1/3/17

Ορισμός: Στατιστική είναι ο επιστημονικός κλάδος που ασχολείται με τη σχεδίαση πειραμάτων και μεθόδων δειγματοληψίας, τη συλλογή και ανάλυση αριθμητικών δεδομένων (μετρήσεων) και την εξαγωγή συμπερασμάτων (σχετικά με κάποιο χαρακτηριστικό - φαινόμενο που είναι άγνωστο και ενδιαφερόμαστε να γνωρίσουμε) για ένα σύνολο με βάση τις πληροφορίες που περιέχονται σε ένα δείγμα (μέρος) του συνόλου αυτού.

Θεωρητικό υπόβαθρο: Θεωρία Πιθανοτήτων
Εφαρμογή: Σε όλες τις επιστήμες

Πληθυσμός: Το σύνολο των τιμών μιας τυχαίας μεταβλητής που μας ενδιαφέρει να μελετήσουμε.

Δείγμα: Μέρος του πληθυσμού που εκλέγεται με στόχο να είναι αντιπροσωπευτικό του πληθυσμού

Δειγματοληψία: Η μέθοδος της επιλογής του δείγματος

Τυχαίο Δείγμα: Εκλέγεται έτσι ώστε όλα τα μέλη του πληθυσμού να έχουν ίση και ανεξάρτητη πιθανότητα να συμπεριληφθούν στο δείγμα. (Κλήρωση, Η/Υ)

Τυχαίο Δείγμα: n ανεξάρτητες κ' ισότιμες μετρήσεις του πληθυσμού (X) τ.μ. $X_1, \dots, X_n$ n: (n) τιμές $X_1, \dots, X_n$ της X του πληθυσμού

Συστηματική Δειγματοληψία: Για n από N μέλη

$k = \frac{N}{n}$, τ. επιλέγω ένα λ, $1 \leq \lambda \leq k$

Τυχαίο Δείγμα: $\lambda, \lambda + k, \ldots, \lambda + (n-1) \cdot k$

π.χ. Τμήμα Μαθηματικών

$N = 1.600$ $\quad k = \frac{1.600}{40} = 40$

$n = 40$

Έστω $\lambda = 25$

$25, 65, \ldots, 1585$

Δειγματοληψία κατά στρώματα (n από N μέλη)

Έστω k υποπληθυσμοί (ξένοι μεταξύ τους) με N_i, $i = 1, \ldots, k$ μέλη ο καθένας κ $N_1 + \ldots + N_k = N$

Επιλέγουμε τυχαία n_i, $i = 1, \ldots, k$ μέλη από τον i-οστό πληθυσμό με $n_1 + \ldots + n_k = n$ και $\frac{N_1}{n_1} = \ldots = \frac{N_k}{n_k} = \frac{N}{n}$

Σχολή Θετικών Επιστημών (Π.Ι)

Τμήμα	Μαθ.	Φυσ.	Χημ.	Πληρ.	Σύνολο
Μέγεθος Στρώματος	1.600	1000	800	600	$4000 = N$
Μέγεθος Δείγματος	$n_1 = 24$	$n_2 = 15$	12	9	$60 = n$

$n = \frac{60}{4000} = 0,015$

$n_1 = N_1 * 0,0,15 = 24$

Περιγραφική Στατιστική

Έστω $x_1, \ldots, x_n$ δείγμα από πληθυσμό X τα x_i δεν είναι κατ' ανάγκη διάφορα μεταξύ τους. Ο αριθμός των φορών που το x_i εμφανίζεται στο δείγμα ονομάζεται συχνότητα της τιμής: f_i

Σχετική Συχνότητα: f_i/n
(Για k διαφορετικές τιμές) $\sum_{i=1}^{k} f_i = n$, $\sum_{i=1}^{k} \frac{f_i}{n} = 1$

Το άθροισμα των συχνοτήτων f_j, για $x_j \leq x_i$ λέγεται αθροιστική συχνότητα: F_i

