ΑΓΓΕΛΙΚΗ ΒΟΥΔΟΥΡΗ

ΜΑΘΗΜΑΤΑ ΠΕΡΙΓΡΑΦΙΚΗΣ ΣΤΑΤΙΣΤΙΚΗΣ

- I -

ΜΟΝΟΜΕΤΑΒΛΗΤΕΣ ΚΑΤΑΝΟΜΕΣ
ΣΥΧΝΟΤΗΤΑΣ
ΜΕΤΡΑ ΘΕΣΗΣ - ΔΙΑΣΠΟΡΑΣ
ΑΣΥΜΜΕΤΡΙΑΣ - ΚΥΡΤΟΤΗΤΑΣ
ΡΟΠΕΣ

ΑΘΗΝΑ 1991

ΕΚΤΥΠΩΣΗ: Σ. ΑΘΑΝΑΣΟΠΟΥΛΟΣ - Σ. ΠΑΠΑΔΑΜΗΣ & ΣΙΑ Ε.Ε.
ΙΩΑΝ. ΘΕΟΛΟΓΟΥ 80 ΖΩΓΡΑΦΟΥ, ΤΗΛ. 77.10.548 - 77.02.033, FAX: 77.10.581

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΕΝΝΟΙΕΣ

[1.1] Στατιστική

Η στατιστική ορίζεται σήμερα ως η επιστήμη η οποία ασχολείται με

[1] συλλογή.

[2] επεξεργασία.

[3] οργάνωση.

[4] παρουσίαση.

[5] ανάλυση.

αριθμητικών στοιχείων, που αναφέρονται σε χαρακτηριστικά και ιδιότητες διαφόρων φαινομένων, με σκοπό τη συστηματική μελέτη των φαινομένων αυτών και την εξαγωγή συμπερασμάτων χρήσιμων στη διαδικασία λήψης αποφάσεων.

[1.2] Στατιστικός Πληθυσμός

Ένα σύνολο στοιχείων των οποίων μία ή περισσότερες ιδιότητες αποτελούν αντικείμενο μελέτης με τη χρήση στατιστικών μεθόδων λέγεται στατιστικός πληθυσμός.

[1.2.1] Παραδείγματα Στατιστικών Πληθυσμών

Το σύνολο των κατοικιών μιας πόλης.

Το σύνολο των εργατών μιας επιχείρησης.

Το σύνολο των φοιτητών ενός πανεπιστημίου.

Το σύνολο των αυτοκινήτων μιας πόλης.

Το σύνολο των νοσοκομείων μιας χώρας.

Το σύνολο των δημοτικών σχολείων μιας χώρας.

[1.3] Στατιστική Μονάδα.

Στατιστική μονάδα λέγεται ένα στοιχείο ενός στατιστικού πληθυσμού. Η στατιστική ενδιαφέρεται για τη μελέτη ιδιοτήτων ή χαρακτηριστικών των στατιστικών μονάδων.

[1.3.1] Παραδείγματα Στατιστικών Μονάδων.

Κατοικία μιας πόλης.

Εργάτης μιας επιχείρησης

Φοιτητής ενός πανεπιστημίου.

Αυτοκίνητο μιας πόλης.

Νοσοκομείο μιας χώρας.

Δημοτικό σχολείο μιας χώρας.

[1.4] Στατιστική Μεταβλητή

Μιά ιδιότητα ή ένα χαρακτηριστικό των στοιχείων ενός στατιστικού πληθυσμού λέγεται στατιστική μεταβλητή.

[1.4.1] Παραδείγματα Στατιστικών Μεταβλητών.

Μεταβλητές	Τιμές
Αριθμός δωματίων κατοικίας	1,2,3,...,
Ετήσιο εισόδημα ατόμου	Κάθε θετικός πραγματικός αριθμός.
Κατάσταση υγείας ατόμου	Κακή, μέτρια, καλή, πολύ κακή, άριστη.
Φύλο ατόμου	Αρσενικό, θηλυκό.
Οικογενειακή κατάσταση ατόμου	Άγαμος, έγγαμος, χήρος, διαζευγμένος.

Ημερήσια τροχαία ατυχήματα μιας πόλης.	0,1,2,3,...,
Ετήσια βροχόπτωση μιας πόλης.	Κάθε θετικός πραγματικός αριθμός.
Μόλυνση περιβάλλοντος μιας περιοχής.	Ασήμαντη, μικρή, μέτρια, μεγάλη, πολύ μεγάλη.

[1.5] Ταξινόμηση Στατιστικών Μεταβλητών.

Ανάλογα με την μέτρηση που δέχονται οι διάφορες στατιστικές μεταβλητές διακρίνονται σε

[1] Κατηγορικές.

[2] Ποιοτικές.

[3] Ποσοτικές.

[1.6] Κατηγορικές Μεταβλητές.

Οι κατηγορικές μεταβλητές επιτρέπουν μόνο την κατάταξη των μονάδων ενός πληθυσμού σε κατηγορίες. Οι "τιμές" των μεταβλητών αυτών εκφράζονται με λέξεις ή με άλλα σύμβολα [αριθμητικά ή μη αριθμητικά]. Σχέσεις μεταξύ των "τιμών" των κατηγορικών μεταβλητών δεν υπάρχουν,. Ακόμη οι τιμές των μεταβλητών αυτών δεν δέχονται καμμία ιεράρχηση δηλαδή κατάταξη σε αύξουσα ή φθίνουσα τάξη μεγέθους.

[1.6.1] Παραδείγματα Κατηγοριών Μεταβλητών.

Μεταβλητές	Τιμές
Φύλο	Αρσενικό, θηλυκό.
Οικογενειακή κατάσταση	Άγαμος, έγγαμος, χήρος.
Θρήσκευμα	Μωαμεθανός, χριστιανός, βουδιστής.
Φυλή	Λευκός, μαύρος, κίτρινος, ερυθρόδερμος.
Επάγγελμα	Εκπαιδευτικός, δικηγόρος, έμπορος.

[1.7] Ποιοτικές Μεταβλητές.

Οι ποιοτικές μεταβλητές επιτρέπουν την κατάταξη των μονάδων ενός πληθυσμού σε κατηγορίες, αλλά και την ιεράρχηση των μονάδων του πληθυσμού. Εκτός από την ιεράρχηση των τιμών των ποιοτικών μεταβλητών δεν είναι δυνατή οποιαδήποτε μέτρηση της αποστάσεως ή της διαφοράς των τιμών αυτών.

[1.7.1] Παραδείγματα Ποιοτικών Μεταβλητών.

Μεταβλητές	Τιμές

Σοβαρότητα ατυχήματος	Σοβαρότερο, σοβαρό, ελαφρότερο, ελαφρό.
Κατάσταση υγείας	Κακή, καλή, άριστη.
Κοινωνική θέση	Κατώτατη, κατώτερη, ανώτερη, ανώτατη.

[1.8] Ποσοτικές Μεταβλητές.

Οι ποσοτικές μεταβλητές επιτρέπουν την κατάταξη των μονάδων σε κατηγορίες, την ιεράρχηση των μονάδων, παρέχουν πληροφόρηση για τις διαφορές μεταξύ των θέσεων δύο μονάδων και την απόλυτη θέση των μονάδων του πληθυσμού.

[1.8.1] Παραδείγματα Ποσοτικών Μεταβλητών.

Μεταβλητές	Τιμές
Ηλικία ατόμου	Θετικοί αριθμοί
Εισόδημα ατόμου	Θετικοί αριθμοί.
Αριθμός παιδιών οικογένειας	0,1,2,3,...
Ημερήσια απασχόληση ατόμου	0,1,2,3,...,8
Διάρκεια τριτοβάθμιας εκπαίδευσης	0,1,2,3,...,6

ατόμου.	
Αριθμός γάμων ατόμου	0,1,2,3.

[1.9] Διακριτές Ποσοτικές Μεταβλητές

Ποσοτικές μεταβλητές που έχουν το πολύ αριθμήσιμο σύνολο τιμών λέγονται διακριτές.

[1.9.1] Παραδείγματα Διακριτών Ποσοτικών Μεταβλητών.

Μεταβλητές	Τιμές
Αριθμός αγοριών μιας πενταμελούς οικογένειας.	0,1,2,3.
Αριθμός μελών οικογένειας.	2,3,4,...,
Αριθμός επιβατικών αυτοκινήτων μιας πόλης.	0,1,2,3,...,
Αριθμός τυπογραφικών λαθών που υπάρχουν σε μιά σελίδα βιβλίου.	0,1,2,3,...,

[1.10] Συνεχείς Ποσοτικές Μεταβλητές

Οι ποσοτικές μεταβλητές που παίρνουν όλες τις τιμές ενός διαστήματος πραγματικών αριθμών λέγονται συνεχείς.

[1.10.1] Παραδείγματα Συνεχών Ποσοτικών Μεταβλητών.

Μεταβλητές	Τιμές
Ηλικία ατόμου	Θετικοί αριθμοί.
Βάρος ατόμου	Θετικοί αριθμοί.
Ανάστημα ατόμου	Θετικοί αριθμοί.
Εισόδημα ατόμου	Θετικοί αριθμοί.

[1.11] Συλλογή Στατιστικών Στοιχείων

Η πρώτη φάση της στατιστικής μελέτης ενός φαινομένου ή προβλήματος είναι η συλλογή των αριθμητικών δεδομένων ή στοιχείων. Η συλλογή στοιχείων είναι η διαδικασία μέτρησης ή απαρίθμησης χαρακτηριστικών ιδιοτήτων των μονάδων ενός συνόλου ή τμήματός του και η καταγραφή των αριθμητικών δεδομένων που προκύπτουν.

[1.11.1] Τεχνικές Συλλογής Στατιστικών Στοιχείων.

Οι τεχνικές συλλογής στατιστικών στοιχείων είναι η απογραφή και η δειγματοληψία.

[1] Απογραφή

Η εφαρμογή της τεχνικής αυτής αποβλέπει στην συγκέντρωση στατιστικών στοιχείων απ' όλες τις μονάδες ενός στατιστικού πληθυσμού.

[2] Δειγματοληψία.

Η εφαρμογή της τεχνικής αυτής αποβλέπει στην συγκέντρωση στατιστικών στοιχείων από ένα υποσύνολο των μονάδων ενός στατιστικού πληθυσμού.

[1.12] Επεξεργασία Στατιστικών Στοιχείων.

Μετά την συγκέντρωση των στατιστικών στοιχείων ακολουθεί η φάση της επεξεργασίας των στοιχείων. Στη δεύτερη αυτή φάση περιλαμβάνονται εκείνες οι ενέργειες οι οποίες είναι απαραίτητες για τη συστηματική παρουσίαση, ανάλυση και αξιοποίηση των στοιχείων αυτών. Στις ενέργειες αυτές συμπεριλαμβάνονται ο έλεγχος πληρότητας, σαφήνειας και συνέπειας των συγκεντρωθέντων στοιχείων καθώς και οι διάφοροι πρόσθετοι αριθμητικοί υπολογισμοί.

[1.13] Οργάνωση Στατιστικών Στοιχείων.

