

Μάθημα 9

§5.2 Γεωμετρικός και Αρμονικός Μέσος

Έχουν περιορισμένη εφαρμογή στη Στατιστική αλλά παρουσιάζουν ενδιαφέρον κατά τη μελέτη ορισμένων οικονομικών μεθεθών όταν για παράδειγμα επιδιώκεται η εντίρμηση μέσω ρυθμών μεταβολής αυτών των μεθεθών ή σε περιπτώσεις που τα δεδομένα εκφράζουν αναλογίες ποσοτήτων (π.χ. ταχύτητα ανά ώρα, αραιβές ανά ώρα κλπ.). Ο γεωμετρικός μέσος εφαρμόζεται στην εντίρμηση της μέσης ετήσιας μεταβολής (αύξησης ή μείωσης) του όγκου ή της αξίας της παραγωγής προϊόντων, στον υπολογισμό του ρυθμού μεταβολής του πληθυσμού κλπ. Ο αρμονικός μέσος ενδύκνεται για την εντίρμηση της μέσης παραγωγικότητας εργαζομένων, μηχανών κλπ.

(α) Γεωμετρικός Μέσος: Για ένα σύνολο N ποσοτικών δεδομένων ο γεωμετρικός μέσος (geometric mean) ορίζεται ως η N -οστή ρίζα του γινομένου των τιμών των διαδέσφων παρατηρήσεων. Για το σύνολο των μετρήσεων $x_1, \dots, x_N$ όπου $x_i > 0 \ \forall i, i=1, \dots, N$, ο πληθυσμιακός γεωμετρικός μέσος G είναι:

$$G = \sqrt[N]{\prod_{i=1}^N x_i} = \sqrt[N]{x_1 x_2 x_3 \dots x_N}$$

Κατ' αναλογία ο γεωμετρικός μέσος του δείγματος ($\bar{g}$) είναι:

$$\bar{g} = \sqrt[n]{\prod_{i=1}^n x_i} = \sqrt[n]{x_1 x_2 x_3 \dots x_n}$$

Η τιμή του γεωμετρικού μέσου εκφράζεται σε μονάδες μέτρησης της μεταβλητής. Ο πληθυσμιακός γεωμετρικός

μέσος γράφεται:

$$G = \sqrt[N]{\prod_{i=1}^N x_i} = \sqrt[N]{x_1 x_2 x_3 \dots x_N} = (x_1 x_2 x_3 \dots x_N)^{\frac{1}{N}}$$

Παίρνοντας τους λογαρίθμους και των δύο μελών έχουμε:

$$\log G = \frac{1}{N} (\log x_1 + \dots + \log x_N) = \frac{\sum_{i=1}^N \log x_i}{N}, \text{ δηλαδή}$$

ο λογαρίθμος του γεωμετρικού μέσου είναι ίσος με τον αριθμητικό μέσο των λογαρίθμων των τιμών των διαθέσιμων παρατηρήσεων.

Παράδειγμα: Έστω δεδομένα 4, 2, 5, 6.

$$\bar{g} = \sqrt[4]{4 \cdot 2 \cdot 5 \cdot 7} = \sqrt[4]{280} = 4,0906$$

$$\begin{aligned} \log(\bar{g}) &= \frac{\log 4 + \log 2 + \log 5 + \log 7}{4} \\ &= \frac{0,60206 + 0,30103 + 0,69897 + 0,84510}{4} \\ &= \frac{2,44716}{4} \Rightarrow \log(\bar{g}) = 0,61179 \Rightarrow \bar{g} = 10^{0,61179} \end{aligned}$$

$$\Rightarrow \bar{g} = 4,0906$$

(β) Αρμονικός Μέσος: Για ένα σύνολο μετρήσεων $\{x_i, i=1, 2, \dots, n\}$ με $x_i > 0 \forall i$ ο αρμονικός μέσος ορίζεται ως το αντίστροφο του αριθμητικού μέσου των αντίστροφων τιμών των παρατηρήσεων. Διαφορετικά, το αντίστροφο του αρμονικού μέσου ενός συνόλου παρατηρήσεων ορίζεται ως ο αριθμητικός μέσος των αντίστροφων τιμών της μεταβλητής.

