Esempio 1: Media aritmetica di una distribuzione unitaria

- Risponde a domande del tipo: Qual è l'anzianità di servizio media dei 15 impiegati ?

10	12	2	3	4	6	8	9	10	11	12	2	3	13	13

$$\mu = \frac{10+12+12+3+4+6+8+9+10+11+2+3+13+13}{15} =$$

$$\frac{118}{15} = 7,86$$

Media aritmetica per una distribuzione di frequenze

- Ogni singolo valore deve pesare a seconda del numero di volte con cui si presenta nella distribuzione in esame
- **La media aritmetica per una distribuzione di frequenze è data da:**

$$\mu = \frac{1}{n} \sum_{j=1}^{H} y_j n_j$$

dove

- y_j per $j = 1, \ldots, H$ sono i valori distinti assunti dalla variabile y
- n_j per $j = 1, \ldots, H$ sono le frequenze assolute con cui ogni valore y_j è osservato nella distribuzione
- se sostituiamo $\frac{n_j}{n} = f_j$

$$\mu = \sum_{j=1}^{H} y_j f_j$$

Esempio 1: Media aritmetica per una distribuzione di frequenze

- y = anzianità di servizio espressa in anni
- i valori osservati sono

$$10, 12, 2, 3, 4, 6, 8, 9, 10, 11, 12, 2, 3, 13, 13$$

- n=15
- i valori distinti (y_j) assunti dalla variabile sono 2,3,4,6,8,9,10,11,12,13
- le corrispondenti frequenze assolute (n_j) con cui i valori y_j si presentano nella distribuzione sono

  2, 2, 1, 1, 1, 1, 2, 1 ,2, 2
- la distribuzione di frequenze è

Esempio 2 -Media aritmetica per una distribuzione di frequenze

La distribuzione di frequenze della variabile y (Anzianità di servizio misurata in anni) è

anzianità y_j	frequenze assolute n_j	frequenze relative f_j
2	2	0,133
3	2	0,133
4	1	0,067
6	1	0,067
8	1	0,067
9	1	0,067
10	2	0, 133
11	1	0,067
12	2	0,133
13	2	0,133
Totale	15	1,000

La media aritmetica della variabile è uguale a

$$\mu = \frac{1}{n}\sum_{j=1}^{H} y_j n_j$$

$$\mu = \frac{2(2) + 3(2) + 4(1) + 6(1) + 8(1) + 9(1) + 10(2) + 11(1) + 12(2) + 13(2)}{15} = 7,87$$

Posso calcolare la media utilizzando le frequenze relative

$$\mu = \sum_{j=1}^{H} y_j f_j$$

$$\mu = 2(0,133)+3(0,133)+4(0,067)+\ldots+11(0,067)+12(0,133)+13(0,133) = 7,87$$

Proprietà della media aritmetica

1 Il totale del carattere posseduto da n unità è uguale alla media aritmetica moltiplicata per n

$$\sum_{i=1}^{n} y_i = n\mu$$

Esempio: Verifichiamo la proprietà partendo dall'esempio relativo all'anzianità di servizio dei 15 impiegati

$$10, 12, 12, 3, 4, 6, 8, 9, 10, 11, 12, 2, 3, 13, 13$$

$\sum_{i=1}^{n} y_i = 118 \quad 15(7,86) = 118$

2 La somma degli scarti delle osservazioni dalla media aritmetica è pari a 0

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

Esempio:

$$(10 - 7,86) + (12 - 7,86) + \cdots + (13 - 7,86) + (13 - 7,86) = 0$$

(Esercizio 3.79 libro Agresti-Finlay)

3 La sommatoria degli scarti al quadrato dei valori della distribuzione da una costante c è minima quando questo valore è uguale alla media aritmentica μ

$$\sum_{i=1}^{n}(y_i - c)^2 \, minima \, se \, c = \mu \tag{1}$$

Esempio: se $c = \mu$

$$(10-7,86)^2+(12-7,86)^2+(2-7,86)^2+\ldots+(13-7,86)^2+(13-7,86)^2 = 241,7$$

- prendiamo un qualsiasi altro valore $\neq$ da μ: $c = 9$

$$(10-9)^2+(12-9)^2+(2-9)^2+\ldots+(3-9)^2+(13-9)^2+(13-9)^2 = 261$$

$241,7 < 261$(Esercizio 3.81 libro Agresti-Finlay)

4 La media aritmetica è sempre compresa tra il minimo e il massimo dei termini della distribuzione

Esempio: la media è maggiore di 2 e minore di 13

5 Se le n osservazioni che costituiscono il collettivo statistico appartengono a H sottoinsiemi (o gruppi) di numerosità $n_1, \ldots, n_H$ (tale che $\sum_{j=1}^{H} n_j = n$) e il valore medio all'interno di ciascun sottocollettivo è pari a $\mu_1, \ldots, \mu_H$, la media aritmetica è uguale a

$$\mu = \frac{1}{n}(\mu_1 n_1 + \ldots + \mu_H n_H)$$

Esempio: Importo medio (in euro) delle pensioni per area geografica in Italia:

- Nord, $\mu_1 = 13046$, $n_1 = 9.304.918$;
- Centro $\mu_2 = 13565$, $n_2 = 3.669.715$;
- Mezzogiorno $\mu_3 = 11845$, $n_3 = 5.135.978$

L'importo medio delle pensioni nelle tre aree geografiche è pari a

$$\mu = \frac{1}{18.110.611}(13.046 \times 9.304.918 + 13.565 \times 3.669.715 + 11.845 \times 5.135.978) =$$

$$\mu = 12810,88$$

Media aritmetica di una variabile quantitativa ripartita in classi

- Se la tabella si riferisce ad una variabile quantitativa ripartita in classi occorre avanzare delle ipotesi semplificatrici per determinare un valore unico da attribuire a ciascuna classe
- La scelta tipica è di attribuire a ciascuna classe il valore centrale:

$$\tilde{y}_j$$

$$\mu = \sum_{j=1}^{H} \tilde{y}_j f_j$$

Media aritmetica di un carattere suddiviso in classi

Tabella: Composizione degli iscritti al primo anno per voto di maturità

Voto di maturità	Frequenze Assolute n_j
60 – 69	122
70 – 79	92
80 – 89	64
90 – 100	31
Totale	**309**

Fonte: Rapporto di Autovalutazione 2011-12

Un'approssimazione del valore medio si può ottenere considerando $y_j = \tilde{y}_j$ valore centrale della classe

$$\bar{y} = \frac{1}{n} \sum_j \tilde{y}_j n_j$$

Media aritmetica di un carattere suddiviso in classi

Tabella: Composizione degli iscritti al primo anno per voto di maturità

Voto di maturità	$\tilde{y}$	Frequenze Assolute n_j
60 – 69	64.5	122
70 – 79	74.5	92
80 – 89	84	64
90 – 100	95	31
Totale		**309**

Fonte: Rapporto di Autovalutazione 2011-12

Tramite la semisomma dei valori della classe otteniamo i valori

$$\frac{60+69}{2} = 64.5$$

$$\frac{70+79}{2} = 74.5$$