Σκοπός της οργάνωσης είναι η κατάλληλη εγγραφή των στατιστικών στοιχείων σε διάφορα μέσα, [ηλεκτρονικούς υπολογιστές] ώστε να επιτυγχάνεται η πλέον εύκολη και αποδοτική χρησιμοποίηση των στοιχείων αυτών.

[1.14] Παρουσίαση Στατιστικών Στοιχείων.

Μετά από την συλλογή, επεξεργασία και οργάνωση των συλλεγέντων στατιστικών στοιχείων ακολουθεί η διαδικασία της συνοπτικής παρουσίασης των στοιχείων αυτών. Η παρουσίαση των στατιστικών στοιχείων γίνεται με τους στατιστικούς πίνακες, τις γραφικές παραστάσεις ή διαγράμματα και με τις συνοπτικές εκθέσεις ή αναφορές.

[1] Στατιστικοί Πίνακες.

Οι πίνακες αυτοί δίνουν συστηματικές κατατάξεις των στατιστικών στοιχείων. Οι κατατάξεις γίνονται με την χρήση ορισμένων κριτηρίων. Τα στοιχεία παρουσιάζονται σε γραμμές και στήλες και κατά ένα τέτοιο τρόπο ο οποίος διευκολύνει την μελέτη και την αξιοποίηση των στοιχείων αυτών.

[2] Γραφικές Παραστάσεις ή Διαγράμματα.

Η παρουσίαση στατιστικών στοιχείων με γραφικές παραστάσεις ή διαγράμματα κάνει χρήση γεωμετρικών σχημάτων ή απεικονήσεων για την αναπαράσταση των στοιχείων αυτών.

[3] Συνοπτικές Εκθέσεις ή Αναφορές.

Στα κείμενα των συνοπτικών εκθέσεων ή αναφορών ενσωματώνονται τα σημαντικότερα αποτελέσματα, σχολιάζονται η σημασία και οι εφαρμογές αυτών και συνοψίζονται τα συμπεράσματα της όλης στατιστικής μελέτης.

Η παρουσίαση των στατιστικών στοιχείων σε πίνακες ή διαγράμματα γίνεται συνήθως με τη βοήθεια μιάς ή περισσοτέρων εκ των παρακάτω κατατάξεων των στοιχείων αυτών.

[1] Γεωγραφικές Κατατάξεις.

Η παρουσίαση υπάρχουσας γεωγραφικής διαφοροποίησης των τιμών μιας στατιστικής μεταβλητής γίνεται με την χρήση γεωγραφικών πινάκων και διαγραμμάτων. Ο απλούστερος γεωγραφικός πίνακας αποτελείται από δύο στήλες. Η πρώτη στήλη περιλαμβάνει ανόματα περιοχών, χωρών ή ηπείρων και η δεύτερη στήλη περιλαμβάνει τις τιμές της μεταβλητής στις αντίστοιχες περιοχές, χώρες ή ηπείρους. Ένα γεωγραφικό διάγραμμα είναι ένας χάρτης μιάς περιοχής, χώρας ή ηπείρου, τα τμήματα του οποίου προυσιάζονται με διαφορετικά χρώματα ή γραμμοσκιάσεις η σημασία των οποίων διευκρινίζεται σε σχετικό υπόμνημα.

[2] Χρονολογικές Κατατάξεις.

Χρονολογική κατάταξη λέγεται η παρουσίαση των τιμών μιάς ποσοτικής στατιστικής μεταβλητής σε σχέση με το χρόνο. Οι χρονολογικοί πίνακες και τα χρονολογικά διαγράμματα επιτρέπουν την παρακολούθηση της διαχρονικής εξέλιξης ενός φαινομένου.

[3] Κατηγορικές, Ποιοτικές και Ποσοτικές Κατατάξεις.

Τα στατιστικά στοιχεία ομαδοποιούνται με κατάλληλες μεθόδους και παρουσιάζονται με τη βοήθεια ειδικών συνοπτικών κατατάξεων οι οποίες είναι γνωστές στη στατιστική ορολογία ως κατανομές συχνότητας.

Οι κατατάξεις αυτές ή κατανομές συχνότητας ονομάζονται κατηγορικές, ποιοτικές ή ποσοτικές ανάλογα με τη φύση της υπό μελέτη στατιστικής μεταβλητής και δίνουν με απλό και συνοπτικό τρόπο την εικόνα του αντίστοιχου στατιστικού πληθυσμού.

Οι χρησιμότητα των κατανομών συχνότητας είναι μεγάλη για την περιγραφή και ανάλυση διαφόρων πληθυσμιακών χαρακτηριστικών.

[1.15] Ανάλυση Στατιστικών Στοιχείων.

Σκοπός της ανάλυσης στατιστικών στοιχείων ή στατιστικής ανάλυσης αριθμητικών στοιχείων είναι να αποκτήσουν οι αριθμοί έννοια και σημασία. Η στατιστική ανάλυση αποτελεί αναγκαίο και πολύτιμο τρόπο ποσοτικής διερεύνησης προβλημάτων πολλών τομέων όπως ο Οικονομικός τομέας, η Κοινωνιολογία, η Βιολογία, η Ανθρωπολογία, η Ιατρική και η Δημόσια Υγεία,

η Ψυχολογία, η Εκπαίδευση.

[1.16] Τομείς της Στατιστικής.

Στο περιεχόμενο της στατιστικής περιλαμβάνονται οι δύο ακόλουθοι τομείς.

[1] Περιγραφική Στατιστική.

[2] Επαγωγική Στατιστική.

Οι τομείς αυτοί καθορίζονται ανάλογα με τον επιδιωκόμενο σκοπό.

[1.17] Περιγραφική Στατιστική.

Ο τομέας αυτός της στατιστικής ασχολείται με τις μεθόδους εξαγωγής περιγραφικών συμπερασμάτων που αφορούν ένα στατιστικό πληθυσμό.

[11.17.1] Τεχνικές Περιγραφικής Στατιστικής.

Οι σημαντικότερες και συνηθέστερα χρησιμοποιούμενες τεχνικές της περιγραφικής στατιστικής είναι :

[1] Μονομεταβλητές Κατανομές Συχνότητας.

Οι κατανομές αυτές και οι αντίστοιχες γραφικές παραστάσεις αυτών παρουσιάζουν συνοπτικά τη δομή ενός στατιστικού πληθυσμού. Οι μονομεταβλητές κατανομές συχνότητας διακρίνονται σε κατηγορικές, ποιοτικές ή ποσοτικές κατανομές.

[2] Οι παράμετροι των Μονομεταβλητών Κατανομών Συχνότητας.

Ο υπολογισμός ορισμένων παραμέτρων [χαρακτηριστικών μεγεθών] των μονομεταβλητών κατανομών συχνότητας, παρέχει χρήσιμες πληροφορίες για τη μορφή των κατανομών αυτών.

[3] Οι Διμεταβλητές ή Πολυμεταβλητές Κατανομές Συχνότητας.

Οι κατανομές αυτές παρουσιάζουν συνοπτικά τη δομή ενός στατιστικού πληθυσμού, ο οποίος μελετάται ως προς δύο ή περισσότερα χαρακτηριστικά. Η παρουσίαση της συνοπτικής δομής του πληθυσμού γίνεται με την χρήση της κατάταξης των μονάδων του στατιστικού πληθυσμού με βάση τις τιμές δύο ή περισσοτέρων στατιστικών μεταβλητών.

[4] Οι Παράμετροι των Διμεταβλητών ή Πολυμεταβλητών Κατανομών Συχνότητας.

Ο υπολογισμός ορισμένων παραμέτρων των διμεταβλητών ή πολυμεταβλητών κατανομών συχνότητας, παρέχει χρήσιμες πληροφορίες για την αλληλοεξάρτηση και τον τρόπο συμμεταβολής δύο ή περισσοτέρων στατιστικών μεταβλητών.

[5] Τα Στατιστικά Μαθηματικά Υποδείγματα.

Η προσαρμογή στατιστικών μαθηματικών υποδειγμάτων σε αριθμητικά στατιστικά στοιχεία είναι σημαντική για την περιγραφή της δομής και την αποκάλυψη ιδιοτήτων των χαρακτηριστικών ενός στατιστικού πληθυσμού.

[6] Οι Χρονολογικές Σειρές.

Η χρήση των υποδειγμάτων αυτών είναι σημαντική για την διαπίστωση, περιγραφή και μελέτη της διαχρονικής εξέλιξης ενός φαινομένου.

[7] Οι Γραμμές Τάσης.

Η προσαρμογή των μαθηματικών υποδειγμάτων τα οποία λέγονται γραμμές τάσης, σε αριθμητικά στατιστικά στοιχεία είναι σημαντική για την περιγραφή της μακροχρόνιας συμπεριφοράς ενός φαινομένου.

[8] Οι Αριθμοδείκτες.

Η κατάρτηση των αριθμοδεικτών είναι σημαντική για την παρακολούθηση της διαχρονικής ή γεωγραφικής μεταβολής ενός σύνθετου μεγέθους.

[1.18] Επαγωγική Στατιστική.

Η επαγωγική στατιστική ασχολείται με τις μεθόδους οι οποίες καθιστούν δυνατή

[1] Τη γενίκευση των συμπερασμάτων που προκύπτουν από τις περιγραφικές στατιστικές αναλύσεις στοιχείων τα οποία αναφέρονται σε τμήμα ενός στατιστικού πληθυσμού που ερευνάται, ώστε να ισχύουν τα συμπεράσματα αυτά για ολόκληρο το στατιστικό πληθυσμό.

[2] Την πρόβλεψη της μελλοντικής συμπεριφοράς στατιστικών μεταβλητών με χρήση στοιχείων του παρελθόντος.

[3] Την εξέλιξη των σχέσεων μεταξύ στατιστικών μεταβλητών με χρήση επίσης στοιχείων του παρελθόντος.

[4] Τη λήψη αποφάσεων κάτω από συνθήκες αβεβαιότητας.

ΚΕΦΑΛΑΙΟ 2

ΜΟΝΟΜΕΤΑΒΛΗΤΕΣ ΚΑΤΑΝΟΜΕΣ ΣΥΧΝΟΤΗΤΑΣ

[2.1] Ταξινόμηση Μονομεταβλητών Κατανομών Συχνότητας

Η παρουσίαση της συνοπτικής δομής ενός στατιστικού πληθυσμού, ο οποίος μελετάται ως πρός ένα χαρακτηριστικό ή ισοδύναμα ως προς μία στατιστική μεταβλητή, γίνεται με τη χρήση της κατανομής συχνότητας των τιμών της στατιστικής μεταβλητής.

Η ακόλουθη ταξινόμηση ισχύει για τις κατανομές συχνότητας των τιμών των στατιστικών μεταβλητών.

[1] Κατηγορικές Κατανομές Συχνότητας.

[2] Ποιοτικές Κατανομές Συχνότητας.

[3] Ποσοτικές Κατανομές Συχνότητας.

Δηλαδή η κατανομή συχνότητας των τιμών μιας στατιστικής μεταβλητής ταξινομείται ανάλογα με τη φύση της στατιστικής μεταβλητής.

