

ΟΡΙΣΜΟΙ ΣΤΑΤΙΣΤΙΚΗΣ

Πληθυσμός :

Καλείται το σύνολο όλων των ατόμων, για τα οποία θέλουμε να βγάλουμε συμπέρασμα, από την έρευνα που κάνουμε. (Κάθε στοιχείο του πληθυσμού το ονομάζουμε άτομο).

Δείγμα:

Καλούμε τα συγκεκριμένα άτομα του πληθυσμού που χρησιμοποιούμε στην έρευνα μας (ποιά άτομα! Όχι πόσα άτομα).

Προσοχή! Για να είναι το δείγμα αξιόπιστο θα πρέπει για τα άτομα που χρησιμοποιούμε να γνωρίζουμε τα πάντα, εκτός από την ιδιότητα ως προς την οποία τα μελετάμε και **να είναι άτομα επιλεγμένα** , ύστερα από μελέτη που θα κάνουμε και όχι τυχαία .

Πλήθος δείγματος ή μέγεθος δείγματος:

Καλείται ο αριθμός των ατόμων του πληθυσμού που χρησιμοποιούμε στην έρευνα που κάνουμε , (**πόσα άτομα ; Όχι ποιά άτομα !**) .

Παρατηρήσεις (ti) :

Καλούμε τις τιμές των μεταβλητών όλων των ατόμων του δείγματος γράφοντας την καθεμία τιμή της μεταβλητής τόσες φορές όσες φορές εμφανίζεται.

Μεταβλητή: (X)

Καλείται η ιδιότητα ως προς την οποία εξετάζουμε ένα δείγμα, για να βγάλουμε συμπέρασμα για όλα τα άτομα του πληθυσμού.

Ποσοτικές μεταβλητές:

Καλούνται αυτές που έχουν τιμές που εκφράζουν ποσότητα και μπορούν να προστεθούν (συνήθως χρησιμοποιούμε αριθμητικά σύμβολα ή λέξεις που κρύβουν ποσότητα).

Ποιοτικές μεταβλητές:

Καλούνται αυτές που έχουν τιμές που εκφράζουν κάποιο μη ποσοτικό χαρακτηριστικό και δεν μπορούμε να τις προσθέσουμε (συνήθως χρησιμοποιούμε λέξεις ή αριθμούς, π.χ. κωδικούς που ορίζουν ποιότητα).

Τιμές της μεταβλητής: (Xi)

Καλούνται οι διαφορετικές τιμές που μπορεί να πάρει η ιδιότητα (μεταβλητή X).. ως προς την οποία εξετάζουμε το δείγμα.

Συχνότητα μιας τιμής της μεταβλητής: (vi)

Καλείται ο αριθμός που μας λέει **πόσα άτομα** του δείγματος εμφανίζουν αυτή την τιμή της μεταβλητής.

Σχετική συχνότητα μιας τιμής της μεταβλητής: (f)

Καλείται ο αριθμός (δεκαδικός ≤ 1) που μας λέει **τι μέρος** (τι **κλάσμα**;) του δείγματος εμφανίζει αυτήν την τιμή της μεταβλητής.

Σχετική συχνότητα τοις εκατό μιας τιμής της μεταβλητής: (f%)

Καλείται ο αριθμός (**ποσοστό ≤ 100**) που μας λέει τι ποσοστό του δείγματος εμφανίζει αυτή την τιμή της μεταβλητής.

Αθροιστική συχνότητα μιας τιμής της μεταβλητής: (Ni)

Καλούμε τον αριθμό που μας λέει **πόσα άτομα** του δείγματος έχουν **ΜΕΧΡΙ ΚΑΙ** αυτήν την τιμή της μεταβλητής, (δηλαδή: έχουν μικρότερη ή ίση από αυτή την τιμή της μεταβλητής).

Αθροιστική σχετική συχνότητα μιας τιμής της μεταβλητής: (Fi)

Καλείται ο αριθμός (δεκαδικός ≤ 1) που μας λέει **τι μέρος (τι κλάσμα)** του δείγματος εμφανίζει **ΜΕΧΡΙ ΚΑΙ** αυτή την τιμή της μεταβλητής, (δηλαδή: **τι μέρος (τι κλάσμα)** του δείγματος έχει μικρότερη ή ίση από αυτή την τιμή της μεταβλητής).

