$pi))
```

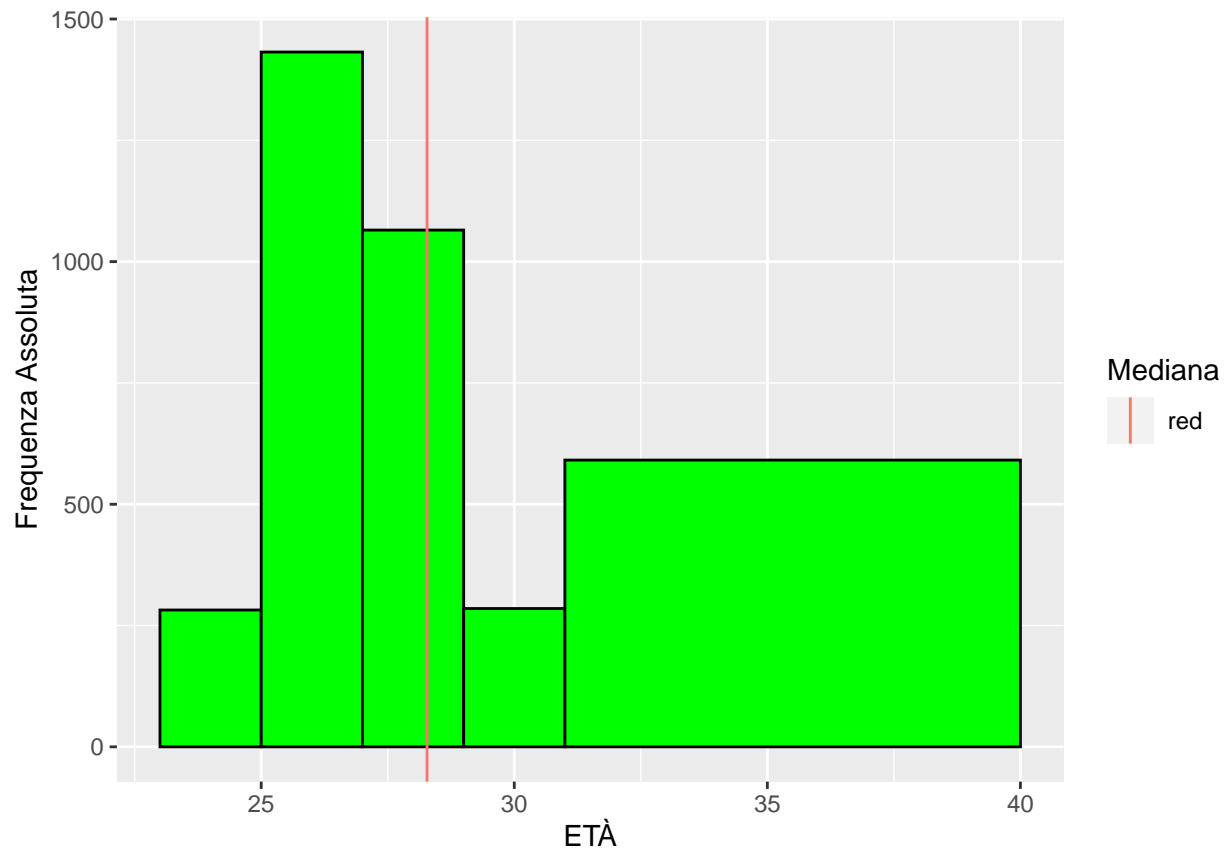
```
## [1] 12.13252
```

ISTOGRAMMA

```
Lettere_Table
```

##	Yci	ETA_MIN	ETA_MAX	fi	Fi	pi	Pi
## 1	24.0	23	25	282	282	0.07715458	0.07715458
## 2	26.0	25	27	1432	1714	0.39179207	0.46894665
## 3	28.0	27	29	1065	2779	0.29138167	0.76032832
## 4	30.0	29	31	285	3064	0.07797538	0.83830369
## 5	35.5	31	40	591	3655	0.16169631	1.00000000

```
ggplot(Lettere_Table,aes(x=Yci,y=fi)) + geom_col( color="black", fill="green",width = Lettere_Table$ET
```



SIMMETRIA

$$\gamma_Y = \frac{E[(Y - E(Y))^3]}{\sigma_Y^3}$$

```
(sigma_LET = Varianza_LET**0.5) #radice quadrata della varianza
```

```
## [1] 3.483177
```

```
(gamma_LET = mean((Lettere_Table$Yci-Media_LET)**3)/(sigma_LET**3))
```

```
## [1] 1.381985
```

Avendo un valore POSITIVO c'è asimmetria positiva, quindi a destra della mediana, come si evince dall'istogramma

CURTOSI

$$\beta_Y = \frac{E[(Y - E(Y))^4]}{\sigma_Y^4}$$

```
(curtosi_LET = mean((Lettere_Table$Yci-Media_LET)**4)/(sigma_LET**4))
```

```
## [1] 4.20227
```

Avendo un valore > 3 vi è una situazione LEPTOCURTICA, quindi le frequenze possiedono un valore elevato ai margini del Range di Y.

Ciò conferma anche la situazione di ASIMMETRIA POSITIVA