[2.2] Κατανομή Συχνότητας μιας Κατηγορικής Μεταβλητής.

Η κατανομή συχνότητας μιας κατηγορικής μεταβλητής περιλαμβάνει :

[1] τις κατηγορίες στις οποίες κατατάσσονται οι μονάδες του πληθυσμού.

[2] τους αριθμούς οι οποίοι δηλώνουν το μέγεθος κάθε κατηγορίας. Οι αριθμοί αυτοί λέγονται απόλυτες συχνότητες. Αρκετές φορές στην κατανομή συχνότητος μιάς κατηγορικής μεταβλητής περιλαμβάνονται και οι σχετικές συχνότητες. Οι συχνότητες αυτές προκύπτουν από την έκφραση των απόλυτων συχνοτήτων ως ποσοστών του πλήθους των "τιμών" της κατηγορικής μεταβλητής.

Ο πίνακας 2.1 δίνει ένα παράδειγμα κατανομής συχνότητας μιας κατηγορικής μεταβλητής.

Πίνακας 2.1

Κατανομή Συχνότητας της Οικογενειακής Κατάστασης των Εργαζομένων μιας Επιχείρησης

Οικογενειακή κατάσταση	Απόλυτες Συχνότητες	Σχετικές Συχνότητες (%)
Άγαμος	80	25
Έγγαμος	160	50
Χήρος	32	10
Διαζευγμένος	48	15
	320	100

Από τον πίνακα 2.1 προκύπτει ότι :

[1] η στατιστική μεταβλητή είναι η οικογενειακή κατάσταση

[2] οι κατηγορίες - τιμές - της στατιστικής μεταβλητής είναι άγαμος, έγγαμος, χήρος και διαζευγμένος.

[3] οι απόλυτες συχνότητες των τεσσάρων κατηγοριών της στατιστικής μεταβλητής είναι 80, 160, 32 και 48.

[4] οι σχετικές συχνότητες των τεσσάρων κατηγοριών της στατιστικής μεταβλητής είναι τα ποσοστά 25%, 50%, 10% και 15%.

[2.3] Γραφική απεικόνιση της Κατανομής Συχνότητας μιας Κατηγορικής Μεταβλητής.

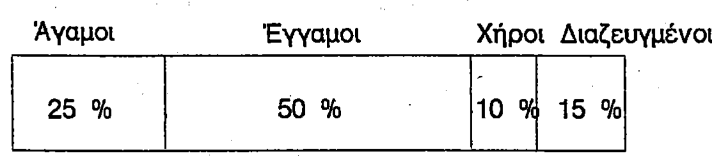
Η γραφική απεικόνιση της κατανομής συχνότητας μιας κατηγορικής μεταβλητής γίνεται με το ορθογώνιο και το κυκλικό διάγραμμα.

[1] Το ορθογώνιο διάγραμμα είναι ένα ορθογώνιο παραλληλόγραμμο το οποίο υποδιαιρείται σε μικρότερα ορθογώνια παραλληλόγραμμα αντίστοιχα προς τις κατηγορίες της κατανομής συχνότητας μιας κατηγορικής μεταβλητής. Τα μικρά ορθογώνια τα οποία παριστάνουν τις διάφορες κατηγορίες της κατανομής έχουν εμβαδά ίσα - υπό κλίμακα - προς τις αντίστοιχες απόλυτες ή σχετικές συχνότητες των κατηγοριών της μεταβλητής.

[2] Το κυκλικό διάγραμμα είναι ένας κύκλος που υποδιαιρείται σε κυκλικούς τομείς αντίστοιχα προς τις κατηγορίες της κατανομής συχνότητας μιάς κατηγορικής μεταβλητής. Οι κυκλικοί τομείς οι οποίοι παριστάνουν τις διάφορες κατηγορίες της κατανομής έχουν εμβαδά ίσα - υπό κλίμακα - προς τις αντίστοιχες, απόλυτες ή σχετικές συχνότητες των κατηγοριών της μεταβλητής.

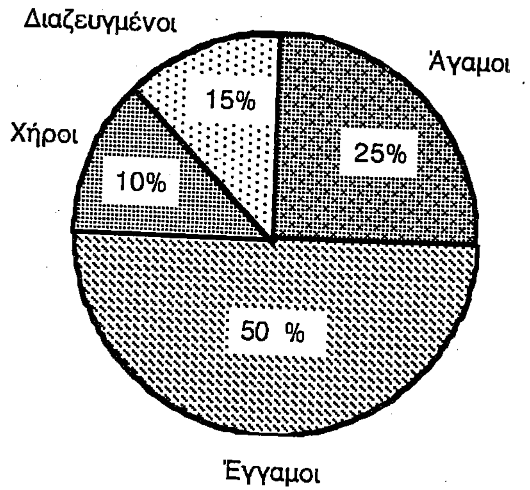
Το σχήμα 2.1 είναι η γραφική απεικόνιση της κατηγορικής κατανομής συχνότητας του πίνακα 2.1 με τη χρήση ορθογωνίου διαγράμματος.

Σχήμα 2.1



Το σχήμα 2.2 είναι η γραφική απεικόνιση της κατηγορικής κατανομής συχνότητας του πίνακα 2.1 με τη χρήση κυκλικού διαγράμματος.

Σχήμα 2.2



[2.4] Κατανομή Συχνότητας μιάς Ποιοτικής Μεταβλητής.

Η κατανομή συχνότητας μιάς ποιοτικής μεταβλητής περιλαμβάνει :

[1] τις κατηγορίες στις οποίες κατατάσσονται ιεραρχικά οι μονάδες του πληθυσμού.

[2] τις απόλυτες και τις σχετικές συχνότητες, όπως και στην περίπτωση μιάς κατηγορικής μεταβλητής.

Ο πίνακας 2.2 δίνει ένα παράδειγμα κατανομής συχνότητας μιάς ποιοτικής μεταβλητής.

Πίνακας 2.2

Κατανομή Συχνότητας της Κατάστασης Υγείας των Κατοίκων Ενός Χωριού

Κατάσταση Υγείας	Απόλυτες Συχνότητες	Σχετικές Συχνότητες (%)
Κακή	32	5
Μέτρια	96	15
Καλή	288	45
Άριστη	224	35
	640	100

Από τον πίνακα 2.2 προκύπτει ότι :

[1] η στατιστική μεταβλητή είναι η κατάσταση υγείας των κατοίκων ενός χωριού.

[2] οι κατηγορίες - τιμές - της στατιστικής μεταβλητής είναι κακή, μέτρια, καλή και άριστη.

[3] οι απόλυτες συχνότητες της στατιστικής μεταβλητής είναι 32, 96, 288 και 224.

[4] οι σχετικές συχνότητες της στατιστικής μεταβλητής είναι τα ποσοστά 5%, 15%, 45% και 35%.

[2.5] Γραφική Απεικόνιση της κατανομής Συχνότητας μιάς Ποιοτικής Μεταβλητής.

Η γραφική απεικόνιση της κατανομής συχνότητας μιάς ποιοτικής μεταβλητής γίνεται με το ορθογώνιο, κυκλικό και το ακιδωτό διάγραμμα.

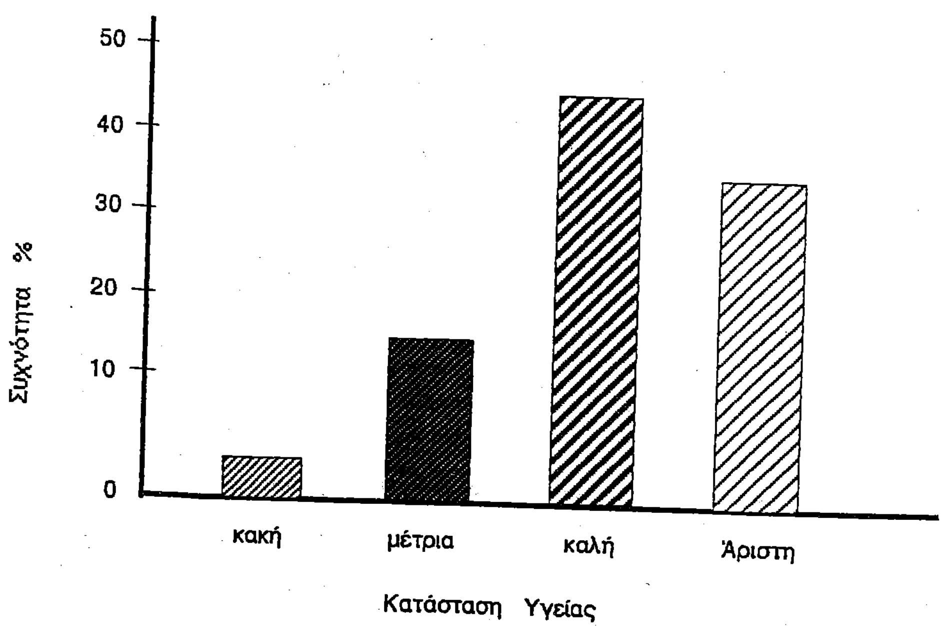
[1] Το ορθογώνιο διάγραμμα είναι ίδιο με το ορθογώνιο διάγραμμα που χρησιμοποιείται για την γραφική απεικόνιση μιάς κατηγορικής κατανομής συχνότητας.

[2] Το κυκλικό διάγραμμα είναι ίδιο με το κυκλικό διάγραμμα που χρησιμοποιείται για την γραφική απεικόνιση μιάς κατηγορικής κατανομής συχνότητας.

[3] Το ακιδωτό διάγραμμα αποτελείται από ορθογώνια παραλληλόγραμμα με ίσες βάσεις και εμβαδά ίσα - υπό κλίμακα - προς τις απόλυτες ή σχετικές συχνότητες της ποιοτικής μεταβλητής. Τα ορθογώνια παραλληλόγραμμα του ακιδωτού διαγράμματος βρίσκονται χωριστά το ένα από το άλλο.

Το σχήμα 2.3 είναι η γραφική απεικόνιση της ποιοτικής κατανομής συχνότητας του πίνακα 2.2 με τη χρήση ακιδωτού διαγράμματος.

Σχήμα 2.3



[2.6] Κατανομή Συχνότητας μιας Διακριτής Ποσοτικής Μεταβλητής.

Η κατανομή συχνότητας μιάς διακριτής ποσοτικής μεταβλητής περιλαμβάνει :

[1] τις τιμές της μεταβλητής.

[2] τις απόλυτες και τις σχετικές συχνότητες, όπως και στη περίπτωση μιάς κατηγορικής ή ποιοτικής μεταβλητής, οι οποίες αντιστοιχούν στις τιμές της μεταβλητής.

Ο πίνακας 2.3 δίνει ένα παράδειγμα κατανομής συχνότητας μιάς διακριτής μεταβλητής.