Αθροιστική σχετική συχνότητα τοις εκατό μιας τιμής της μεταβλητής: (Fi%)

Καλείται ο αριθμός (**ποσοστό ≤ 100**) που μας λέει **τι ποσοστό** του δείγματος εμφανίζει **μέχρι και** αυτή την τιμή της μεταβλητής, (δηλαδή: **τι ποσοστό του δείγματος** έχει μικρότερη ή ίση από αυτή την τιμή της μεταβλητής)

ΠΑΡΑΜΕΤΡΟΙ ΘΕΣΗΣ

Επικρατούσα τιμή:

Καλείται ο η τιμή της μεταβλητής με την μεγαλύτερη συχνότητα, δηλαδή η τιμή της μεταβλητής που εμφανίστηκε περισσότερες φορές. Μπορώ να έχω περισσότερες από μία επικρατούσες τιμές και τότε η κατανομή θα λέγεται δικόρυφη ,τρικόρυφη κλπ.

Μέση τιμή ($\bar{X}$):

Καλούμε το μέσο όρο των παρατηρήσεων και τον βρίσκουμε από τους τύπους:

$$1) \bar{x} = \frac{X_1 V_1 + \dots + X_k V_k}{V}$$

$$2) \bar{X} = \frac{t_1 + \dots + t_v}{v}$$

$$3) \bar{x} = \chi_1 f_1 + \dots + \chi_k f_k$$

Διάμεσος (δ) :

- 1) Τοποθετώ όλες τις παρατηρήσεις n του δείγματος σε **αύξουσα σειρά**.
- 2) Υπολογίζω το $\frac{n+1}{2}$
- 3) Η διάμεσος (δ) ισούται :
 - i. με τη **μεσαία παρατήρηση** όταν το πλήθος του δείγματος είναι **περιττό** ή
 - ii. με το **μέσο όρο των δύο μεσαίων παρατηρήσεων** όταν το πλήθος του δείγματος είναι **άρτιο**

ΜΕΘΟΔΟΣ

Αν n περιττός: Τότε το $\frac{n+1}{2} = \lambda$ είναι φυσικός αριθμός και τότε έχω μία μεσαία παρατήρηση την (t_λ) και επομένως **$\delta = t_\lambda$**

Αν n άρτιος: Τότε το $\frac{n+1}{2} = \lambda,5$ είναι δεκαδικός αριθμός και τότε έχω δύο μεσαίες παρατηρήσεις την (t_λ) και την ($t_{\lambda+1}$) και επομένως **$\delta = \frac{t_\lambda + t_{\lambda+1}}{2}$**

ΠΑΡΑΜΕΤΡΟΙ ΔΙΑΣΠΟΡΑΣ

Εύρος μεταβολής:

Καλείται η απόσταση της μικρότερης από την μεγαλύτερη παρατήρηση.

Αν έχω γράψει τις κ τιμές της μεταβλητής κατά αύξουσα σειρά τότε:

$$\text{εύρος μεταβολής} = \chi_k - \chi_1 \quad , \quad \kappa: \text{πλήθος τιμών της μεταβλητής}$$

Διακύμανση S^2 :

καλούμε τον αριθμό που μας λέει πόσο κατά μέσο όρο απέχουν τα τετράγωνα της απόστασης, των παρατηρήσεων από τη μέση τιμή.

$$S^2 = \frac{V1(x_{1-\bar{x}})^2 + \dots + V\kappa(x_{\kappa-\bar{x}})^2}{V}$$

$$S^2 = \frac{(\tau_{1-\bar{x}})^2 + \dots + (t_{\nu-\bar{x}})^2}{V}$$

Τυπική απόκλιση S:

καλείται η τετραγωνική ρίζα της διακύμανση και υπολογίζεται από τους τύπους:

$$1) \quad S = \sqrt{S^2}$$

$$2) \quad S = \sqrt{\frac{V1(x_{1-\bar{x}})^2 + \dots + V\kappa(x_{\kappa-\bar{x}})^2}{V}}$$

$$3) \quad S = \sqrt{\frac{(\tau_{1-\bar{x}})^2 + \dots + (t_{\nu-\bar{x}})^2}{V}}$$

ΠΡΟΣΟΧΗ!!

Οι μονάδες της διακύμανσης γίνονται τετραγωνικές των παρατηρήσεων. Δηλαδή αν οι παρατηρήσεις είναι σε km η διακύμανση είναι σε km², ενώ όλων των άλλων παραμέτρων οι μονάδες είναι ίδιες με των παρατηρήσεων.

CV

$$CV = \frac{\text{τυπική απόκλιση}}{\text{μέση τιμή}} = \frac{S}{\bar{X}} \quad \text{Το γράφουμε σε ποσοστό τοις \% .}$$

Ισχύει: $CV = \frac{S}{\bar{X}} = 100 \cdot \frac{S}{\bar{X}} \cdot \%$

και είναι καθαρός αριθμός χωρίς μονάδες!

