

Εισαγωγή στη θεωρία Πιθανοτήτων και την Στατιστική

Οργάνωση & Παρουσίαση Δεδομένων

Ακατέργαστα ή Πρωτογενή δεδομένα

**Σε κλίμακα από 0 έως 5 (0=καθόλου, 5=πανικός) πόσο φοβόσασταν
τα μαθηματικά στο λύκειο;**

Οι απαντήσεις 50 φοιτητών στην ερώτηση ήταν οι εξής:

2, 3, 3, 2, 4, 4, 4, 5, 3, 2, 5, 1, 4, 4, 1, 0, 4, 3, 3, 4, 5, 3, 1, 3, 5,

1, 5, 4, 2, 5, 3, 4, 3, 5, 2, 4, 5, 4, 5, 5, 5, 5, 2, 4, 4, 2, 4, 5, 2, 4

Ακατέργαστα ή Πρωτογενή δεδομένα

Οι μη επεξεργασμένες μετρήσεις ονομάζονται **ακατέργαστα** ή **πρωτογενή δεδομένα** (raw data).

Κατάλληλη στατιστική επεξεργασία → οργάνωση των δεδομένων με τέτοιο τρόπο, ώστε να αναδεικνύονται οι πληροφορίες που περιέχουν & να περιγράφονται αποτελεσματικά οι κύριες τάσεις που επικρατούν στο δείγμα.

Πίνακας κατανομής συχνοτήτων

Απόλυτη συχνότητα ή **συχνότητα** ονομάζεται ο φυσικός αριθμός n_i που δείχνει πόσες φορές εμφανίζεται μια τιμή της μεταβλητής στο σύνολο των παρατηρήσεων.

Το άθροισμα όλων των συχνοτήτων είναι ίσο με το μέγεθος του δείγματος

Ένας πίνακας στον οποίο παρουσιάζονται οι διαφορετικές τιμές της μεταβλητής και οι αντίστοιχες συχνότητες ονομάζεται **πίνακας κατανομής συχνοτήτων**.

Πίνακας κατανομής συχνοτήτων

Σε κλίμακα από 0 έως 5 (0=καθόλου, 5=πανικός) πόσο φοβόσασταν τα μαθηματικά ;

2, 3, 3, 2, 4, 4, 4, 5, 3, 2, 5, 1, 4, 4, 1, 0, 4, 3, 3, 4, 5, 3, 1, 3, 5,
1, 5, 4, 2, 5, 3, 4, 3, 5, 2, 4, 5, 4, 5, 5, 5, 5, 2, 4, 4, 2, 4, 5, 2, 4

Πίνακας 2.1. Κατανομή συχνοτήτων τιμών της κλίμακας του φόβου.

Πόσο φοβόσασταν τα μαθηματικά στο λύκειο;	Αριθμός φοιτητών
0 (καθόλου)	1
1 (πολύ λίγο)	4
2 (λίγο)	8
3 (αρκετά)	9
4 (πολύ)	15
5 (πάρα πολύ)	13
	Σύνολο: 50

Κατασκευή πίνακα κατανομής συχνοτήτων

Βήμα 1. Σε μια στήλη αναγράφουμε όλες τις πιθανές τιμές της μεταβλητής. Αν η μεταβλητή είναι τακτική ή ποσοτική, οι τιμές της πρέπει να τοποθετούνται είτε σε αύξουσα είτε σε φθίνουσα σειρά.

Βήμα 2. Αρχίζοντας από την πρώτη κατά σειρά μέτρηση, βάζουμε μια γραμμή | στην στήλη δίπλα στην τιμή της και συνεχίζουμε μέχρι να εξαντλήσουμε όλες τις τιμές του δείγματος. Διαγράφουμε τη μέτρηση, μόλις αυτή καταχωριστεί.

Βήμα 3. Καθαρογράφουμε τα παραπάνω σε έναν πίνακα όπου δείχνουμε τη συχνότητα εμφάνισης της κάθε τιμής. Συμπεριλαμβάνουμε και το γενικό άθροισμα των συχνοτήτων που πρέπει να ισούται με το μέγεθος του δείγματος.

