Corsi di Laurea in SES, SEFA e SG

Statistica di Base (a.a. 2019/2020)

Valeria Sambucini

Appunti Integrativi - PROPRIETA' della MEDIA ARITMETICA

Data la distribuzione unitaria $x_1, ..., x_i, ..., x_n$ si definisce la media aritmetica come

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Le principali proprietà della media aritmetica sono:

1. **Internalità.** La media aritmetica è sempre compresa tra il valore minimo e il valore massimo delle modalità assunte dal carattere.

   Dimostrazione. Si considerino le modalità osservate ordinate in senso non decrescente

   $$x_{(1)} \leq x_{(2)} \leq ... \leq x_{(i)} \leq ... \leq x_{(n)}.$$

   Per ogni $i = 1, ..., n$, avremo che $x_{(1)} \leq x_i \leq x_{(n)}$; quindi $\sum_{i=1}^{n} x_{(1)} \leq \sum_{i=1}^{n} x_i \leq \sum_{i=1}^{n} x_{(n)}$, cioè $nx_{(1)} \leq \sum_{i=1}^{n} x_i \leq nx_{(n)}$, da cui $x_{(1)} \leq \mu \leq x_{(n)}$.

2. La somma dei valori assunti dal carattere è uguale alla media aritmetica moltiplicata per il numero di unità

   $$\sum_{i=1}^{n} x_i = n\mu.$$

3. La somma degli scarti tra i valori osservati e la media aritmetica è nulla, per cui μ costituisce il *baricentro* di una distribuzione.

   Dimostrazione.

   $$\sum_{i=1}^{n}(x_i - \mu) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \mu = n\mu - n\mu = 0$$

4. La media aritmetica è l'unico valore che rende minima la somma degli scarti al quadrato.

   In altre parole, la somma degli scarti al quadrato da qualsiasi valore diverso dalla media aritmetica è superiore alla somma dei quadrati degli scarti dalla media aritmetica,

   $$\sum_{i=1}^{n}(x_i - c)^2 > \sum_{i=1}^{n}(x_i - \mu)^2, \quad \forall\, c \neq \mu.$$

   Dimostrazione. Partiamo dal considerare la somma degli scarti al quadrato rispetto ad un generico valore c ed eseguiamo alcuni semplici passaggi algebrici

   $$\begin{aligned}
   \sum_{i=1}^{n}(x_i - c)^2 &= \sum_{i=1}^{n}(x_i - \mu + \mu - c)^2 \\
   &= \sum_{i=1}^{n}\Big[(x_i - \mu) + (\mu - c)\Big]^2 \\
   &= \sum_{i=1}^{n}\Big[(x_i - \mu)^2 + (\mu - c)^2 + 2(x_i - \mu)(\mu - c)\Big] \\
   &= \sum_{i=1}^{n}(x_i - \mu)^2 + \sum_{i=1}^{n}(\mu - c)^2 + 2(\mu - c)\sum_{i=1}^{n}(x_i - \mu) \\
   &= \sum_{i=1}^{n}(x_i - \mu)^2 + n(\mu - c)^2.
   \end{aligned}$$

   Poiché il secondo termine $n(\mu - c)^2$ è strettamente maggiore di 0 per $c \neq \mu$, possiamo concludere che $\sum_{i=1}^{n}(x_i - c)^2 > \sum_{i=1}^{n}(x_i - \mu)^2$ qualunque sia la costante c diversa dalla media aritmetica.

5. **Proprietà associativa.** Se una popolazione è suddivisa in H gruppi disgiunti di numerosità $n_1, n_2, ..., n_H$ al cui interno il carattere X presenta le medie aritmetiche $\mu_1, \mu_2, ..., \mu_H$, la media aritmetica complessiva nella popolazione si ottiene come

$$\mu = \frac{\mu_1 n_1 + \mu_2 n_2 + ... + \mu_H n_H}{n_1 + n_2 + ... + n_H} = \frac{1}{n} \sum_{h=1}^{H} \mu_h n_h.$$

Dimostrazione. Indichiamo con x_i^h il valore osservato del carattere sull'i-esima unità appartenente all'h-esimo gruppo, $\forall\ h = 1, ..., H$ e $i = 1, ..., n_h$. La media aritmetica complessiva nella popolazione si ottiene come

$$\begin{aligned} \mu &= \frac{1}{n} \Big[\sum_{i=1}^{n_1} x_i^1 + \sum_{i=1}^{n_2} x_i^2 + ... + \sum_{i=1}^{n_H} x_i^H \Big] \\ &= \frac{1}{n} \Big[\mu_1 n_1 + \mu_2 n_2 + ... + \mu_H n_H \Big] = \frac{1}{n} \sum_{h=1}^{H} \mu_h n_h \end{aligned}$$

6. **Linearità.** La media aritmetica della variabile $Y = \alpha + \beta X$ (ottenuta da X mediante una trasformazione lineare con coefficienti α e β costanti) è pari a $\mu_Y = \alpha + \beta \mu_X$, dove μ_X è la media aritmetica di X.

Dimostrazione. La distribuzione unitaria della variabile Y risulta essere

$$y_1 = \alpha + \beta x_1, y_2 = \alpha + \beta x_2, ..., y_n = \alpha + \beta x_n$$

e quindi la media aritmetica di Y è data da

$$\mu_Y = \frac{1}{n} \sum_{i=1}^{n} y_i = \frac{1}{n} \sum_{i=1}^{n} (\alpha + \beta x_i) = \frac{1}{n} n\alpha + \frac{\beta}{n} \sum_{i=1}^{n} x_i = \alpha + \beta \mu_X.$$

7. **Monotonia.** Si consideri la distribuzione ordinata delle modalità osservate

$$x_{(1)} \le x_{(2)} \le ... \le x_{(i)} \le ... \le x_{(j)} \le ... \le x_{(n)}.$$

Indicando con μ_i la media aritmetica dei primi i valori e con μ_j la media aritmetica dei primi j valori, con $i < j$, si ha che

$$\mu_i \le \mu_j,$$

dove il segno di uguaglianza vale se e soltanto se

$$x_{(1)} = x_{(2)} = ... = x_{(i)} = ... = x_{(j)},$$

mentre in ogni altro caso

$$\mu_i < \mu_j.$$