Έστω H ο πληθυσμιακός και $\bar{h}$ ο δειγματικός αρμονικός μέσος, αντίστοιχα. Οι αντίστοιχοι τύποι είναι:

Για τον πληθυσμό μεγέθους N :

-51-

$$\frac{L}{H} = \frac{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_N}}{N} = \frac{\sum_{i=1}^N \frac{1}{x_i}}{N} \Rightarrow H = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}}$$

Για το δείγμα μεγέθους n :

$$\frac{l}{h} = \frac{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}{n} = \frac{\sum_{i=1}^n \frac{1}{x_i}}{n} \Rightarrow$$

$$\Rightarrow \bar{h} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}, \text{ όπου } x_i > 0 \forall i, i=1, \dots, N \text{ ή } i=1, \dots, n$$

Για ένα σύνολο δεδομένων, η τιμή του αρμονικού μέσου εκφράζεται στις μονάδες μέτρησης της μεταβλητής.

Παράδειγμα Ο αρμονικός μέσος των δεδομένων 2, 5, 8, 4 είναι

$$\bar{h} = \frac{4}{\frac{1}{2} + \frac{1}{5} + \frac{1}{8} + \frac{1}{4}} = \frac{4}{1,075} = 3,7209$$

(1) Ιδιότητες του Γεωμετρικού Μέσου

$$\text{Έστω } u_i = \frac{x_i}{G} \text{ τότε } \prod_{i=1}^N u_i = \prod_{i=1}^N \left(\frac{x_i}{G} \right) = 1$$

Απόδειξη

$$\text{Είναι } \prod_{i=1}^N u_i = \prod_{i=1}^N \frac{x_i}{G} \Rightarrow \log \left(\prod_{i=1}^N u_i \right) = \log \left(\prod_{i=1}^N \frac{x_i}{G} \right)$$

$$\Rightarrow \sum_{i=1}^N \log u_i = \sum_{i=1}^N (\log x_i - \log G) = \sum_{i=1}^N \log x_i - N \log G$$

$$\text{Όπως } N \log G = \sum_{i=1}^N \log x_i$$

$$\text{Άρα } \sum_{i=1}^N \log u_i = \sum_{i=1}^N \log x_i - N \log G = \sum_{i=1}^N \log x_i - \sum_{i=1}^N \log x_i$$

$$\Rightarrow \sum_{i=1}^N \log u_i = 0 \Rightarrow \prod_{i=1}^N u_i = 1$$

$$(2) \min \sum_{i=1}^N \left(\frac{x_i}{a} \right)^2 = G$$

Απόδειξη

$$\frac{\partial}{\partial a} \sum_{i=1}^N (\log x_i - \log a)^2 = 0 \Rightarrow \sum_{i=1}^N (\log x_i - \log a) = 0$$

$$\Rightarrow \sum_{i=1}^N \log x_i = N \log a \Rightarrow \log a = \frac{\sum_{i=1}^N \log x_i}{N} = \log G.$$

Ιδιότητες του Αρμονικού Μέσου

(1) Οι απομείσεις $u_i = \frac{1}{x_i} - \frac{1}{H}$ λαμβάνονται αρμονικές απομείσεις. Ισχύει ότι $\sum_{i=1}^N u_i = \sum_{i=1}^N \left(\frac{1}{x_i} - \frac{1}{H} \right) = 0$

Απόδειξη

$$\sum_{i=1}^N \left(\frac{1}{x_i} - \frac{1}{H} \right) = \sum_{i=1}^N \frac{1}{x_i} - \frac{N}{H} = \sum_{i=1}^N \frac{1}{x_i} - \sum_{i=1}^N \frac{1}{x_i} = 0$$

Διότι από τον τύπο του αρμονικού μέσου, έχουμε:

$$\frac{1}{H} = \frac{\sum_{i=1}^N \frac{1}{x_i}}{N} \Rightarrow \frac{N}{H} = \sum_{i=1}^N \frac{1}{x_i}$$

(2) Το άθροισμα των τετραγώνων των αρμονικών απομείσεων γίνεται ελάχιστο για την τιμή του αρμονικού μέσου.