Εφαρμογή 1

Σε έναν δήμο υπάρχουν 4 υποψήφιοι Α, Β, Γ, Δ για τη δημαρχία. Σε μια στατιστική έρευνα που έγινε ρωτήσαμε 30 άτομα ποιον από τους 4 υποψήφιους θα ψηφίσουν στις επόμενες εκλογές και πήραμε τις εξής απαντήσεις:

Β	Γ	Α	Γ	Δ	Β
Γ	Δ	Β	Α	Γ	Γ
Α	Β	Α	Β	Γ	Δ
Β	Γ	Γ	Δ	Β	Α
Γ	Β	Β	Α	Γ	Γ

Να φτιάξετε έναν πίνακα κατανομής συχνοτήτων

Ομαδοποιημένες παρατηρήσεις

Για να μελετηθεί και να αξιολογηθεί η αναγνωστική ικανότητα παιδιών μειονοτικών ομάδων δόθηκε ένα σταθμισμένο τεστ σε 89 παιδιά Ρομά.

Οι βαθμοί τους παρατίθενται στον πίνακα:

14, 6, 8, 5, 13, 12, 3, 8, 7, 1, 3, 10, 0, 5, 13, 6, 14, 7, 11, 5, 21, 14, 11, 5, 3, 19, 31, 36, 13, 15, 13, 11, 11, 9, 2, 20, 20, 12, 10, 13, 11, 6, 10, 38, 6, 11, 5, 4, 11, 10, 22, 1, 9, 11, 18, 14, 17, 27, 30, 15, 38, 33, 37, 41, 30, 26, 28, 11, 31, 10, 4, 13, 7, 21, 23, 39, 23, 28, 21, 20, 22, 9 10, 12, 22, 38, 37, 41, 42

Ομαδοπο

Πίνακας 2.2β. Κατασκευή του πίνακα κατανομής του βαθμού AI (βήμα 2).

Βαθμός στο τεστ AI	Συχνότητα (f)
0	I
1	II
2	I
3	III
4	II
5	HH
6	IIII
7	III
8	II
9	III
10	HHI
11	HHIIII
12	III
13	HHI
14	IIII
15	II
16	
17	I
18	I
19	I
20	III

Βαθμός στο τεστ AI	Συχνότητα (f)
21	III
22	III
23	II
24	
25	
26	I
27	I
28	II
29	
30	II
31	II
32	
33	I
34	
35	
36	I
37	II
38	III
39	I
40	
41	II
42	I

Ομαδοποι

Πίνακας 2.2γ. Πίνακας κατανομής του βαθμού AI (βήμα 3).

Βαθμός στο τεστ AI	Συχνότητα (f)
0	1
1	2
2	1
3	3
4	2
5	5
6	4
7	3
8	2
9	3
10	6
11	9
12	3
13	6
14	4
15	2
16	0
17	1
18	1
19	1
20	3

Βαθμός στο τεστ AI	Συχνότητα (f)
21	3
22	3
23	2
24	0
25	0
26	1
27	1
28	2
29	0
30	2
31	2
32	0
33	1
34	0
35	0
36	1
37	2
38	3
39	1
40	0
41	2
42	1

Ομαδοποιημένες παρατηρήσεις

- Επειδή υπάρχουν πολλές τιμές στην κλίμακα μέτρησης του βαθμού στο τεστ AI, ο πίνακας συχνότητων είναι μακρύς και δύσχρηστος.
- Η εξαγωγή των πληροφοριών δεν είναι εύκολη και άμεση

Ομαδοποιημένες παρατηρήσεις

Βαθμός στο τεστ AI	Συχνότητα
[0...5)	9
[5...10)	17
[10...25)	28
[15...20)	5
[20...25)	11
[25...30)	4
[30...35)	5
[35...40)	7
[40...45)	3

Ομαδοποιημένες παρατηρήσεις

Όταν το πλήθος των τιμών μιας ποσοτικής μεταβλητής (είτε διακριτής είτε συνεχούς) είναι αρκετά μεγάλο, τότε τα δεδομένα ταξινομούνται (ομαδοποιούνται) σε μικρό πλήθος ομάδων, οι οποίες ονομάζονται **κλάσεις**.

