

§5.5 Σταθμικός Γεωμετρικός Μέσος (Weighted geometric mean)

Για ένα σύνολο n παρατηρήσεων ομαδοποιημένων σε μία κατανομή συχνοτήτων με k τάξεις, ο σταθμικός γεωμετρικός μέσος δίνεται από τον τύπο:

$$\bar{g} = \sqrt[n]{\underbrace{w_1 w_1 \dots w_1}_{f_1} \underbrace{w_2 \dots w_2}_{f_2} \dots \underbrace{w_k w_k \dots w_k}_{f_k}} = \sqrt[n]{w_1^{f_1} w_2^{f_2} \dots w_k^{f_k}}$$

όπου οι κεντρικές τιμές των τάξεων $w_1, \dots, w_k$ εμφανίζονται με συχνότητες $f_1, \dots, f_k$ των οποίων το άθροισμα είναι $f_1 + f_2 + \dots + f_k = n$.

Η παραπάνω σχέση γράφεται: $\bar{g} = (w_1^{f_1} \dots w_k^{f_k})^{\frac{1}{n}}$

Παίρνοντας λογαρίθμους και των δύο μερών:

$$\log(\bar{g}) = \frac{1}{n} [f_1 \log w_1 + f_2 \log w_2 + \dots + f_k \log w_k]$$

$$\Rightarrow \log(\bar{g}) = \frac{\sum_{i=1}^k f_i \log w_i}{\sum_{i=1}^k f_i}, \text{ δηλαδή ο λογαρίθμος του στα-}$$

θμικού γεωμετρικού μέσου μία συνεχούς κατανομής συχνοτήτων είναι ίσος με τον σταθμικό αριθμητικό μέσο που υπολογίζεται χρησιμοποιώντας ως βάρη (σταθμίσες) τους λογαρίθμους των κεντρικών τιμών της κατανομής.

Παράδειγμα	Ομάδες Ηλικιών	f_i	w_i	$\log w_i$	$f_i \log w_i$
	20 - 30	2	25	1,398	2,796
	30 - 40	4	35	1,544	6,176
	40 - 50	3	45	1,653	4,959
	50 - 60	1	55	1,740	1,740

$$\log(\bar{g}) = \frac{\sum_{i=1}^4 f_i \log w_i}{\sum_{i=1}^4 f_i} = \frac{15,671}{10} = 1,5671 \Rightarrow \bar{g} = 10^{1,5671} = 36,9 \text{ έτη.}$$

§ 5.6 Σταθμισμός Αρμονικός Μέσος (Weighted harmonic mean)

Ορίζεται ως το πηλίκο του πηλούς των διαθέσιμων παρατηρήσεων προς το άθροισμα των γινόμενων των ταξικών συχνοτήτων (f_i) επί τα αντίστροφα των κεντρικών τιμών των διαστημάτων τάξεων ($\frac{1}{w_i}$) της κατανομής:

$$\frac{1}{\bar{h}} = \frac{f_1 \frac{1}{w_1} + f_2 \frac{1}{w_2} + \dots + f_k \frac{1}{w_k}}{f_1 + f_2 + \dots + f_k}$$

$$\Rightarrow \bar{h} = \frac{\sum_{i=1}^k f_i}{\sum_{i=1}^k f_i \frac{1}{w_i}} = \frac{n}{\sum_{i=1}^k \frac{f_i}{w_i}}$$

Παράδειγμα

Ομάδες Ηλικιών	f_i	w_i	f_i/w_i
20-30	2	25	0,080
30-40	4	35	0,114
40-50	3	45	0,066
50-60	1	55	0,018
Σύνολο	10		0,278

$$\bar{h} = \frac{\sum_{i=1}^4 f_i}{\sum_{i=1}^4 \frac{f_i}{w_i}} = \frac{10}{0,278} = 35,97 \text{ έτη.}$$

§ 5.7 Επιμερατούσα Τιμή (Mode) Σ' ένα σύνολο δεδομένων είναι η τιμή με την μεγαλύτερη συχνότητα. Καλείται επιμερατούσα ή κορυφή (mode). Τα δεδομένα δεν εμφανίζουν πάντοτε επιμερατούσα τιμή ενώ υπάρχουν περιπτώσεις όπου εντοπίζονται περισσότερα του ενός σημεία μέγιστης συχνότητας. Είναι το μόνο στατιστικό