$$\min \sum_{i=1}^N \left(\frac{1}{x_i} - \frac{1}{a} \right)^2 = H$$

Απόδειξη $\frac{\partial}{\partial a} \sum_{i=1}^N \left(\frac{1}{x_i} - \frac{1}{a} \right)^2 = \frac{2}{a^2} \sum_{i=1}^N \frac{1}{x_i} - \frac{2aN}{a^4} = 0$

$$\Rightarrow \frac{1}{a^2} \sum_{i=1}^N \frac{1}{x_i} = \frac{aN}{a^4} \Rightarrow \frac{1}{a^2} \sum_{i=1}^N \frac{1}{x_i} = \frac{N}{a^3} \text{ και συνεπώς}$$

$$\sum_{i=1}^N \frac{1}{x_i} = \frac{N}{a} \Rightarrow a = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}} = H.$$

§ 5.3 Σύγκριση μεταξύ των τριών μέτρων κεντρικής τάσης
Ο αριθμητικός μέσος "επηρεάζεται" περισσότερο από τα
άλλα δύο μέτρα από ακραίες τιμές των δεδομένων. Για
το ίδιο σύνολο μετρήσεων η τιμή του αριθμητικού μέσου
είναι η μεγαλύτερη μεταξύ των τριών. Γενικά ισχύει ότι:

$$\bar{x} > \bar{g} > \bar{h}$$

Παράδειγμα 1 Έστω τα δεδομένα 8, 7, 10, 14, 11.

$$\text{Είναι } \bar{x} = \frac{50}{5} = 10, \quad \bar{g} = \sqrt[5]{86240} = 9,708$$

$$\bar{h} = \frac{5}{0,5302} = 9,430$$

Παράδειγμα 2 Έστω ότι αντί για 11 θάζουμε την ακραία
τιμή 110 δηλ. τα νέα δεδομένα είναι: 8, 7, 10, 14, 110. Τότε

$$\bar{x} = \frac{149}{5} = 29,8, \quad \bar{g} = \sqrt[5]{862400} = 15,286$$

$$\bar{h} = \frac{5}{0,44837} = 11,151. \text{ Η τιμή του αριθμητικού μέσου}$$

αυξήθηκε σε σχέση με το Παράδειγμα 1 κατά 19,8
μονάδες ή 198% η τιμή του γεωμετρικού μέσου αυξή-
θηκε κατά 5,578 μονάδες ή 57% και η τιμή του
αρμονικού μέσου αυξήθηκε κατά 1,721 μονάδες
(ή 18%).

§ 5.4 Σταθμισμένος Αριθμητικός Μέσος

Έστω η κατανομή 10 εργαζομένων πίας επιχειρήσεων κατά
ομάδες ηλικιών:

-54-

Ομάδες ηλικιών σε έτη	# Εργαζομένων (f_i)	Μεντριές τιμές τάξεων (w_i)
20-30	2	25
30-40	4	35
40-50	3	45
50-60	1	55
Σύνολο	10	

Η υπόθεση ότι οι συχνότητες f_i αναφέρονται στα κέντρα των διαστημάτων τάξεων ισοδυναμεί με την παραδοχή ότι 2 εργαζόμενοι έχουν ηλικία 25 ετών, 4 εργαζόμενοι έχουν ηλικία 35 ετών κ.ο.κ. Η μέση ηλικία μπορεί να επισημωθεί ως εξής:

$$\bar{x} = \frac{2 \cdot 25 + 4 \cdot 35 + 3 \cdot 45 + 1 \cdot 55}{2 + 4 + 3 + 1} = \frac{390}{10} = 39$$

Γενικά $\bar{x} = \frac{\sum_{i=1}^K f_i w_i}{\sum_{i=1}^K f_i}$ όπου $K := \# \text{ τάξεων}$. Ο $\bar{x}$ καλείται

σταθμικός μέσος όρος διότι για τον υπολογισμό του σταθμίζονται οι συχνότητες (f_i) με τα κέντρα των τάξεων (w_i). Καλείται και βαρυμετρικός μέσος.