Κάθε τιμή μπορεί να ανήκει μόνο σε μία κλάση.

Καμία παρατήρηση δεν μπορεί να μείνει έξω από κάποια κλάση.

Τα άκρα των κλάσεων ονομάζονται **όρια** των κλάσεων.

Η διαφορά του κατώτερου από το ανώτερο όριο μιας κλάσης ονομάζεται **πλάτος** της κλάσης.

Ομαδοποιημένες παρατηρήσεις

Βήμα 1: Αφαιρούμε τη μικρότερη από τη μεγαλύτερη τιμή για να βρούμε το εύρος (range) των δεδομένων. Στο παράδειγμά μας: $42 - 0 = 42$.

Βήμα 2: Διαιρούμε το εύρος με κάποιο βολικό μήκος διαστήματος (2, 3, 5, 10 κ.λπ.). Στο παράδειγμα με τον βαθμό στο τεστ AI, διαιρώντας το 42 με το 5, βρίσκουμε 8,4. Επειδή ο αριθμός των διαστημάτων πρέπει να είναι ακέραιος, στρογγυλοποιούμε προς τον αμέσως μεγαλύτερο ακέραιο αριθμό, άρα θα χρειαστούμε 9 κλάσεις μήκους 5.

Ομαδοποιημένες παρατηρήσεις

Βήμα 3: Αναγράφουμε τις κλάσεις αρχίζοντας από τη μικρότερη και καταλήγοντας στη μεγαλύτερη. Επειδή δεν επιτρέπονται επικαλύψεις μεταξύ των κλάσεων, το πάνω όριο κάθε κλάσης είναι ανοιχτό και το κάτω όριο κλειστό. Αν ήταν (0-5, 5-10 κ.λπ.) δεν θα μπορούσαμε να ξέρουμε σε ποια κλάση καταμετρήθηκε ο βαθμός 5.

Βήμα 4: Συνεχίζουμε όπως και σε έναν απλό πίνακα συχνοτήτων, καταμετρώντας δηλαδή τους βαθμούς που εμπίπτουν σε κάθε κλάση.

Εφαρμογή 2

Τα ύψη (σε cm) 40 μαθητών της Γ' Λυκείου ενός σχολείου είναι τα εξής:

168	160	164	154	169
158	162	176	165	161
163	155	164	166	172
174	164	156	159	165
166	161	167	163	167
165	172	157	175	160
168	160	164	159	167
156	163	177	155	164

Να σχηματίσετε κλάσεις της μορφής $[α, β)$ πλάτους 4 cm

Να κατασκευάσετε πίνακα κατανομής συχνοτήτων

Εφαρμογή 3

Ρωτήσαμε 50 μαθητές ενός Λυκείου πόσα παιδιά υπάρχουν στην οικογένειά τους και πήραμε τις απαντήσεις του πίνακα.

Να συμπληρώσετε τη συχνότητα της τιμής $x=3$

Αριθμός παιδιών	Συχνότητα n
1	13
2	30
3	
4	2

Σχετική συχνότητα / Σχετική εκατοστιαία συχνότητα

Η απόλυτη συχνότητα κάποιας τιμής δεν παρέχει από μόνη της μεγάλη πληροφόρηση.

Αυτό που χρειάζεται να γνωρίζουμε είναι το μέγεθος μιας συχνότητας σχετικά με το μέγεθος όλου του δείγματος.

Βαθμός	Συχνότητα (f)	
	Αγγλικά	Γερμανικά
[0, 5)	9	3
[5, 10]	31	6
Σύνολο	40	9

Σχετική συχνότητα / Σχετική εκατοστιαία συχνότητα

Αν διαιρέσουμε την (απόλυτη) συχνότητα με το μέγεθος του δείγματος προκύπτει η **σχετική συχνότητα f** .

Το άθροισμα των σχετικών συχνοτήτων ισούται με 1.

Τη σχετική συχνότητα μπορούμε να την εκφράσουμε και επί τοις εκατό οπότε ονομάζεται **εκατοστιαία σχετική συχνότητα $f\%$**

Το άθροισμα των εκατοστιαίων σχετικών συχνοτήτων ισούται με 100.