Πίνακας 2.3

Κατανομή Συχνότητας του Αριθμού των Δωματίων των Κατοικιών μιάς Πόλης

Αριθμός Δωματίων	Απόλυτες Συχνότητες	Σχετικές Συχνότητες (%)
1	243	5
2	1215	25
3	1701	35
4	972	20
5	729	15
	4860	100

Από τον πίνακα 2.3 προκύπτει ότι :

[1] η στατιστική μεταβλητή είναι ο αριθμός των δωματίων των κατοικιών μιάς πόλης.

[2] οι τιμές της στατιστικής μεταβλητής είναι 1, 2, 3 , 4 και 5.

[3] οι απόλυτες συχνότητες της στατιστικής μεταβλητής είναι : 243, 1215, 1701, 972 και 729.

[4] οι σχετικές συχνότητες της στατιστικής μεταβλητής είναι τα ποσοστά : 5%, 25%, 35%, 20% και 15%.

[2.7] Γραφική Απεικόνιση της Κατανομής Συχνότητας μιάς Διακριτής Ποσοτικής Μεταβλητής.

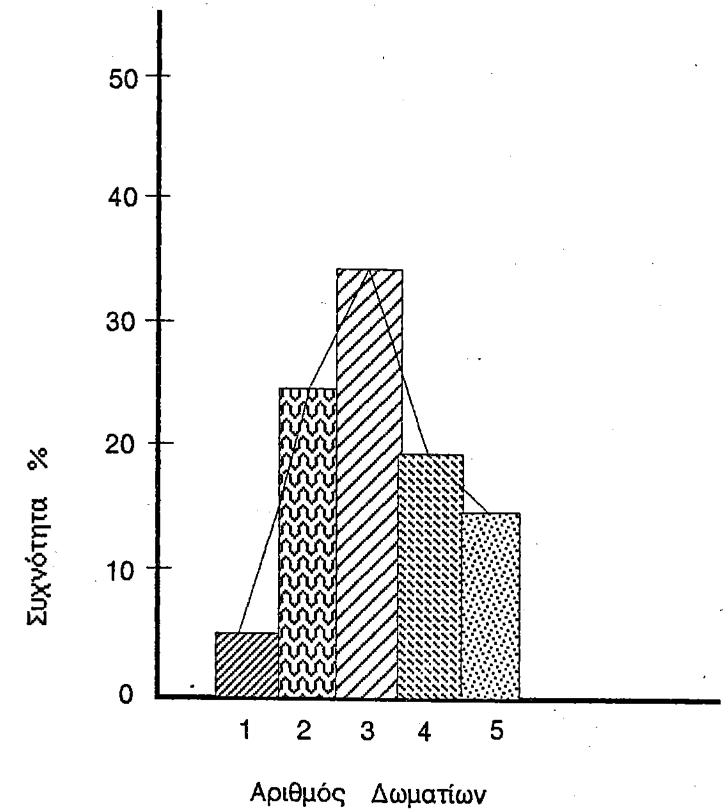
Η γραφική απεικόνιση της κατανομής συχνότητας μιάς διακριτής ποσοτικής μεταβλητής γίνεται με το ιστόγραμμα και το πολύγωνο συχνότητας.

[1] Το ιστόγραμμα μιάς διακριτής ποσοτικής κατανομής συχνότητας αποτελείται από συνεχόμενα ορθογώνια με ίσες βάσεις και εμβαδά ίσα - υπό κλίμακα - πρός τις απόλυτες ή σχετικές συχνότητες της διακριτής ποσοτικής μεταβλητής. Οι τιμές της μεταβλητής αναγράφονται στο μέσον κάθε βάσης.

[2] Το πολύγωνο συχνότητας είναι μιά τεθλασμένη γραμμή η οποία προκύπτει από το αντίστοιχο ιστόγραμμα της κατανομής συχνότητας. Κάθε ευθύγραμμο τμήμα του πολυγώνου συχνότητας ενώνει τα μέσα των άνω βάσεων δύο διαδοχικών ορθογωνίων του ιστογράμματος.

Το σχήμα 2.4 είναι η γραφική απεικόνιση της διακριτής ποσοτικής κατανομής συχνότητας του πίνακα 2.3 με τη χρήση ιστογράμματος και πολυγώνου συχνότητας.

Σχήμα 2.4



[2.8] Κατανομή Συχνότητας μιάς Συνεχούς Ποσοτικής Μεταβλητής.

Η κατανομή συχνότητας μιάς συνεχούς ποσοτικής μεταβλητής περιλαμβάνει :

[1] τα διαστήματα των τιμών της μεταβλητής,

[2] τις απόλυτες και τις σχετικές συχνότητες όπως και στη περίπτωση μιάς κατηγορικής, ποιοτικής ή διακριτής ποσοτικής μεταβλητής, οι οποίες αντιστοιχούν στα διαστήματα της μεταβλητής.

Ο πίνακας 2.4 δίνει ένα παράδειγμα κατανομής συχνότητας μιάς συνεχούς ποσοτικής μεταβλητής.

Πίνακας 2.4

Κατανομή Συχνότητας της Ηλικίας των Ατόμων Ενός Χωριού

Ηλικία Ατόμων	Απόλυτες Συχνότητες	Σχετικές Συχνότητες (%)
20 - 30	216	15
30 - 40	504	35
40 - 50	360	25
50 - 60	288	20
60 - 70	72	5
	1440	100

Από τον πίνακα 2.4 προκύπτει ότι :

[1] η στατιστική μεταβλητή είναι η ηλικία των ατόμων του χωριού.

[2] τα διαστήματα της στατιστικής μεταβλητής είναι (20-30], (30-40], (40-50], (50-60] και (60-70].

[3] οι απόλυτες συχνότητες της στατιστικής μεταβλητής είναι 216, 504, 360, 288 και 72.

[4] οι σχετικές συχνότητες της στατιστικής μεταβλητής είναι τα ποσοστά 15%, 35%, 25%, 20% και 5%.

[2.9] Γραφική Απεικόνιση της Κατανομής Συχνότητας μιάς Συνεχούς Ποσοτικής Μεταβλητής.

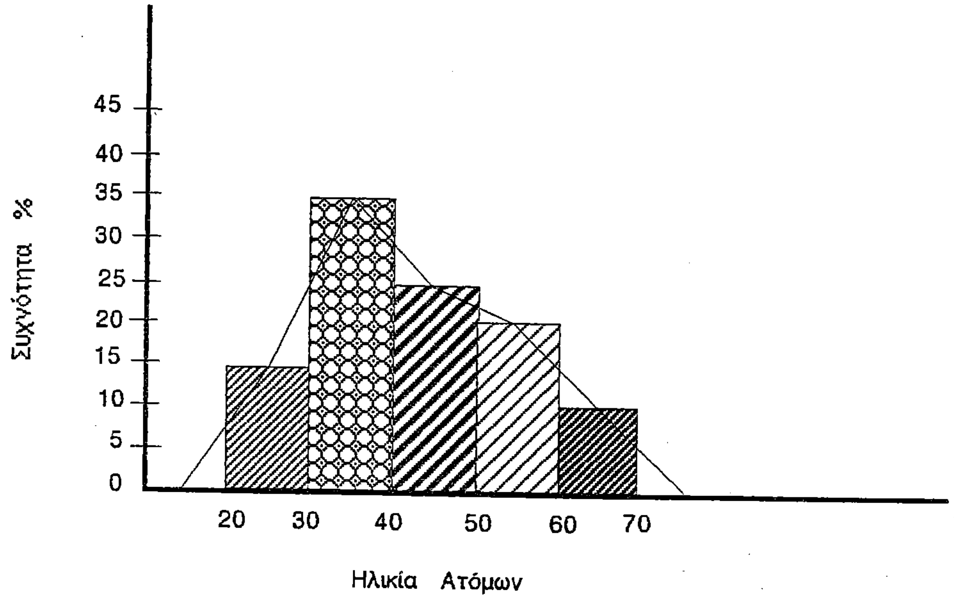
Η γραφική απεικόνιση της κατανομής συχνότητας μιάς συνεχούς ποσοτικής μεταβλητής γίνεται με το ιστόγραμμα και το πολύγωνο συχνότητας.

[1] Το ιστόγραμμα αποτελείται από συνεχόμενα ορθογώνια που έχουν βάσεις τα διαστήματα της μεταβλητής και εμβαδά ίσα - υπό κλίμακα - προς τις αντίστοιχες απόλυτες ή σχετικές συχνότητες.

[2] Το πολύγωνο συχνότητας είναι μία τεθλασμένη γραμμή η οποία προκύπτει από το αντίστοιχο ιστόγραμμα της κατανομής συχνότητας. Κάθε ευθύγραμμο τμήμα του πολυγώνου συχνότητας ενώνει τα μέσα των άνω βάσεων δύο διαδοχικών ορθογωνίων του ιστογράμματος. Συνήθως τα άκρα του πολυγώνου συχνότητας ενώνονται με τα μέσα των δύο γειτονικών τμημάτων - του προηγούμενου της πρώτης τάξης και του επόμενου της τελευταίας τάξης - θεωρουμένων και αυτών ως τάξεων.

Το σχήμα 2.5 είναι η γραφική απεικόνιση της συνεχούς ποσοτικής κατανομής συχνότητας του πίνακα 2.4 με τη χρήση ιστογράμματος και πολυγώνου συχνότητας.

Σχήμα 2.5



Η κατασκευή του ιστογράμματος και του πολυγώνου συχνότητας για κατανομή συχνότητας διακριτής ή συνεχούς ποσοτικής μεταβλητής είναι δυνατή μόνο για κλειστές κατανομές.

ΚΕΦΑΛΑΙΟ 3

ΜΟΝΟΜΕΤΑΒΛΗΤΕΣ ΑΘΡΟΙΣΤΙΚΕΣ ΚΑΤΑΝΟΜΕΣ

[3.1] Ταξινόμηση Μονομεταβλητών Αθροιστικών κατανομών.

Ο ορισμός της έννοιας της αθροιστικής κατανομής μιάς στατιστικής μεταβλητής απαιτεί την δυνατότητα ιεράρχησης των τιμών της στατιστικής μεταβλητής. Επειδή η δυνατότητα ιεράρχησης υπάρχει για τις τιμές των ποιοτικών και ποσοτικών στατιστικών μεταβλητών τότε η ακόλουθη ταξινόμηση ισχύει για τις αθροιστικές κατανομές των στατιστικών μεταβλητών.

[1] Ποιοτικές Αθροιστικές Κατανομές.

[2] Ποσοτικές Αθροιστικές Κατανομές.

[3.2] Αθροιστική Κατανομή μιάς Ποιοτικής Μεταβλητής.

Η αθροιστική κατανομή μιάς ποιοτικής μεταβλητής περιλαμβάνει :

[1] τις κατηγορίες στις οποίες κατατάσσονται ιεραρχικά οι μονάδες του πληθυσμού.

