MATHEMATICS

Pure mathematics is, in its way, the poetry of logical ideas.

ΠΊΝΑΚΕΣ ΣΥΧΝΟΤΉΤΩΝ

Μετά τη συλλογή των στατιστικών δεδομένων είναι αναγκαία η κατασκευή συνοπτικών πινάκων που περιέχουν όλες τις απαραίτητες πληροφορίες του δείγματος. Ας υποθέσουμε ότι εξετάζουμε ένα δείγμα μεγέθους ν ως προς ένα χαρακτηριστικό που εκφράζεται από την τυχαία μεταβλητή X. Υποθέτουμε ότι οι διαφορετικές τιμές που πήρε η μεταβλητή είναι: $x_1,\ x_2,\ ...,\ x_k$, προφανώς $k \leqslant \nu$. Γενικά, την τυχαία τιμή που παίρνει η μεταβλητή τη συμβολίζουμε με x_i όπου ο δείκτης i παίρνει τις τιμές $i = 1,\ 2, ..., k$.

Άσκηση 1

Στο πλαίσιο μιας έρευνας ρωτήσαμε 20 μαθητές Γ' Λυκείου πόσα αδέλφια έχουν. Οι απαντήσεις που πήραμε είναι:

$$2,\ 1,\ 0,\ 0,\ 1,\ 1,\ 1,\ 2,\ 3,\ 1,$$
$$0,\ 1,\ 1,\ 1,\ 0,\ 2,\ 2,\ 0,\ 3,\ 1.$$

1. Ποιος είναι ο πληθυσμός που μελετάμε;
2. Ποιο είναι το μέγεθος ν του δείγματος;
3. Ποια είναι η τυχαία μεταβλητή X και τι είδους είναι;
4. Ποιες είναι οι διαφορετικές τιμές που παίρνει η μεταβλητή και ποιο το πλήθος k των διαφορετικών τιμών;

Λύση

1. Ο πληθυσμός που μελετάμε είναι **οι μαθητές της Γ' Λυκείου**.
2. Το μέγεθος του δείγματος είναι $\nu = 20$.
3. Το χαρακτηριστικό ως προς το οποίο μελετάμε τον πληθυσμό είναι ο **αριθμός των αδελφών** άρα η μεταβλητή X παίρνει για κάθε άτομο του δείγματος τιμή ίση με τον αριθμό αδελφών του ατόμου, δηλαδή $X = 0,\ 1,\ 2,\ 3,\ 4, ...$ Η μεταβλητή είναι **ποσοτική** και **διακριτή**.
4. Οι διαφορετικές τιμές που παίρνει η μεταβλητή στο συγκεκριμένο δείγμα είναι $x_1 = 0,\ x_2 = 1,\ x_3 = 2,\ x_4 = 3$, άρα $k = 4$.

-Τέλος Λύσης-

Στην τιμή x_i αντιστοιχίζεται η (απόλυτη) **συχνότητα** (frequency) ν_i, δηλαδή ο φυσικός αριθμός που δείχνει πόσες φορές εμφανίζεται η τιμή x_i της εξεταζόμενης μεταβλητής X στο σύνολο των παρατηρήσεων. Είναι φανερό ότι το άθροισμα όλων των συχνοτήτων είναι ίσο με το μέγεθος ν του δείγματος, δηλαδή:

$$\nu_1+\nu_2+\cdots+\nu_k=\nu \tag{1}$$

Αν διαιρέσουμε τη συχνότητα ν_i με το μέγεθος ν του δείγματος, προκύπτει η σχετική συχνότητα (relative frequency) f_i της τιμής x_i, δηλαδή:

$$f_i=\frac{\nu_i}{\nu},\ i=1,2,...,k \tag{2}$$

Θεώρημα

Για τη σχετική συχνότητα ισχύουν οι ιδιότητες:

1. $0\leqslant f_i\leqslant 1$ για $i=1,2,...,k$
2. $f_1+f_2+\cdots+f_k=1$

Απόδειξη

1. Προφανώς για τη συχνότητα ν_i ισχύει $0\leqslant \nu_i\leqslant \nu$ οπότε:

$$0\leqslant \nu_i\leqslant \nu \Rightarrow \frac{0}{\nu}\leqslant\frac{\nu_i}{\nu}\leqslant\frac{\nu}{\nu}\Rightarrow 0\leqslant f_i\leqslant 1$$

2.

$$f_1+f_2+\cdots+f_k=\frac{\nu_1}{\nu}+\frac{\nu_2}{\nu}+\cdots+\frac{\nu_k}{\nu}=\frac{\nu_1+\nu_2+\cdots+\nu_k}{\nu}=\frac{\nu}{\nu}=1$$

Συνήθως, τις σχετικές συχνότητες f_i τις εκφράζουμε ως επί τοις εκατό ποσοστό και συμβολίζονται με $f_i\%$, δηλαδή $f_i\%=100\cdot f_i$.

Οι ποσότητες x_i, ν_i, f_i για ένα δείγμα συγκεντρώνονται σε ένα συνοπτικό πίνακα, που ονομάζεται **πίνακας κατανομής συχνοτήτων** ή απλά **πίνακας συχνοτήτων**. Για μια μεταβλητή, το σύνολο των ζευγών (x_i,ν_i) λέμε ότι αποτελεί την **κατανομή συχνοτήτων** και το σύνολο των ζευγών (x_i,f_i), ή των ζευγών $(x_i,f_i\%)$, την **κατανομή των σχετικών συχνοτήτων**.

Παρακάτω παρουσιάζεται η γενική μορφή που έχει ο πίνακας συχνοτήτων. Ο αριθμός των γραμμών του πίνακα χωρίς την πρώτη γραμμή (τίτλοι στηλών) και την τελευταία γραμμή (σύνολα) ισούται με k.

Αύξων Αριθμός	Τυχαία Μεταβλητή	Συχνότητα	Σχετική Συχνότητα	Επί τοις εκατό Σχετική Συχνότητα
1	x_1	ν_1	f_1	$f_1\%$
2	x_2	ν_2	f_2	$f_2\%$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	x_k	ν_k	f_k	$f_k\%$
	Σύνολο	ν	1	100

Η πρώτη στήλη (Αύξων Αριθμός) συνήθως παραλείπεται. Για κάθε στοιχείο του πίνακα ο δείκτης i αντιστοιχεί στη γραμμή του πίνακα στην οποία βρίσκεται το στοιχείο, έτσι για παράδειγμα το στοιχείο ν_3 αντιστοιχεί στη συχνότητα της 3ης γραμμής.

Άσκηση 2

Για τα δεδομένα της άσκησης 1 να κατασκευαστεί ο πίνακας συχνοτήτων, σχετικών συχνοτήτων και επί τοις εκατό σχετικών συχνοτήτων.

Λύση

Στη άσκηση 1 είδαμε ότι οι τιμές που παίρνει η μεταβλητή X είναι $x_1 = 0, x_2 = 1, x_3 = 2$ και $x_4 = 3$. Οι αντίστοιχες συχνότητες είναι $\nu_1 = 5, \nu_2 = 9, \nu_3 = 4$ και $\nu_4 = 2$. Για τις σχετικές συχνότητες έχουμε:

$$
\begin{aligned}
f_1 &= \frac{\nu_1}{\nu} = \frac{5}{20} = 0.25 \\
f_2 &= \frac{\nu_2}{\nu} = \frac{9}{20} = 0.45 \\
f_3 &= \frac{\nu_3}{\nu} = \frac{4}{20} = 0.2 \\
f_4 &= \frac{\nu_4}{\nu} = \frac{2}{20} = 0.1
\end{aligned}
$$