Ο αντίστοιχος πίνακας ονομάζεται **πίνακας κατανομής σχετικών συχνοτήτων**

Σχετική συχνότητα / Σχετική εκατοστιαία συχνότητα

Πίνακας 2.4α. Ομαδοποιημένος πίνακας συχνοτήτων για τους βαθμούς στα Αγγλικά και τα Γερμανικά.

Βαθμός	Συχνότητα (<i>f</i>)		Σχετική συχνότητα (<i>rf</i>)	
	Αγγλικά	Γερμανικά	Αγγλικά	Γερμανικά
[0, 5)	9	3	22,5%	33,3%
[5, 10]	31	6	77,5%	66,7%
Σύνολο	40	9	100%	100%

Εφαρμογή 4

Ρωτήσαμε 50 μαθητές ενός Λυκείου πόσα παιδιά υπάρχουν στην οικογένειά τους και πήραμε τις απαντήσεις του πίνακα.

Να κατασκευάσετε τον πίνακα κατανομής σχετικών συχνοτήτων και σχετικών εκατοστιαίων συχνοτήτων.

Αριθμός παιδιών	Συχνότητα n	Σχετική συχν. f	Σχετική συχν. $f\%$
1	13		
2	30		
3	5		
4	2		
Σύνολο	50		

Εφαρμογή 5

Για τα ύψη των μαθητών της Εφαρμογής 2 να κατασκευάσετε πίνακα κατανομής σχετικών συχνοτήτων και σχετικών εκατοστιαίων συχνοτήτων.

Ύψη	Συχνότητα n	Σχετική συχν. f	Σχετική συχν. $f\%$
[154, 158)	6		
[158, 162)	8		
[162, 166)	12		
[166, 170)	8		
[170, 174)	2		
[174, 178)	4		

Εφαρμογή 6α

Δίνεται ο παρακάτω πίνακας που αναφέρεται σε δεδομένα που έχουν ομαδοποιηθεί σε κλάσεις ίσου πλάτους.

A) Να βρείτε ποιες είναι οι κλάσεις.

Κλάσεις	Συχνότητα n	Σχετική συχν. f
[.....,)	4	
[8,)		0,2
[.....,)		
[....., 20)		
Σύνολο		

Εφαρμογή 6β

Αν γνωρίζουμε ότι:

- 12 παρατηρήσεις είναι μικρότερες του 12 και ότι
- το 70% των παρατηρήσεων είναι μικρότερες του 16.

Να συμπληρώσετε
τον πίνακα

Κλάσεις	Συχνότητα n	Σχετική συχν. f
[.....,)	4	
[8,)		0,2
[.....,)		
[....., 20)		
Σύνολο		

Εφαρμογή 6γ

- Να βρείτε πόσες παρατηρήσεις είναι μικρότερες του 16.

Κλάσεις	Συχνότητα n	Σχετική συχν. f
[.....,)	4	0,1
[8,)	8	0,2
[.....,)	16	0,4
[....., 20)	12	0,3
Σύνολο	40	1

Εφαρμογή 7

Τα αποτελέσματα των εκλογών σε ένα εκλογικό τμήμα δίνονται παρακάτω:

Πόσοι εκλογείς ψήφισαν συνολικά;

Πόσες ψήφους πήρε κάθε κόμμα;

Κόμμα	Συχνότητα ν	Σχετική συχν. f
A		0,15
B	150	0,30
Γ		0,35
Δ		
Σύνολο		

Αθροιστική συχνότητα / Αθροιστική εκατοστιαία συχν.

«Πόσοι φοιτητές απέτυχαν σε μια εξέταση;»

«Τι ποσοστό παιδιών αρχίζει να περπατά πριν την ηλικία του ενός έτους;»

Χρειάζεται να αθροιστούν οι επιμέρους συχνότητες, είτε αυτές είναι απόλυτες (όπως στην πρώτη ερώτηση), είτε είναι σχετικές (όπως στη δεύτερη).

Αθροιστική συχνότητα / Αθροιστική εκατοστιαία συχν.