[2] τις απόλυτες και σχετικές αθροιστικές συχνότητες. Η πρώτη απόλυτη

αθροιστική συχνότητα δηλώνει το πλήθος των μονάδων του πληθυσμού οι οποίες κατατάσσονται στη πρώτη κατηγορία. Η δεύτερη απόλυτη αθροιστική συχνότητα δηλώνει το πλήθος των μονάδων του πληθυσμού οι οποίες κατατάσσονται στην πρώτη και την δεύτερη κατηγορία. Αντίστοιχα η τελευταία απόλυτη αθροιστική συχνότητα - δηλώνει το πλήθος των μονάδων του πληθυσμού. Οι σχετικές αθροιστικές συχνότητες προκύπτουν από την έκφραση των απόλυτων συχνοτήτων ως ποσοστών του πλήθους των μονάδων του πληθυσμού.

Ο πίνακας 3.1 δίνει ένα παράδειγμα αθροιστικής κατανομής μιάς ποιοτικής μεταβλητής.

Πίνακας 3.1

Αθροιστική Κατανομή της Κατάστασης Υγείας των Κατοίκων Ενός Χωριού

Κατάσταση Υγείας	Απόλυτες Συχνότητες	Σχετικές Συχνότητες (%)	Απόλυτες Αθροιστικές Συχνότητες	Σχετικές Αθροιστικές Συχνότητες (%)
Κακή	32	5	32	5
Μέτρια	96	15	128	20
Καλή	288	45	416	65
Άριστη	224	35	640	100
	640	100		

Οι τρείς πρώτες στήλες του πίνακα 3.1 είναι η κατανομή συχνότητας του πίνακα 2.2.

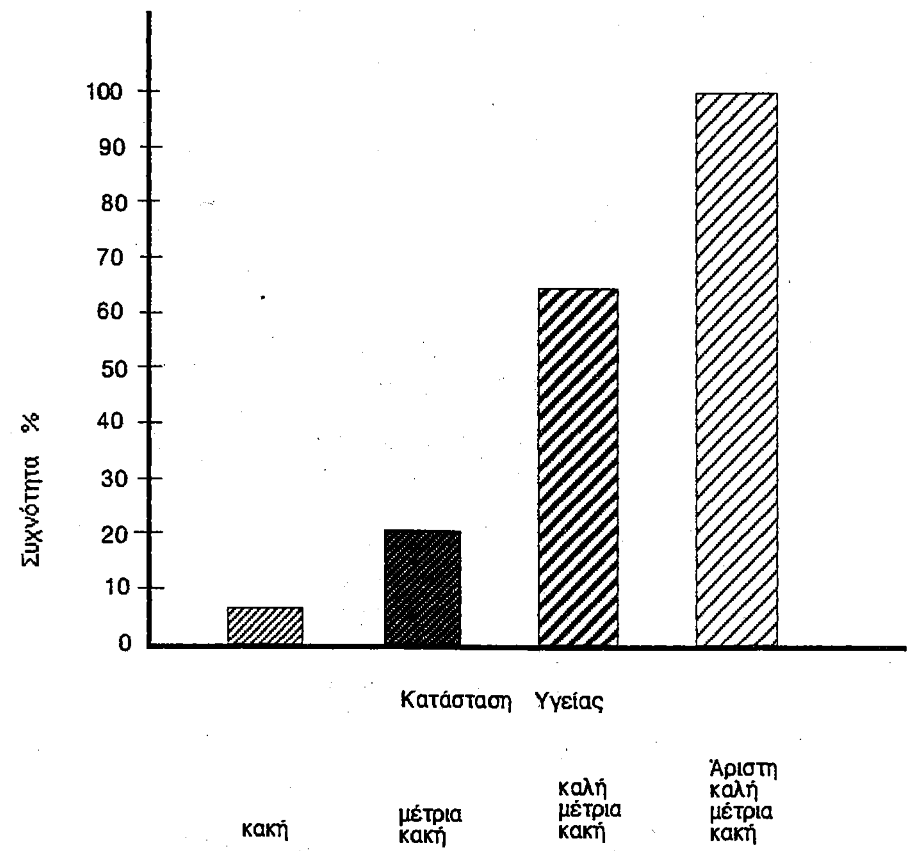
Η πρώτη, η τέταρτη και η πέμπτη στήλη του πίνακα 3.1 είναι η αθροιστική κατανομή η οποία αντιστοιχεί στη κατανομή συχνότητας του πίνακα 2.2.

[3.3] Γραφική Απεικόνιση της Αθροιστικής Κατανομής μιάς Ποιοτικής Μεταβλητής.

Η γραφική απεικόνιση της αθροιστικής κατανομής μιάς ποιοτικής μεταβλητής γίνεται με το αθροιστικό ακιδωτό διάγραμμα. Το διάγραμμα αυτό αποτελείται από ορθογώνια παραλληλόγραμμα με ίσες βάσεις και εμβαδά ίσα - υπό κλίμακα - προς τις απόλυτες ή σχετικές αθροιστικές συχνότητες της ποιοτικής μεταβλητής. Τα ορθογώνια παραλληλόγραμμα του αθροιστικού ακιδωτού διαγράμματος βρίσκονται χωριστά το ένα από το άλλο.

Το σχήμα 3.1 είναι η γραφική απεικόνιση της αθροιστικής κατανομής του πίνακα 3.1 με τη χρήση ακιδωτού διαγράμματος.

Σχήμα 3.1



[3.4] Αθροιστική Κατανομή μιας Διακριτής Ποσοτικής Μεταβλητής.

Η αθροιστική κατανομή μιας διακριτής ποσοτικής μεταβλητής περιλαμβάνει :

[1] τις τιμές της μεταβλητής.

[2] τις απόλυτες και τις σχετικές αθροιστικές συχνότητες. Η απόλυτη αθροιστική συχνότητα που αντιστοιχεί σε μιά δοθείσα τιμή της μεταβλητής δηλώνει τον αριθμό των μονάδων του πληθυσμού για τις οποίες οι τιμές της μεταβλητής είναι μικρότερες ή ίσες της δοθείσης τιμής. Οι σχετικές αθροιστικές συχνότητες προκύπτουν από την έκφραση των απόλυτων αθροιστικών συχνοτήτων ως ποσοστών του πλήθους των μονάδων του πληθυσμού.

Ο πίνακας 3.2 δίνει ένα παράδειγμα αθροιστικής κατανομής μιάς διακριτής ποσοτικής μεταβλητής.

Πίνακας 3.2

Αθροιστική Κατανομή του Αριθμού των Δωματίων των Κατοίκων μιάς Πόλης.

Αριθμός Δωματίων	Απόλυτες Συχνότητες	Σχετικές Συχνότητες (%)	Απόλυτες Αθροιστικές Συχνότητες	Σχετικές Αθροιστικές Συχνότητες (%)
1	243	5	243	5
2	1215	25	1458	30
3	1801	35	3159	65
4	972	20	4131	85
5	729	15	4860	100
	4860	100		

Οι τρείς πρώτες στήλες του πίνακα 3.2 είναι η κατανομή συχνότητας του πίνακα 2.3.

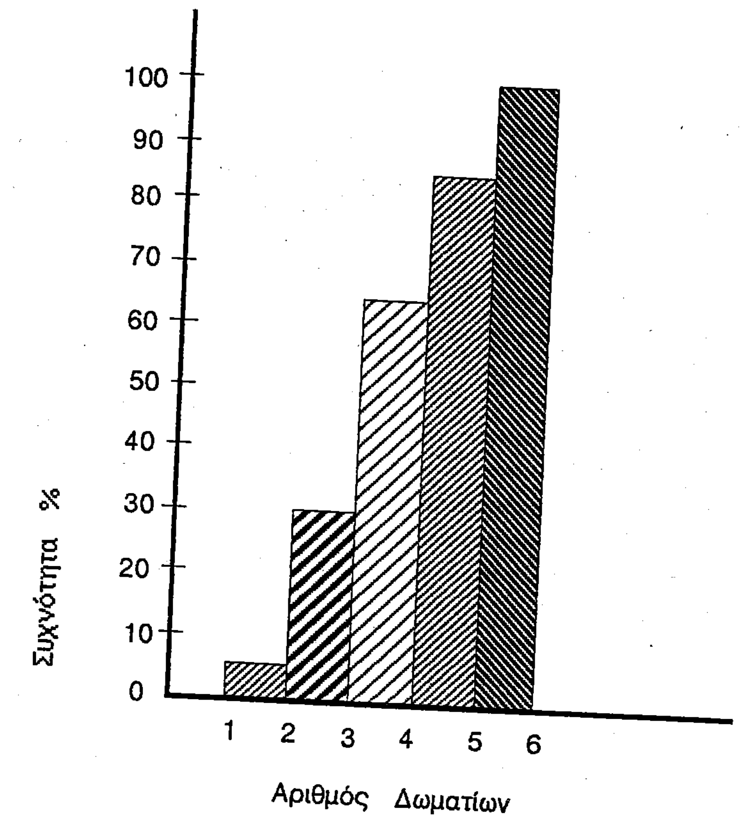
Η πρώτη, η τέταρτη και η πέμπτη στήλη του πίνακα 3.2 είναι η αθροιστική κατανομή η οποία αντιστοιχεί στη κατανομή συχνότητας του πίνακα 2.3.

[3.5] Γραφική Απεικόνιση της Αθροιστικής Κατανομής μιάς Διακριτής Ποσοτικής Μεταβλητής.

Η γραφική απεικόνιση της αθροιστικής κατανομής μιάς διακριτής ποσοτικής μεταβλητής γίνεται με το κλιμακωτό διάγραμμα. Το διάγραμμα αυτό αποτελείται από συνεχόμενα ορθογώνια με ίσες βάσεις και εμβαδά ίσα - υπό κλίμακα - προς τις απόλυτες αθροιστικές ή σχετικές αθροιστικές συχνότητες της διακριτής ποσοτικής μεταβλητής. Οι τιμές της μεταβλητής αναγράφονται στα άκρα κάθε βάσης.

Το σχήμα 3.2 είναι η γραφική απεικόνιση της αθροιστικής κατανομής του πίνακα 3.2 με τη χρήση κλιμακωτού διαγράμματος.

Σχήμα 3.2



[3.6] Αθροιστική Κατανομή μιας Συνεχούς Ποσοτικής Μεταβλητής.

Η αθροιστική κατανομή μιάς συνεχούς ποσοτικής μεταβλητής περιλαμβάνει :

[1] τα διαστήματα των τιμών της μεταβλητής.

[2] τις απόλυτες και τις σχετικές αθροιστικές συχνότητες. Η απόλυτη αθροιστική συχνότητα που αντιστοιχεί στο δεύτερο άκρο ενός δοθέντος διαστήματος της μεταβλητής δηλώνει τον αριθμό των μονάδων του πληθυσμού για τις οποίες οι τιμές της μεταβλητής είναι μικρότερες ή ίσες

του δεύτερου άκρου του δοθέντος διαστήματος. Οι σχετικές αθροιστικές συχνότητες προκύπτουν από την έκφραση των απολύτων αθροιστικών συχνοτήτων ως ποσοστών του πλήθους των μονάδων του πληθυσμού.

Ο πίνακας 3.3 δίνει ένα παράδειγμα αθροιστικής κατανομής μιάς συνεχούς ποσοτικής μεταβλητής.