Τυχαία Μεταβλητή	Συχνότητα	Σχετική Συχνότητα	Επί τοις εκατό Σχετική Συχνότητα
0	5	0.25	25
1	9	0.45	45
2	4	0.2	20
3	2	0.1	10
Σύνολο	20	1	100

-Τέλος Λύσης-

Στην περίπτωση των **ποσοτικών μεταβλητών** εκτός από τις συχνότητες ν_i και τις σχετικές συχνότητες f_i χρησιμοποιούνται και οι λεγόμενες **αθροιστικές συχνότητες** (cumulative frequencies) N_i και οι **αθροιστικές σχετικές συχνότητες** (cumulative relative frequencies) F_i, οι οποίες εκφράζουν το πλήθος και το ποσοστό αντίστοιχα των παρατηρήσεων που είναι **μικρότερες ή ίσες** της τιμής x_i. Επομένως, η αθροιστική συχνότητα N_i υπολογίζεται από τη σχέση:

$$N_i = \nu_1 + \nu_2 + \cdots + \nu_i$$

και η αθροιστική σχετική συχνότητα F_i υπολογίζεται από τη σχέση:

$$F_i = f_1 + f_2 + \cdots + f_i$$

Πιο αναλυτικά για τις σχετικές συχνότητες έχουμε:

$$\begin{aligned} N_1 &= \nu_1 \\ N_2 &= \nu_1 + \nu_2 \\ N_3 &= \nu_1 + \nu_2 + \nu_3 \\ &\vdots \\ N_k &= \nu_1 + \nu_2 + \cdots + \nu_k \end{aligned}$$

και αντίστοιχα για τις σχετικές αθροιστικές συχνότητες έχουμε:

$$
\begin{aligned}
F_1 &= f_1 \\
F_2 &= f_1 + f_2 \\
F_3 &= f_1 + f_2 + f_3 \\
&\vdots \\
F_k &= f_1 + f_2 + \cdots + f_k
\end{aligned}
$$

Παρατηρούμε ότι η τελευταία (k) αθροιστική συχνότητα ισούται με το μέγεθος του δείγματος:

$$N_k = \nu_1 + \nu_2 + \cdots + \nu_k = \nu$$

και η τελευταία (k) αθροιστική σχετική συχνότητα ισούται με μονάδα:

$$F_k = f_1 + f_2 + \cdots + f_k = 1$$

Επίσης είναι φανερό ότι ισχύουν οι σχέσεις:

$$
\begin{aligned}
\nu_1 &= N_1 \\
\nu_2 &= N_2 - N_1 \\
\nu_3 &= N_3 - N_2 \\
&\vdots \\
\nu_k &= N_k - N_{k-1}
\end{aligned}
$$

και

$$
\begin{aligned}
f_1 &= F_1 \\
f_2 &= F_2 - F_1 \\
f_3 &= F_3 - F_2 \\
&\vdots \\
f_k &= F_k - F_{k-1}
\end{aligned}
$$

Άσκηση 3

Στον πίνακα της άσκησης 2 να προστεθούν οι στήλες αθροιστικών συχνοτήτων, αθροιστικών σχετικών συχνοτήτων και επί τοις εκατό αθροιστικών σχετικών συχνοτήτων.

Λύση

Για τις αθροιστικές συχνότητες έχουμε:

$$N_1 = \nu_1 = 5$$
$$N_2 = \nu_1 + \nu_2 = 5 + 9 = 14$$
$$N_3 = \nu_1 + \nu_2 + \nu_3 = 5 + 9 + 4 = 18$$
$$N_4 = \nu_1 + \nu_2 + \nu_3 + \nu_4 = 5 + 9 + 4 + 2 = 20$$

Για τις αθροιστικές σχετικές συχνότητες έχουμε:

$$F_1 = f_1 = 0.25$$
$$F_2 = f_1 + f_2 = 0.25 + 0.45 = 0.7$$
$$F_3 = f_1 + f_2 + f_3 = 0.25 + 0.45 + 0.2 = 0.9$$
$$F_4 = f_1 + f_2 + f_3 + f_4 = 0.25 + 0.45 + 0.2 + 0.1 = 1$$