Αν $CV < 0,10$ ή 10% ΕΙΝΑΙ ΟΜΟΙΟΓΕΝΗΣ.

Αν $CV \geq 0,10$ ή 10% ΔΕΝ ΕΙΝΑΙ ΟΜΟΙΟΓΕΝΗΣ.

Προσοχή !!: Άλλο σημαίνει το να είναι μια κατανομή **περισσότερο ομοιογενής** από μια άλλη και άλλο σημαίνει αν είναι ομοιογενής ή όχι.

Περισσότερο ομοιογενής: Μία κατανομή είναι περισσότερο ομοιογενής από μία άλλη όταν έχει μικρότερο CV από αυτήν.

Xi ΤΙΜΗ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ	Vi ΑΠΟΛΥΤΗ ΣΥΧΝΟΤΗΤΑ ΜΙΑΣ ΤΙΜΗΣ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ	f ΣΧΕΤΙΚΗ ΣΥΧΝΟΤΗΤΑ ΜΙΑΣ ΤΙΜΗΣ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ	f % ΣΧΕΤΙΚΗ ΣΥΧΝΟΤΗΤΑ ΜΙΑΣ ΤΙΜΗΣ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ ΤΟΙΣ ΕΚΑΤΟ	Xi vi ΥΠΟΛΟΓΙΖΕΙ ΜΕΣΗ ΤΙΜΗ	Ni ΑΘΡΟΙΣΤΙΚΗ ΣΥΧΝΟΤΗΤΑ ΜΙΑΣ ΤΙΜΗΣ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ Σ	Fi ΑΘΡΟΙΣΤΙΚΗ ΣΧΕΤΙΚΗ ΣΥΧΝΟΤΗΤΑ ΜΙΑΣ ΤΙΜΗΣ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ Σ	Fi% ΑΘΡΟΙΣΤΙΚΗ ΣΧΕΤΙΚΗ ΣΥΧΝΟΤΗΤΑ ΜΙΑΣ ΤΙΜΗΣ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ Σ ΤΟΙΣ ΕΚΑΤΟ %	Vi(xi-$\bar{x}$)² ΥΠΟΛΟΓΙΖΕΙ ΔΙΑΚΥΜΑΝΣΗ
-------------------------------------	--	--	--	---	--	---	--	--

Xi	Vi	f	f %	Xi vi	Ni	Fi	Fi%	Vi(xi- $\bar{x}$) ²
1	3	3/25=0.12	12	3	3	0.12	12	6.2208
2	12	12/25=0,48	48	24	15	0.60	60	2.3232
3	7	7/25=0,28	28	21	22	0.88	88	2.1952
4	2	2/25=0,08	8	8	24	0.96	96	4.8672
5	1	1/25=0,04	4	5	25	1.00	100	6.5536
άθροισμα	V=25	=1	=100	=61	----	----	-----	22,16

ΥΠΟΛΟΓΙΣΜΟΣ $\sum V_i (X_i - \bar{x})^2$

Βρίσκουμε τη μέση τιμή $\bar{x}$ (μέσο όρο)
1. βρίσκουμε πόσο απέχει το X_1 από τη μέση τιμή $\bar{x}$
2. (με αφαίρεση μικρότερου από μεγαλύτερο)
3. τον αριθμό που θα βρούμε τον πολλαπλασιάζουμε με τον εαυτό του
4. και μετά με τον αριθμό v_1
5. το αποτέλεσμα που βρίσκουμε το θέτουμε στον πίνακα.
Με τον ίδιο αντίστοιχο τρόπο βρίσκουμε και τα υπόλοιπα $Vi(xi-\bar{x})^2$.

ΥΠΟΛΟΓΙΣΜΟΣ ΜΕΣΗΣ ΤΙΜΗΣ

$$\bar{x} = \frac{X_1 V_1 + \dots + X_k V_k}{V} = \frac{61}{25} = 2.44$$

k : πλήθος τιμών της μεταβλητής

$$\bar{X} = \frac{t_1 + \dots + t_v}{v} = \frac{61}{25} = 2.44$$

v : πλήθος δείγματος

$$\bar{x} = X_1 f_1 + \dots + X_k f_k = 1 \cdot 0.12 + 2 \cdot 0.48 + 3 \cdot 0.28 + 4 \cdot 0.08 + 5 \cdot 0.04 = 2.44$$

k : πλήθος τιμών της μεταβλητής

ΟΛΟΙ ΟΙ ΤΥΠΟΙ:

1.	$V = V_1 + \dots + V_k$ k=πλήθος τιμών της μεταβλητής.
2.	$f_1 + \dots + f_k = 1$
3.	$f_i = \frac{V_i}{V} \quad 1 \leq i \leq k$
4.	$f_i \% = 100 \cdot f_i$
5.	$f_1 \% + \dots + f_k \% = 100$
6.	$N_i = V_1 + \dots + V_i \quad 1 \leq i \leq k$
7.	$F_i = f_1 + \dots + f_i \quad 1 \leq i \leq k$
8.	$F_i = \frac{N_i}{V} \quad 1 \leq i \leq k$
9.	$F_i \% = 100 \cdot F_i$
10.	Μέση τιμή $\bar{x} = \frac{x_1 V_1 + \dots + x_k V_k}{V}$ $\bar{x} = x_1 f_1 + \dots + x_k f_k$ $\bar{X} = \frac{t_1 + \dots + t_v}{v}$

11.	Διακύμανση $S^2 = \frac{V_1(x_1 - \bar{x})^2 + \dots + V_k(x_k - \bar{x})^2}{V}$ k=πλήθος τιμών της μεταβλητής. $S^2 = \frac{(\tau_1 - \bar{x})^2 + \dots + (t_v - \bar{x})^2}{V}$
12.	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ $S = \sqrt{S^2}$ $S = \sqrt{\frac{V_1(x_1 - \bar{x})^2 + \dots + V_k(x_k - \bar{x})^2}{V}}$ $S = \sqrt{\frac{(\tau_1 - \bar{x})^2 + \dots + (t_v - \bar{x})^2}{V}}$
13.	ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ $CV = \frac{\text{τυπική απόκλιση}}{\text{μέση τιμή}} = \frac{S}{\bar{x}}$

Ερώτημα: Πότε $k=v$;

Απάντηση: Όταν $V_i = 1$ για κάθε $1 \leq i \leq v$

ΟΜΑΔΟΠΟΙΗΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

Όταν οι παρατηρήσεις είναι σχεδόν όλες διαφορετικές ,

θεωρώ **σαν τιμές της μεταβλητής** διαστήματα (κλάσεις) , π.χ :[150 έως 160).

Θεωρώ **σαν παρατηρήσεις κάθε κλάσης** , όλες τις παρατηρήσεις που είναι ενδιάμεσα της.

ΠΩΣ ΔΗΜΙΟΥΡΓΩ ΚΛΑΣΕΙΣ v ΠΑΡΑΤΗΡΗΣΕΩΝ ;

Βήμα 1=⁰:

Βρίσκω ένα διάστημα στρογγυλοποιημένο (συνήθως σε 0 ή σε 5) το οποίο περιέχει όλες τις παρατηρήσεις.

Βήμα=2=⁰:

Βρίσκω το πλάτος του διαστήματος, π.χ Το διάστημα (α , β) έχει πλάτος = (β-α) .

Βήμα 3=⁰:

Επιλέγω τον αριθμό κλάσεων (έστω κ) που θέλω να χωρίσω το δείγμα και από αυτό βρίσκω το πλάτος κάθε κλάσης.

(πλάτος κλάσης (α , β) = $\frac{\beta - \alpha}{\kappa}$).

Διάστημα:

καλείται ένα κομμάτι της ευθείας των πραγματικών αριθμών και περιέχει τους άπειρους αριθμούς ,που βρίσκονται ανάμεσα στα 2 άκρα του, αν και τα άκρα ανήκουν ή 'όχι σ' αυτό εξαρτάται απ' την μορφή τους (**< ανοιχτό >**) ή [**≤ κλειστό ≥**].

Κέντρο κλάσης ki :

Στις ομαδοποιημένες κατανομές, (δηλαδή παρατηρήσεις που παίρνω σε ομάδες) ,για τον προσδιορισμό των παραμέτρων, **η κάθε κλάση αντιπροσωπεύεται από το κέντρο της.**

Και κέντρο π.χ του (α,β) είναι το $\frac{\alpha+\beta}{2}$ δηλαδή αντί για x_i χρησιμοποιώ k_i .

ΠΑΡΑΜΕΤΡΟΙ ΟΜΑΔΟΠΟΙΗΜΕΝΩΝ ΚΑΤΑΝΟΜΩΝ

Στις ομαδοποιημένες κατανομές δηλαδή σ' αυτές που τα x_i (τιμές των μεταβλητών) είναι διαστήματα.