Πίνακας 3.3

Αθροιστική Κατανομή της Ηλικίας των Ατόμων Ενός Χωριού

Ηλικία Ατόμων	Απόλυτες Συχνότητες	Σχετικές Συχνότητες (%)	Απόλυτες Αθροιστικές Συχνότητες	Σχετικές Αθροιστικές Συχνότητες (%)
20 - 30	216	15	216	15
30 - 40	504	35	720	50
40 - 50	360	25	1080	75
50 - 60	288	20	1368	95
60 - 70	72	5	1440	100
	1440	100		

Οι τρεις πρώτες στήλες του πίνακα 3.3 είναι η κατανομή συχνότητας του πίνακα 2.4.

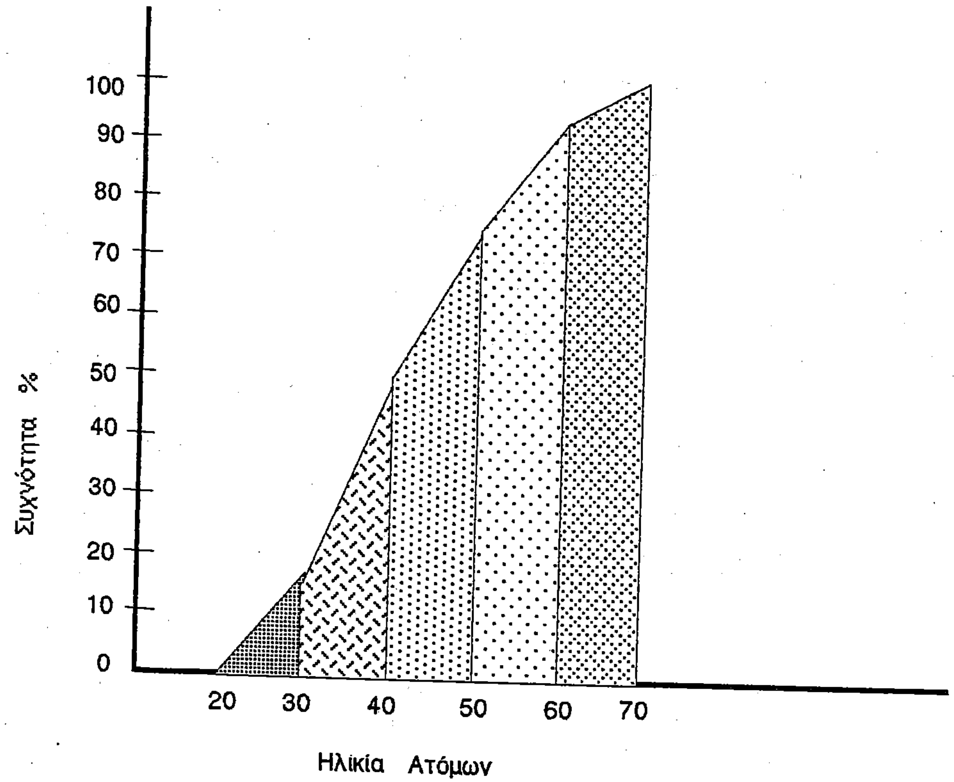
Η πρώτη, η τέταρτη και η πέμπτη στήλη του πίνακα 3.3 είναι η αθροιστική κατανομή η οποία αντιστοιχεί στη κατανομή συχνότητας του πίνακα 2.4.

[3.7] Γραφική Απεικόνιση της Αθροιστικής Κατανομής μιάς Συνεχούς Ποσοτικής Μεταβλητής.

Η γραφική απεικόνιση της αθροιστικής κατανομής μιάς συνεχούς ποσοτικής μεταβλητής γίνεται με το αθροιστικό διάγραμμα. Το διάγραμμα αυτό αποτελείται από συνεχόμενα ορθογώνια τραπέζια. Τα μήκη των παραλλήλων πλευρών ενός ορθογωνίου τραπεζίου αντιπροσωπεύουν - υπό κλίμακα - διαδοχικές απόλυτες ή σχετικές αθροιστικές συχνότητες.

Το σχήμα 3.3 είναι η γραφική απεικόνιση της αθροιστικής κατανομής του πίνακα 3.3 με τη χρήση αθροιστικού διαγράμματος.

Σχήμα 3.3



[3.8] <u>Συμβολική Παράσταση της Κατανομής Συχνότητας και Αθροιστικής Κατανομής Κατηγορικής και Ποιοτικής Μεταβλητής.</u>

Η συμβολική παράσταση της κατανομής συχνότητας και της αθροιστικής κατανομής μιάς κατηγορικής ή ποιοτικής αθροιστικής μεταβλητής είναι σχεδόν ίδια με την συμβολική παράσταση της κατανομής

συχνότητας και της αθροιστικής κατανομής μιάς διακριτής ποσοτικής μεταβλητής που παρουσιάζεται στο εδάφιο 3.9. Οι κατηγορίες της κατηγορικής ή ποιοτικής μεταβλητής αντιστοιχούν στις τιμές της διακριτής ποσοτικής μεταβλητής.

[3.9] <u>Συμβολική Παράσταση της Κατανομής Συχνότητας και Αθροιστικής Κατανομής μιάς Διακριτής Ποσοτικής Μεταβλητής.</u>

Το εδάφιο αυτό παρουσιάζει τον ενιαίο τρόπο συμβολικής παράστασης της κατανομής συχνότητας και της αθροιστικής κατανομής μιάς διακριτής ποσοτικής μεταβλητής.

Συμβολική Παράσταση της Κατανομής Συχνότητας και της Αθροιστικής Κατανομής μιάς Διακριτής Μεταβλητής.

Πίνακας 3.4

Τιμές της Μεταβλητής	Συχνότητες		Αθροιστικές Συχνότητες	
X	Απόλυτες f_i	Σχετικές f_i/N	Απόλυτες F_i	Σχετικές F_i/N
x_1	f_1	f_1/N	F_1	F_1/N
x_2	f_2	f_2/N	F_2	F_2/N
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
x_i	f_i	f_i/N	F_i	F_i/N
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
x_k	f_k	f_k/N	F_k	F_k/N

Στις επόμενες παραγράφους του εδαφίου αυτού γίνεται επεξήγηση των χρησιμοποιούμενων συμβόλων που υπεισέρχονται στον πίνακα 2.1.

[1] Στατιστική Μεταβλητή.

Οι διακριτές στατιστικές μεταβλητές - όπως όλες οι στατιστικές μεταβλητές - παριστάνονται με κεφαλαία γράμματα.
X,Y,Z,...

[2] Τιμές της μεταβλητής.

Οι τιμές της μεταβλητής X παριστάνονται με x_i, i=1,2,...,κ. Το σύνολο των τιμών μιάς διακριτής μεταβλητής είναι πεπερασμένο ή το πολύ αριθμήσιμο.

[3] Απόλυτες Συχνότητες.

Η απόλυτη συχνότητα f_i συμβολίζει τον αριθμό των μονάδων του πληθυσμού των οποίων οι τιμές είναι ίσες με x_i, i=1,2,3,...,k.

[4] Απόλυτη Συνολική Συχνότητα.

Το άθροισμα

$$N = \sum_{i=1}^{k} f_i$$

λέγεται συνολική απόλυτη συχνότητα και είναι ίσο με το συνολικό αριθμό των μονάδων πληθυσμού.

[5] Σχετικές Συχνότητες.

Η σχετική συχνότητα f_i/N είναι το πηλίκο της αντίστοιχης απόλυτης συχνότητας f_i δια της συνολικής απόλυτης συχνότητας N και συμβολίζει την αναλογία των μονάδων του πληθυσμού των οποίων οι τιμές είναι ίσες με x_i, i =1,2,...,k.

[6] Απόλυτες Αθροιστικές Συχνότητες.

Η απόλυτη αθροιστική συχνότητα

$$F_i = \sum_{j=1}^{i} f_j$$

$$= F_{i-1} + f_i, \quad i =1,2,...,k$$

συμβολίζει τον αριθμό των μονάδων του πληθυσμού των οποίων οι τιμές είναι μικρότερες ή ίσες του αριθμού x_i, i =1,2,...,k.

[7] Σχετικές Αθροιστικές Συχνότητες.

Η σχετική αθροιστική συχνότητα F_i /N είναι το πηλίκο της αντίστοιχης απόλυτης αθροιστικής συχνότητας F_i δια της συνολικής απόλυτης συχνότητας N και συμβολίζει την αναλογία των μονάδων του πληθυσμού οι τιμές των οποίων είναι μικρότερες ή ίσες του αριθμού x_i, i =1,2,...,k.

[3.10] Συμβολική Παράσταση της Κατανομής Συχνότητας και Αθροιστικής Κατανομής μιάς Συνεχούς Ποσοτικής Μεταβλητής.

Το εδάφιο αυτό παρουσιάζει τον ενιαίο τρόπο συμβολικής παράστασης της κατανομής συχνότητας και της αθροιστικής κατανομής μιάς συνεχούς ποσοτικής μεταβλητής.

Συμβολική Παράσταση της Κατανομής Συχνότητας και της Αθροιστικής Κατανομής Συχνότητας μιας Συνεχούς Μεταβλητής.

Πίνακας 3.5

Τάξεις Τιμών της Μεταβλητής	Κεντρικές Τιμές των Τάξεων	Ταξικές Συχνότητες		Αθροιστικές Συχνότητες	
X	x_i	Απόλυτες f_i	Σχετικές f_i/N	Απόλυτες F_i	Σχετικές F_i/N
α_0 — α_1	x_1	f_1	f_1/N	F_1	F_1/N
α_0 — α_1	x_2	f_2	f_2/N	F_2	F_2/N
. — .	.	.	.	.	.
. — .	.	.	.	.	.
. — .	.	.	.	.	.
α_{i-1} — α_i	x_i	f_i	f_i/N	F_i	F_i/N
. — .	.	.	.	.	.
. — .	.	.	.	.	.
. — .	.	.	.	.	.
α_{k-1} — α_k	x_k	f_k	f_k/N	F_k	F_k/N

Στις επόμενες παραγράφους του εδαφίου αυτού γίνεται επεξήγηση των χρησιμοποιούμενων συμβόλων που υπεισέρχονται στο πίνακα 2.1.

[1] Στατιστική Μεταβλητή.

Οι στατιστικές μεταβλητές - κατηγορικές, ποιοτικές ή ποσοτικές - παριστάνονται με τα κεφαλαιά γράμματα. X, Y, Z, ...

[2] Συνολικό Εύρος Μεταβολής της Μεταβλητής.

Το διάστημα με άκρα α_0 και α_K - ανοικτό, ημιανοικτό, κλειστό - στο οποίο μεταβέλλεται η μεταβλητή X λέγεται συνολικό εύρος μεταβολής της X. Αν η X μπορεί να πάρει τις τιμές α_0 και α_K τότε οι τιμές αυτές λέγονται αντίστοιχα ελάχιστη και μέγιστη τιμή της μεταβλητής.

[3] Τάξεις Τιμών της Μεταβλητής.