περιγραφικό μέτρο που δεν λαμβάνει απαραίτητα μία και -57-
μόνο τιμή. Κατανομές με ένα μόνο σημείο μέγιστης συχνότητας
καλούνται μονοόρυφες (unimodal) με δύο διόρυφες (bimodal)
κλπ. Η επικρατούσα τιμή συμβολίζεται με T αν πρόκειται για
πληθυσμό και με t αν πρόκειται για δείγμα.

Παραδείγματα

- (α) Τα δεδομένα 3, 5, 1, 7, 10, 7, 5, 7, 9, 2, 7 έχουν επικρατούσα τιμή
το 7 με συχνότητα εμφάνισης 4.
- (β) Τα δεδομένα 10, 11, 14, 20, 14, 11, 7, 18, 14, 11 παρουσιάζουν
δύο κορυφές, την τιμή 11 με συχνότητα 3 και την τιμή 14
με την ίδια συχνότητα.
- (γ) Τα δεδομένα 4, 7, 1, 9, 2, 10 δεν εμφανίζουν κανένα σημείο μέ-
γιστης συχνότητας.
- (δ) Τροποποιώντας τα δεδομένα του (α) και προσθέτοντας δύο
υψηλές τιμές: 3, 5, 1, 7, 10, 7, 5, 7, 9, 2, 7, 200, 500. Τα νέα δεδο-
μένα εξακολουθούν να έχουν επικρατούσα τιμή το 7, δηλαδή
το σημείο μέγιστης συχνότητας δεν επηρεάζεται από αιρ-
ες τιμές της μεταβλητής.
- (ε) Οικογενειακή κατάσταση 25 ατόμων:

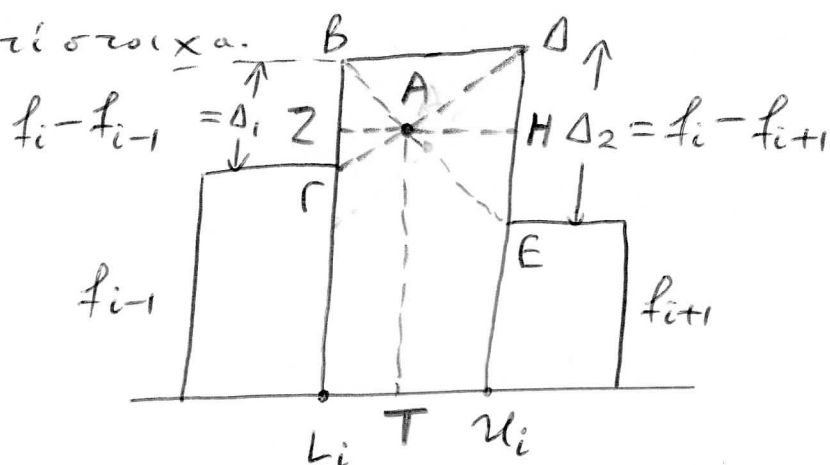
Ε Ε Ε Α Χ Δ Ε Ε Χ Α Ε Α Χ Ε Ε Δ Χ Α Ε Α Ε Α Ε Α Α, Α ν

1: Άγαροι, 2: Έγγαροι, 3: Χήροι, 4: Διαζευχθέντες, τα παραπάνω
δεδομένα γράφονται:

2 2 2 1 3 4 2 2 3 1 2 1 3 2 2 4 3 1 2 1 2 1 2 1 1

Οι Έγγαροι (συμβολική τιμή: Ε, αριθμητική τιμή: 2) είναι
η πολύ πληθύνουσα κατηγορία. Η επικρατούσα τιμή είναι
η 2 ή η Ε με συχνότητα $f=11$.