Υπάρχει επιπλέον και η επικρατούσα κλάση,

Επικρατούσα κλάση

Καλούμε την κλάση με τη μεγαλύτερη συχνότητα, (πιο ψηλή) ,

μπορώ να έχω περισσότερες από 1 επικρατούσες κλάσεις π.χ. (155,165) και (175,185) , τότε η κατανομή θα λέγεται δικόρυφη , τρικόρυφη κλπ.

Ενώ η διάμεσος η επικρατούσα τιμή και το μεγαλύτερο ή μικρότερο από ενδιάμεσο αριθμό μιας κλάσης προσδιορίζεται γραφικά.

ΙΣΤΟΓΡΑΜΜΑΤΑ ΣΥΧΝΟΤΗΤΩΝ

ΙΣΤΟΓΡΑΜΜΑ ΣΥΧΝΟΤΗΤΩΝ , ΣΧΕΤΙΚΩΝ ΣΥΧΝΟΤΗΤΩΝ ΚΑΙ ΣΧΕΤΙΚΩΝ ΣΥΧΝΟΤΗΤΩΝ ΤΟΙΣ ΕΚΑΤΟ:

Χρησιμοποιώ 2 παραπάνω κλάσεις, την προηγούμενη και την επόμενη του διαστήματος που περιέχει τις παρατηρήσεις και θεωρώ ότι έχουν συχνότητα μηδέν.

Πολύγωνο συχνοτήτων:

Ενώνω το μέσο της προηγούμενης κλάσης, με το μέσο της πάνω βάσης του ορθογωνίου, και κλείνω με το μέσο της επόμενης κλάσης.

Όλο μαζί ονομάζεται διάγραμμα συχνοτήτων ή γραφική παράσταση.

ΙΣΤΟΓΡΑΜΜΑ ΑΘΡΟΙΣΤΙΚΩΝ ΣΥΧΝΟΤΗΤΩΝ , ΑΘΡΟΙΣΤΙΚΩΝ ΣΧΕΤΙΚΩΝ ΣΥΧΝΟΤΗΤΩΝ ΚΑΙ ΑΘΡΟΙΣΤΙΚΩΝ ΣΧΕΤΙΚΩΝ ΣΥΧΝΟΤΗΤΩΝ ΤΟΙΣ ΕΚΑΤΟ:

Δε χρησιμοποιούμε προηγούμενη και επόμενη κλάση, ενώνουμε την αρχή των αξόνων με την πάνω δεξιά κορυφή των ορθογωνίων.

Διάμεσος σε ομαδοποιημένες παρατηρήσεις:

Αν έχω μόνο πίνακα δηλαδή δεν ξέρω τις παρατηρήσεις τότε η διάμεσος βρίσκεται **ΜΟΝΟ** γραφικά από οποιοδήποτε ιστόγραμμα αθροιστικών συχνοτήτων.

ΠΡΟΣΟΧΗ!! Αν μας δίνονται οι παρατηρήσεις και εμείς κατασκευάσουμε τον πίνακα , τότε μπορώ τη διάμεσο, τη μέση τιμή ,τη διακύμανση ,τη τυπική απόκλιση και το εύρος μεταβολής να τα προσδιορίσω όπως και στις διακριτές μεταβλητές (με τύπους).

ΣΕ ΟΜΑΔΟΠΟΙΗΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ: Μπορώ να βρω τη διάμεσο γραφικά, από οποιοδήποτε ιστόγραμμα αθροιστικών συχνοτήτων ή ιστόγραμμα αθροιστικών σχετικών συχνοτήτων

ΓΡΑΦΙΚΟΣ ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΕΠΙΚΡΑΤΟΥΣΑΣ ΤΙΜΗΣ ΟΜΑΔΟΠΟΙΗΜΕΝΗΣ ΚΑΤΑΝΟΜΗΣ

1. Ενώνω την πάνω αριστερή κορυφή του ορθογωνίου της επικρατούσας κλάσης με την αριστερή κορυφή της διπλανής επόμενης κλάσης
2. και την **πάνω δεξιά κορυφή** του ορθογωνίου της επικρατούσας κλάσης, **χιαστί**, με την **δεξιά κορυφή της διπλανής** προηγούμενης κλάσης.
3. Το σημείο τομής του χιαστί το προβάλλω στον άξονα των κλάσεων ($X'X$). Ο αριθμός που αντιστοιχεί στην προβολή στον άξονα ($X'X$) αποτελεί την επικρατούσα τιμή.

.....
τηλ. 697 369 2006 x11k10@sch.gr

Μαθηματικός: Χριστόφορος Κατσίμπρας