Το συνολικό εύρος μεταβολής της μεταβλητής X διαιρείται σε κ υποδιαστήματα τα οποία λέγονται τάξεις. Η επιθυμητή απλότητα και η απαιτούμενη ομογένεια των δεδομένων καθορίζουν τον αριθμό κ των τάξεων. Η διαμόρφωση των τάξεων γίνεται κατά τέτοιο τρόπο ώστε κάθε τιμή της μεταβλητής να ανήκει σε μία και μόνο τάξη.

Οι αριθμοί α_{i-1} και α_i λέγονται άκρα της αντίστοιχης τάξης, i = 1,2,3,...,k. O αριθμός α_{i-1} λέγεται πρώτο άκρο και ο αριθμός α_i λέγεται δεύτερο άκρο της αντίστοιχης τάξης. Το ένα από τα άκρα κάθε τάξης - συνήθως το πρώτο - δεν ανήκει στη τάξη.

Δηλαδή οι τάξεις είναι ημιανοικτά διαστήματα.

[4] Πλάτη των Τάξεων.

Η διαφορά $\delta_i = \alpha_i - \alpha_{i-1}$ λέγεται πλάτος της αντίστοιχης τάξης, i=1,2,...,κ. Η κατανομή λέγεται ίσου πλάτους όταν όλες οι τάξεις έχουν το ίδιο πλάτος δηλαδή $\delta_i = \delta$, i =1,2,...κ. Όταν όλες οι τάξεις δεν έχουν το ίδιο πλάτος η κατανομή λέγεται άνισου πλάτους.

[5] Κεντρικές Τιμές των Τάξεων

Ο αριθμός

$$x_i = \frac{\alpha_{i-1} + \alpha_i}{2}$$

λέγεται κεντρική τιμή της αντίστοιχης τάξης, i =1,2,...κ.

[6] Απόλυτες Ταξικές Συχνότητες.

Η απόλυτη ταξική συχνότητα f_i συμβολίζει τον αριθμό των μονάδων του πληθυσμού των οποίων οι τιμές περιλαμβάνονται στην αντίστοιχη τάξη, $i=1,2,...к$.

[7] Απόλυτη Συνολική Συχνότητα.

Το άθροισμα

$$N = \sum_{i=1}^{k} f_i$$

λέγεται απόλυτη συνολική συχνότητα και είναι ίσο με το συνολικό αριθμό των μονάδων πληθυσμού.

[8] Σχετικές Ταξικές Συχνότητες.

Η σχετική ταξική συχνότητα f_i/N είναι το πηλίκο της αντίστοιχης απόλυτης ταξικής συχνότητας f_i δια της απόλυτης συνολικής συχνότητας N και συμβολίζει την αναλογία των μονάδων του πληθυσμού οι οποίες ανήκουν στην αντίστοιχη τάξη.

[9] Απόλυτη Αθροιστική Συχνότητα

Η απόλυτη αθροιστική συχνότητα

$$F_i = \sum_{j=1}^{i} f_j$$

$$=F_{i-1} + f_i, \quad i = 1,2,...,k$$

συμβολίζει τον αριθμό των μονάδων του πληθυσμού των οποίων οι τιμές

είναι μικρότερες ή ίσες του αριθμού a_i.

[10] Σχετικές Αθροιστικές Συχνότητες.

Η συνεχή αθροιστική συχνότητα F_i/N είναι το πηλίκο της αντίστοιχης απόλυτης αθροιστικής συχνότητας F_i δια της συνολικής απόλυτης συχνότητας N και συμβολίζει την αναλογία των μονάδων του πληθυσμού οι τιμές των οποίων είναι μικρότερες ή ίσες του αριθμού a_i.

ΚΕΦΑΛΑΙΟ 4

ΜΕΤΡΑ ΘΕΣΗΣ ΜΟΝΟΜΕΤΑΒΛΗΤΩΝ ΚΑΤΑΝΟΜΩΝ

[4.1] Κατηγορίες Μέτρων Θέσης.

Μία πολύ συνοπτική παρουσίαση ποσοτικών στατιστικών στοιχείων γίνεται συνήθως με τον υπολογισμό όρισμένων χαρακτηριστικών μεγεθών γνωστών ως παραμέτρων ή μέτρων. Οι παράμετροι ή τα μέτρα είναι αριθμητικές εκφράσεις που χαρακτηρίζουν τη θέση, διασπορά και την μορφή της κατανομής συχνότητας μιάς ποσοτικής στατιστικής μεταβλητής.

Θέση ή σημείο συγκέντρωσης μιάς ποσοτικής κατανομής συχνότητας είναι ένας αριθμός περί τον οποίον τείνουν να συγκεντρωθούν οι τιμές της αντίστοιχης ποσοτικής στατιστικής μεταβλητής. Σκοπός των παραμέτρων ή μέτρων θέσης είναι να καθορίσει μία θέση ή ισοδύναμα ένα σημείο συγκέντρωσης μιάς ποσοτικής κατανομής συχνότητας.

Το κεφάλαιο αυτό διαπραγματεύεται τα ακόλουθα μέτρα θέσης μονομεταβλητών ποσοτικών κατανομών συχνότητας :

[1] Διάμεσος.

[2] Τεταρτημόρια.

[3] Δεκατημόρια.

[4] Εκατοστημόρια.

[5] Αριθμητικός Μέσος Όρος.

[6] Σημείο Μέγιστης Συχνότητας.

[4.2] Διάμεσος Ποσοτικής Μεταβλητής.

Διάμεσος μιάς ποσοτικής μεταβλητής ή ισοδύναμα διάμεσος της αντίστοιχης κατανομής συχνότητας λέγεται μιά τιμή της μεταβλητής τέτοια ώστε N/2 από τις τιμές της μεταβλητής να είναι μικρότερες ή ίσες της τιμής αυτής.

Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού της διαμέσου μιάς ποσοτικής μεταβλητής.

[1] Διάμεσος Διακριτής Ποσοτικής Μεταβλητής.

[2] Διάμεσος Συνεχούς Ποσοτικής Μεταβλητής.

[4.3] Διάμεσος Διακριτής Ποσοτικής Μεταβλητής.

Ο υπολογισμός της διαμέσου M μιάς διακριτής ποσοτικής μεταβλητής X κάνει χρήση των ακόλουθων μεγεθών :

[1] N, το πλήθος των τιμών της μεταβλητής.

[2] F_{i-1}, F_i, οι διαδοχικές αθροιστικές συχνότητες μεταξύ των οποίων βρίσκεται ο αριθμός $\frac{N}{2}$. Δηλαδή

$$F_{i-1} < \frac{N}{2} < F_i \qquad (4.3.1)$$

Η διάμεσος M της διακριτής ποσοτικής μεταβλητής X είναι :

$$M = x_i \tag{4.3.2}$$

όπου x_i η τιμή της μεταβλητής η οποία αντιστοιχεί στην αθροιστική συχνότητα F_i.

Επειδή η τιμή xi της μεταβλητής X επαναλαμβάνεται fi φορές τότε ως διάμεσος λαμβάνεται η τιμή xi που κατέχει τη θέση τάξης $\frac{N}{2}$ - άν N είναι άρτιος - ή συμβατικά τη θέση τάξης $\frac{(N+1)}{2}$ - αν N είναι περιττός - όταν οι τιμές της X τοποθετούνται κατά αύξουσα τάξη μεγέθους.

Ο υπολογισμός της διαμέσου M της διακριτής μεταβλητής του πίνακα 3.2 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε $N = 4860$, $\frac{N}{2} = 2430$

και

$$F_2 = 1458 < 2430 < F_3 = 3159.$$

[2] Επομένως η διάμεσος M είναι $M = 3$.

Επειδή $N = 4860$ είναι άρτιος αριθμός η τιμή 3 που κατέχει την $\frac{N}{2} = 2430$ θέση τάξης είναι διάμεσος της μεταβλητής X του πίνακα 3.2.

[4.4] Διάμεσος Συνεχούς Ποσοτικής Μεταβλητής

Ο υπολογισμός της διαμέσου M μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται γενικά κατά προσέγγιση με τη βοήθεια γραμμικής παρεμβολής και κάνει χρήση των ακόλουθων μεγεθών :

[1] α_{i-1}, α_i, τα άκρα του διαστήματος στο οποίο εντοπίζεται η διάμεσος.

[2] δ_i, το μήκος του διαστήματος στο οποίο εντοπίζεται η διάμεσος. Δηλαδή $\delta_i = \alpha_i - \alpha_{i-1}$.

[3] f_i, η απόλυτη συχνότητα του διαστήματος στο οποίο εντοπίζεται η διάμεσος.

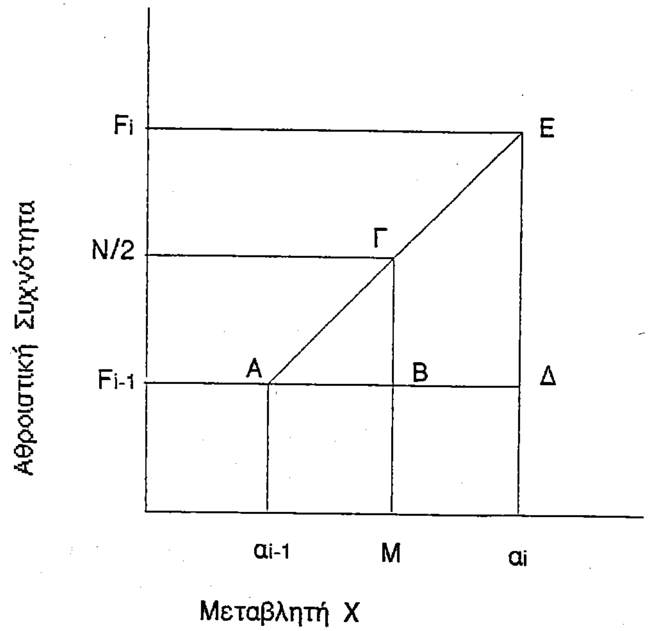
[4] N, το πλήθος των τιμών της μεταβλητής.

[5] F_{i-1}, η αθροιστική συχνότητα του διαστήματος που προηγείται εκείνου στο οποίο εντοπίζεται η διάμεσος.

[6] F_i, η αθροιστική συχνότητα του διαστήματος στο οποίο εντοπίζεται η διάμεσος.

Η χρήση γραμμικής παρεμβολής για τον υπολογισμό της διαμέσου M της μεταβλητής X υποθέτει ότι τα σημεία A, Γ, E του σχήματος 4.1 βρίσκονται σε ευθεία γραμμή.