Η **αθροιστική συχνότητα** μιας τιμής είναι το άθροισμα των επιμέρους συχνοτήτων (απόλυτων ή σχετικών) των τιμών, οι οποίες είναι *μικρότερες* ή *ίσες* με την εν λόγω τιμή.

Προσοχή: Αθροίζουμε τις συχνότητες, όχι τις τιμές της μεταβλητής.

Η απόλυτη αθροιστική συχνότητα συμβολίζεται με **N**, ενώ η σχετική αθροιστική συχνότητα με **F**.

Αθροιστική συχνότητα / Αθροιστική εκατοστιαία συχν.

Πόσοι
φοιτητές
απέτυχαν;;

Βαθμός	Επιμέρους συχνότητα		Αθροιστική συχνότητα	
	Απόλυτη ν	Σχετική (%) f %	Απόλυτη N	Σχετική (%) F %
1	1	2,50	1	2,50
2	3	7,50	4	10,00
3	4	10,00	8	20,00
4	1	2,50	9	22,50
5	5	12,50	14	35,00
6	6	15,50	20	50,00
7	8	20,00	28	70,00
8	5	12,50	33	82,50
9	4	10,00	37	92,50
10	3	7,50	40	100,00

Αθροιστική συχνότητα / Αθροιστική εκατοστιαία συχν.

- Η αθροιστική συχνότητα της μικρότερης τιμής είναι ίση με την αντίστοιχη απόλυτη συχνότητα.
- Η αθροιστική συχνότητα της μεγαλύτερης τιμής είναι ίση με το μέγεθος του δείγματος.
- Η σχετική αθροιστική συχνότητα της μεγαλύτερης τιμής είναι ίση με 100%.
- **Αθροιστικές συχνότητες δεν έχουν νόημα για μεταβλητές με ονομαστική κλίμακα μέτρησης**

Εφαρμογή 8

Ο παρακάτω πίνακας δείχνει την κατανομή του βαθμού σε ένα τεστ.

α) Συμπληρώστε τα κενά του πίνακα.

β) Πόσους μαθητές έχει η τάξη;

γ) Τι ποσοστό μαθητών έγραψε πάνω από τη βάση;

Βαθμός	Μέση τιμή	Απλή συχνότητα		Αθροιστική συχνότητα	
		Απόλυτη n	Σχετική $f\%$	Απόλυτη N	Σχετική $F\%$
[1, 6)			10	5	
[6, 11)		13			
[11, 16)		17			
[16, 20]					100

Εφαρμογή 8

Βαθμός	Μέση τιμή	Απλή συχνότητα		Αθροιστική συχνότητα	
		Απόλυτη n	Σχετική $f\%$	Απόλυτη N	Σχετική $F\%$
[1, 6)			10	5	
[6, 11)		13			
[11, 16)		17			
[16, 20]					100

Άσκηση για το σπίτι

Τριάντα φοιτητές ρωτήθηκαν πόσες ώρες μελέτησαν το περασμένο Σαββατοκύριακο. Οι απαντήσεις που έδωσαν είναι οι εξής:

4, 10, 2, 0, 3, 5, 6, 10, 7, 1, 8, 2, 4, 7, 8, 10, 7, 8, 9, 10, 7, 1, 7, 10, 7, 3, 4, 8, 5, 3.

A) Να σχηματίσετε 5 κλάσεις ίσου πλάτους της μορφής $[α, β)$

B) Να συμπληρώσετε τον ομαδοποιημένο πίνακα κατανομής αυτών των δεδομένων.

Ώρες μελέτης	Συχνότητα	Αθροιστική συχνότητα	Σχετική αθροιστική συχνότητα

Ύψη	Συχνότητα ν	Σχετική συχν. f	Σχετική συχν. f%
[154, 158)	6		
[158, 162)	8		
[162, 166)	12		
[166, 170)	8		
[170, 174)	2		
[174, 178)	4		

Κλάσεις	Συχνότητα n	Σχετική συχν. f
[.....,)	4	
[8,)		0,2
[.....,)		
[....., 20)		
Σύνολο		

Αριθμός παιδιών	Συχνότητα n	Σχετική συχν. f	Σχετική συχν. $f\%$
1	13		
2	30		
3	5		
4	2		
Σύνολο	50		