Αθροιστική σχετική συχνότητα: F_i/n

Παράδειγμα 1

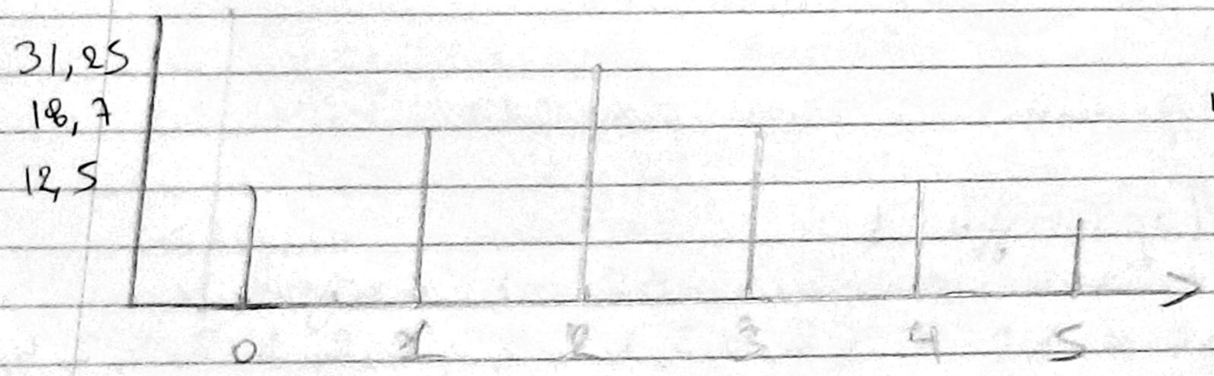
Αριθμός παιδιών 32 $(=n)$ οικογενειών
1, 0, 2, 1, 0, 2, 3, 2, 4, 3, 1, 2, 5, 3, 2, 2, 1, 2, 3, 2, 4, 1, 3, 0, 4, 1, 4, 2, 5, 0, 2, 3

Διατεταγμένος Πίνακας Συχνοτήτων

Τιμή x_i Αριθμός Παιδιών	Συχ. f_i	Σχ. Συχ. f_i/n	Αθρ. Συχ. F_i	F_i/n
0	4	4/32 = 0,125	4	4/32
1	6	6/32 = 18,75%	10	10/32
2	10	10/32 = 31,25%	20	20/32
3	6	6/32 = 18,75%	26	26/32
4	4	4/32 = 12,5%	30	30/32
5	2	2/32 = 6,75%	32	1
	$n = 32$	1		

Παράδειγμα (ομαδοποιημένα)

Τιμή x_i (αρ. παιδιών)	Συχνότητα f_i	Σχ. Συχ. f_i/n	Αθρ. Συχ. F_i
0	4	4/32 = 12,5%	4
1	6	6/32 = 18,75%	10
2	10	10/32 = 31,25%	20
3	6	6/32 = 18,75%	26
4	4	4/32 = 12,5%	30
5	2	2/32 = 6,25%	32 = n



n = 32

$R = \max x_i - \min x_i$ $(= 355 - 161 = 194)$

k ομάδες , $5 \le k \le 30$, $k = 1 + 3{,}32 \log_{10} n$

$(= 1 + 3{,}32 \log_{10} 60 = 6{,}9 \approx 7)$

k = 8

μήκος $d = \frac{R}{k} = \frac{194}{8} = 24{,}25 \approx 25$

161 355

[160,5 ,]

+25

Ομάδα i	Όρια [L, U]	x_i	Συχν f_i	Σχ.Συχ. $f_i/ν$
1	[160,5 , 18,5]	175	8	5/60 = 0,1333
2	[185,5 , 210,5]	198	11	11/60
3	[210,5 , 235,5]	223	13	10/60
4	[235,5 , 262,5]	248	10	6/60
5	[260,5 , 285,5]	273	6	7/60
6	[285,5 , 310,5]	298	7	3/60
7	[310,5 , 335,5]	323	3	2/60
8	[335,5 , 365,5]	348	0 ν=60	1

Αθρ. Συχ. F_i	Αθρ. Σχ. Συχ. $F_i/ν$
8	8/60
19	19/60
32	32/60 = 0,533
42	42/60
48	48/60
55	55/60
50	8/60
60 = 1	1

Από φυλλάδιο (π.χ. πολύγωνο συχνοτήτων)