5.2.4. La media aritmetica

Se e solo se il carattere è quantitativo è possibile calcolare la media aritmetica della distribuzione.

Il concetto di media aritmetica è intuitivo. Se si chiede a ciascuno degli studenti presenti in classe di quanti componenti è formata la famiglia alla quale appartiene, e si chiede poi di calcolare il numero medio dei componenti la famiglia per la classe, gli studenti la ottengono senza tentennamenti, sommando tutti i componenti e dividendo per il numero delle risposte. Se un capofamiglia vuol conoscere nel mese di gennaio il reddito medio mensile pro-capite per ciascun componente della propria famiglia, egli sa che deve sommare il reddito mensile di gennaio, qualunque ne sia la fonte, e dividere il reddito totale ottenuto per il numero dei componenti la famiglia. Il concetto di media aritmetica come di uno strumento che equiripartisce un totale fra ciascuno degli addendi che ad esso hanno contribuito è immediato, ed il calcolo non pone problemi particolari quando si conosce la modalità del carattere presente in ciascuna unità, ossia quando si ha a disposizione una distribuzione statistica unitaria. La determinazione può porre qualche problema se si ha una distribuzione di frequenze. Si distinguerà dunque la trattazione a seconda del modo in cui si presentano le informazioni.

a) Distribuzione unitaria

Se in una distribuzione unitaria, le n osservazioni del carattere X assumono rispettivamente i valori ξ_1, ξ_2,..., ξ_i,...,ξ_n, la media aritmetica (μ) si può esprimere come:

$$\mu = \frac{\xi_1 + \xi_2 + \ldots + \xi_i + \ldots + \xi_n}{n} = \frac{\sum_{i=1}^{n} \xi_i}{n},$$

dove $\sum_{i=1}^{n} \xi_i$ è un simbolo compatto, un modo abbreviato di indicare la somma.

E' evidente che se il carattere è espresso in cm, anche μ è espresso in cm, se il carattere è espresso in £, anche μ è espresso in £, e così via. Se min(X) è la modalità minima del carattere e max (X) la massima, è immediatamente comprensibile che $\min(X) < \mu < \max(X)$ (si dice che la media aritmetica gode della proprietà di internalità) e se tutte le modalità sono uguali fra loro, ossia se ξ_i = cost., per i=1,2,...,n, allora è μ=cost.=max(X)=min(X). L'internalità della media aritmetica, consente da un lato di tenere sotto controllo il risultato, infatti se la diseguaglianza non è rispettata, il calcolo della media è senz'altro errato; ma consente anche di affermare che μ può non coincidere con i valori osservati.

Si è visto, tuttavia, che i dati possono essere disposti anche in forma tabellare ed è perciò necessario saper calcolare la media anche delle distribuzioni di frequenze.

b) Distribuzione di frequenze per carattere discreto con modalità espresse singolarmente

La classificazione del collettivo secondo il voto alla maturità ha fornito come prima elaborazione la frequenza degli studenti voto per voto (Tab.6).

Tab. 6 - Studenti presenti alla lezione di statistica del 6-10-1994 per voto alla maturità

	Voto alla maturità																						
	36	37	38	40	41	42	43	44	45	46	48	50	51	52	53	54	55	56	57	58	60	n.i	Tot.
N. studenti	5	1	8	8	2	12	1	8	2	3	6	8	1	8	1	7	1	3	1	5	17	1	109

Si ha dunque che 5 studenti hanno superato l'esame con 36; 1 con 37; 8 con 38 e così via. Il concetto della media aritmetica, non dipende, ovviamente, dalla tabella sulla quale essa viene calcolata. Si tratta di sommare i voti di tutti gli studenti e dividere il "monte voto" per il loro numero. In questo caso si può però sfruttare il fatto che nella somma vi sono addendi uguali e scrivere:

$$\mu = \frac{36x5 + 37x1 + 38x8 + \ldots + 60x17}{108} = 48{,}39$$

Perciò si ottiene ~~, ancora una volta,~~ che il voto medio alla maturità è 48,39/60.

Formalmente dunque

$$\mu = \frac{\sum_{i=1}^{k} x_t n_t}{n} = \sum_{i=1}^{k} x_t f_t = \frac{\sum_{i=1}^{K} x_t p_t}{100}, \qquad \text{(ii)}$$