Τυχαία Μεταβλητή	Συχνότητα	Σχετική Συχνότητα	Επί τοις εκατό Σχετική Συχνότητα	Αθροιστική Συχνότητα	Αθροιστική Σχετ. Συχνότητα	Επί τοις εκατό Αθροιστική Σχετ. Συχνότητα
0	5	$0,25$	25	5	0.25	25
1	9	$0,45$	45	14	0.7	70
2	4	$0,2$	20	18	0.9	90
3	2	$0,1$	10	20	1	100
Σύνολο	20	1	100			

-Τέλος Λύσης-

Προσπαθήστε μόνοι σας

Άσκηση 4

Να συμπληρωθεί ο πίνακας

Τυχαία Μεταβλητή	Συχνότητα	Σχετική Συχνότητα	Επί τοις εκατό Σχετική Συχνότητα	Αθροιστική Συχνότητα	Αθροιστική Σχετ. Συχνότητα	Επί τοις εκατό Αθροιστική Σχετ. Συχνότητα
x_1	5					

Τυχαία Μεταβλητή	Συχνότητα	Σχετική Συχνότητα	Επί τοις εκατό Σχετική Συχνότητα	Αθροιστική Συχνότητα	Αθροιστική Σχετ. Συχνότητα	Επί τοις εκατό Αθροιστική Σχετ. Συχνότητα
x_2	8					
x_3	12					
x_4	15					
x_5	10					
Σύνολο						

Άσκηση 5

Να συμπληρωθεί ο πίνακας

Τυχαία Μεταβλητή	Συχνότητα	Σχετική Συχνότητα	Επί τοις εκατό Σχετική Συχνότητα	Αθροιστική Συχνότητα	Αθροιστική Σχετ. Συχνότητα	Επί τοις εκατό Αθροιστική Σχετ. Συχνότητα
1			20			
2				25		
3					0.82	
4	9					
Σύνολο						

Άσκηση 6

Να συμπληρωθεί ο πίνακας

Τυχαία Μεταβλητή	Συχνότητα	Σχετική Συχνότητα	Επί τοις εκατό Σχετική Συχνότητα	Αθροιστική Συχνότητα	Αθροιστική Σχετ. Συχνότητα	Επί τοις εκατό Αθροιστική Σχετ. Συχνότητα

Τυχαία Μεταβλητή	Συχνότητα	Σχετική Συχνότητα	Επί τοις εκατό Σχετική Συχνότητα	Αθροιστική Συχνότητα	Αθροιστική Σχετ. Συχνότητα	Επί τοις εκατό Αθροιστική Σχετ. Συχνότητα
x_1	α					
x_2	β		20			
x_3	γ			33		
x_4	β				0.86	
x_5	α					
Σύνολο						

Άσκηση 7

Στο πλαίσιο μίας έρευνας ρωτήθηκαν ν άτομα πόσες ημέρες το μήνα πηγαίνουν σινεμά. Οι απαντήσεις που δόθηκαν ήταν από 0 έως 4 ημέρες. Ισχύουν:

- 5 άτομα πηγαίνουν σινεμά 4 ημέρες το μήνα.
- 85 άτομα πηγαίνουν σινεμά το πολύ 2 ημέρες το μήνα.
- Τα άτομα που πηγαίνουν σινεμά 2 ημέρες το μήνα είναι διπλάσια από τα άτομα που δεν πηγαίνουν καθόλου σινεμά.
- Το 45% πηγαίνει σινεμά τουλάχιστον 2 ημέρες το μήνα.
- Το 15% δεν πηγαίνει σινεμά.

Να συμπληρωθεί ο πίνακας συχνοτήτων με στήλες: x_i, ν_i, f_i, $f_i\%$, N_i, F_i και $F_i\%$.

Στείλε την προσπάθειά σου