Στη περίπτωση οραδοποιημένων δεδομένων θεωρούμε ότι η -58-
 υπό μελέτη κατανομή είναι μονοκύρφη διγαδρή υπάρχει μία
 μόνο τάξη που παρουσιάζει μέγιστη συχνότητα. Έστω ότι η
 μέγιστη συχνότητα παρατηρείται στην i -οστή τάξη και f_i
 είναι η μέγιστη συχνότητα της κατανομής. Έστω L_i και
 U_i αντίστοιχα το κατώτερο και το ανώτερο όριο αυτής
 της τάξης. Οι συχνότητες της αμέσως προηγούμενης τάξης
 και της αμέσως επόμενης τάξης συμβολίζονται ως f_{i-1} και
 f_{i+1} αντίστοιχα.



Το σχήμα παρουσιάζει το ισόγραμμα μιας τέτοιας κατανομής.
 Δημιουργούνται δύο όμοια τρίγωνα, ABG , ADE για τα οποία

$$\text{έχουμε: } \frac{(AZ)}{(BG)} = \frac{(AH)}{(DE)} \Rightarrow \frac{T - L_i}{\Delta_1} = \frac{U_i - T}{\Delta_2}$$

$$\Rightarrow T = \frac{\Delta_1 U_i + \Delta_2 L_i}{\Delta_1 + \Delta_2} = \frac{\Delta_1 (L_i + \delta_i) + \Delta_2 L_i}{\Delta_1 + \Delta_2}$$

$$= \frac{L_i (\Delta_1 + \Delta_2) + \delta_i \Delta_1}{\Delta_1 + \Delta_2} = L_i + \delta_i \frac{\Delta_1}{\Delta_1 + \Delta_2}, \quad \delta_i = U_i - L_i$$

$$\text{δηλαδή } T = L_i + \delta_i \frac{\Delta_1}{\Delta_1 + \Delta_2}, \text{ όπου}$$

T = το σημείο μέγιστης συχνότητας της κατανομής
 i = η τάξη με την μεγαλύτερη συχνότητα από όλες τις

άλλες τάξεις της μονοκύριας κατανομής

f_i = η μέγιστη (απόλυτη ή σχετική) συχνότητα της μονοκύριας κατανομής

L_i = το κατώτερο όριο της τάξης η οποία παρουσιάζει την μέγιστη συχνότητα

δ_i = το πλάτος της τάξης που παρουσιάζει τη μέγιστη συχνότητα

$\Delta_1 = f_i - f_{i-1}$ = η διαφορά ανάμεσα στη μέγιστη συχνότητα και τη συχνότητα της αμέσως προηγούμενης τάξης

$\Delta_2 = f_i - f_{i+1}$ = η διαφορά ανάμεσα στη μέγιστη συχνότητα και τη συχνότητα της αμέσως επόμενης τάξης

Ο τύπος δίνει κατά προσέγγιση την επικρατούσα τιμή και η εφαρμογή του δεν είναι εφικτή εάν η καρπύση είναι πορτοκύρια.

Παράδειγμα Κατανομή των γάμων κατά ηλικία της νέας το έτος 1988 στις ημιαστικές περιοχές της Ελλάδας.

Στόχος μας να επισημειώσουμε την επικρατούσα ηλικία της νέας.

Ομάδες Ηλικιών	Αριθμός Γάμων	Πηγή: ΕΣΥΕ Πίνακας 4, σελ. 120.
Μέχρι 15 ετών	17	
15-20	1352	
20-25	2107	
25-30	875	
30-35	276	
35-40	115	
40-45	76	
45-50	50	
50-55	41	
55-60	35	
60-65	20	
65-άνω	12	
Σύνολο	4976	

Τα στοιχεία δείχνουν ότι πρόκειται για μονοκύρφη κατανομή. Η μέγιστη συχνότητα παρατηρείται στην τρίτη τάξη δηλαδή στις ηλικίες μεταξύ 20 και 25 ετών.

$$\begin{aligned} \text{Είναι: } T &= L_3 + \delta_3 \frac{\Delta_1}{\Delta_1 + \Delta_2} = L_3 + \delta_3 \frac{(f_3 - f_2)}{(f_3 - f_2) + (f_3 - f_4)} \\ &= 20 + 5 \frac{(2107 - 1352)}{(2107 - 1352) + (2107 - 875)} = 20 + 5 \frac{755}{755 + 1232} \\ &= 21,9 \text{ έτη.} \end{aligned}$$