Σχήμα 4.1



Από την ομοιότητα των ορθογωνίων τριγώνων ΑΒΓ και ΑΔΕ προκύπτει

$$\frac{M - \alpha_{i-1}}{\alpha_i - \alpha_{i-1}} = \frac{\frac{N}{2} - F_{i-1}}{F_i - F_{i-1}} \qquad (4.4.1)$$

Επειδή $\alpha_i - \alpha_{i-1} = \delta_i$ και $F_i - F_{i-1} = f_i$ από τη σχέση (4.4.1) προκύπτει ότι :

$$M = \alpha_{i-1} + \frac{\delta_i}{f_i}\left(\frac{N}{2} - F_{i-1}\right) \qquad (4.4.2)$$

Η εφαρμογή της (4.4.2) για τον υπολογισμό της διαμέσου Μ της κατανομής του πίνακα 3.3 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε $N = 1440$, $\frac{N}{2} = 720$

και

$F_1 = 216 < 720 \le F_2 = 720.$

[2] Επομένως η διάμεσος Μ περιλαμβάνεται στο δεύτερο διάστημα της κατανομής. Τότε έχουμε :

$i = 2$
$\alpha_1 = 30$
$\alpha_2 = 40$
$\delta_2 = 10$
$f_2 = 504$
$F_1 = 216$

[3] Επομένως από τη σχέση (4.4.2) προκύπτει ότι η διάμεσος Μ είναι :

$$M = 30 + \frac{10}{504}(720 - 216) = 40$$

[4.5] Τεταρτημόρια Ποσοτικής Μεταβλητής.

Πρώτο τεταρτημόριο μιάς ποσοτικής μεταβλητής λέγεται μία τιμή της μεταβλητής τέτοια ώστε $\frac{N}{4}$ από τις Ν τιμές της μεταβλητής να είναι μικρότερες ή ίσες της τιμής αυτής. Δεύτερο τεταρτημόριο λέγεται η διάμεσος της μεταβλητής. Τρίτο τεταρτημόριο μιάς ποσοτικής μεταβλητής λέγεται μιά τιμή της μεταβλητής τέτοια ώστε $\frac{3N}{4}$ από τις Ν τιμές της μεταβλητής να είναι μικρότερες ή ίσες της τιμής αυτής.

Επειδή ο υπολογισμός της διαμέσου έχει παρουσιαστεί στα εδάφια 4.3 και 4.4 στα επόμενα εδάφια θα γίνει παρουσίαση του υπολογισμού του πρώτου και τρίτου τεταρτημορίου.

Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού τεταρτημορίων μιάς ποσοτικής μεταβλητής.

[1] Πρώτο και Τρίτο Τεταρτημόριο Διακριτής Ποσοτικής Μεταβλητής.

[2] Πρώτο και Τρίτο Τεταρτημόριο Συνεχούς Ποσοτικής Μεταβλητής.

[4.6] <u>Πρώτο Τεταρτημόριο Διακριτής Ποσοτικής Μεταβλητής.</u>

Ο υπολογισμός του πρώτου τεταρτημορίου Q_1 μιάς διακριτής ποσοτικής μεταβολής Χ κάνει χρήση των ακόλουθων μεγεθών.:

[1] Ν, το πλήθος των τιμών της μεταβλητής.

[2] F_{i-1}, F_1, οι διαδοχικές αθροιστικές συχνότητες μεταξύ των οποίων βρίσκεται ο αριθμός $\frac{N}{4}$. Δηλαδή

$$F_{i-1} < \frac{N}{4} < F_i \qquad (4.6.1)$$

Το πρώτο τεταρτημόριο της διακριτής ποσοτικής μεταβλητής Χ είναι :

$$Q_1 = x_i \qquad (4.6.2)$$

όπου x_i η τιμή της μεταβλητής η οποία αντιστοιχεί στην αθροιστική συχνότητα F_i.

Ο υπολογισμός του πρώτου τεταρτημορίου Q_1 της διακριτής ποσοτικής μεταβλητής του πίνακα 3.2 γίνεται με το ακόλουθο τρόπο.

[1] Έχουμε $N = 4860$, $\frac{N}{4} = 1215$

και

$$F_1 = 243 < 1215 < F_2 = 1458.$$

[2] Επομένως το πρώτο τερτημόριο Q_1 είναι $Q_1 = 2$.

Επειδή $N = 4680$ είναι πολλαπλάσιο του 4 η τιμή 2 κατέχει την $\frac{4860}{4} = 1215$ θέση τάξης είναι το πρώτο τεταρτημόριο της μεταβλητής Χ του πίνακα 3.2.

Όταν το Ν δεν είναι πολλαπλάσιο του 4 η θέση τάξης του πρώτου τεταρτημορίου Q_1 μπορεί να ληφθεί - συμβατικά - ίση με $\left[\frac{N}{4}\right]$ ή $\left[\frac{N}{4}\right] + 1$.

[4.7] <u>Τρίτο Τεταρτημόριο Διακριτής Ποσοτικής Μεταβλητής.</u>

Ο υπολογισμός του τρίτου τεταρτημορίου Q_3 μιάς διακριτής ποσοτικής μεταβλητής Χ κάνει χρήση των ακόλουθων μεγεθών.

[1] Ν, το πλήθος των τιμών της μεταβλητής.

[2] F_{i-1}, F_i, οι διαδοχικές αθροιστικές συχνότητες μεταξύ των οποίων βρίσκεται ο αριθμός $\frac{3N}{2}$.. Δηλαδή

$$F_{i-1} < \frac{3N}{2} < F_i \qquad (4.7.1)$$

Το τρίτο τεταρτημόριο της διακριτής ποσοτικής μεταβλητής Χ είναι :

$$Q_3 = x_i \qquad (4.7.2)$$

όπου x_i η τιμή της μεταβλητής η οποία αντιστοιχεί στην αθροιστική συχνότητα F_i.

Ο υπολογισμός του τρίτου τεταρτημορίου Q_3 της διακριτής ποσοτικής μεταβλητής του πίνακα 3.2 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε $N = 4860$, $\frac{3N}{4} = 3645$

και

$F_3 = 3159 < 3645 < F_4 = 4131$.

[2] Επομένως το τρίτο τεταρτημόριο Q_3 είναι $Q_3 = 4$.

Επειδή N = 4860 είναι πολλαπλάσιο του 4 η τιμή 4 που κατέχει την $\frac{3.4860}{4} = 3645$ θέση τάξης είναι το τρίτο τερτημόριο της μεταβλητής X του πίνακα 3.2.

Όταν το N δεν είναι πολλαπλάσιο του 4 η θέση τάξης του τρίτου τεταρτημορίου Q3 μπορεί να ληφθεί - συμβατικά - ίση με $\left[\frac{3N}{4}\right]$ ή $\left[\frac{3N}{4}\right] + 1$.

[4.8] Πρώτο Τεταρτημόριο Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του πρώτου τεταρτημορίου Q_1 μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται γενικά κατά προσέγγιση με τη βοήθεια γραμμικής παρεμβολής και κάνει χρήση των ακόλουθων μεγεθών.

[1] α_i, α_{i-1}, τα άκρα του διαστήματος στο οποίο εντοπίζεται το πρώτο τεταρτημόριο.

[2] δ_i, το μήκος του διατήματος στο οποίο εντοπίζεται το πρώτο τεταρτημόριο. Δηλαδή $\delta_i = \alpha_i - \alpha_{i-1}$.

[3] f_i, η απόλυτη συχνότητα του διαστήματος στο οποίο εντοπίζεται το πρώτο τεταρτημόριο.

[4] N, το πλήθος των τιμών της μεταβλητής.

[5] F_{i-1}, η αθροιστική συχνότητα του διαστήματος που προηγείται εκείνου στο οποίο εντοπίζεται το πρώτο τεταρτημόριο.

[6] F_i, η αθροιστική συχνότητα του διαστήματος στο οποίο εντοπίζεται το πρώτο τεταρτημόριο.

Η χρήση γραμμικής παρεμβολής για τον υπολογισμό του πρώτου τεταρτημορίου Q_1 μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται κατά τρόπο ανάλογο της χρήσης γραμμικής παρεμβολής για τον υπολογισμό της διαμέσου M της X. Επομένως το πρώτο τεταρτημόριο Q_1 της X είναι :

$$Q_1 = \alpha_{i-1} + \frac{\delta_i}{f_i}\left(\frac{N}{4} - F_{i-1}\right) \qquad (4.8.1)$$

Η εφαρμογή της (4.8.1) για τον υπολογισμό του πρώτου τεταρτημορίου Q_1 της κατανομής του πίνακα 3.3 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε N = 1440, $\frac{N}{4} = 360$

και

$F_1 = 216 < 360 < F_2 = 720$.

[2] Επομένως το πρώτο τεταρτημόριο Q_1 περιλαμβάνεται στο δεύτερο διάστημα της κατανομής. Τότε έχουμε :

$i = 2$
$\alpha_1 = 30$
$\alpha_2 = 40$
$\delta_2 = 10$
$f_2 = 504$
$F_1 = 216$

[3] Επομένως από τη σχέση (4.8.1) προκύπτει ότι το πρώτο τεταρτημόριο Q_1 είναι :

$$Q_1 = 30 + \frac{10}{504}(360 - 216) = 32,8$$

[4.9] Τρίτο Τεταρτημόριο Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του τρίτου τεταρτημορίου Q_3 μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται κατά ανάλογο τρόπο προς τον υπολογισμό του πρώτου τεταρτημορίου Q_1 και κάνει χρήση των ακόλουθων μεγεθών.

[1] α_i, α_{i-1}, τα άκρα του διαστήματος στο οποίο εντοπίζεται το τρίτο τεταρτημόριο.

[2] δ_i, το μήκος του διαστήματος στο οποίο εντοπίζεται το τρίτο τεταρτημόριο. Δηλαδή $\delta_i = \alpha_i - \alpha_{i-1}$.

[3] f_i, η απόλυτη συχνότητα του διαστήματος στο οποίο εντοπίζεται το τρίτο τεταρτημόριο.

[4] N, το πλήθος των τιμών της μεταβλητής.

[5] F_{i-1}, η αθροιστική συχνότητα του διαστήματος που προηγείται εκείνου στο οποίο εντοπίζεται το τρίτο τεταρτημόριο.

[6] F_i, η αθροιστική συχνότητα του διαστήματος στο οποίο εντοπίζεται το τρίτο τεταρτημόριο.

Επομένως το τρίτο τεταρτημόριο Q_3 της X είναι :

$$Q_3 = \alpha_{i-1} + \frac{\delta_i}{f_2}\left(\frac{3N}{4} - F_{i-1}\right) \tag{4.9.1}$$

Η εφαρμογή της (4.9.1) για τον υπολογισμό του τρίτου τεταρτημορίου Q_3 της κατανομής του πίνακα 3.3 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε $N = 1440$, $\frac{3N}{4} = 1080$

και

$$F_2 = 720 < 1080 \leq F_3 = 1080.$$

[2] Επομένως το τρίτο τεταρτημόριο Q_3 περιλαμβάνεται στο τρίτο διάστημα της κατανομής. Τότε έχουμε :

$i = 3$
$\alpha_2 = 40$
$\alpha_3 = 50$
$\delta_3 = 10$
$f_3 = 360$
$F_2 = 720$

[3] Επομένως από τη σχέση (4.9.1) προκύπτει ότι το τρίτο τεταρτημόριο Q_3 είναι :

$$Q_1 = 40 + \frac{10}{360}(1080 - 720) = 50$$

[4.10] Δεκατημόρια Ποσοτικής Μεταβλητής.

Το k - Δεκατημόριο κ = 1,2,...,9 μιάς ποσοτικής μεταβλητής λέγεται μία τιμή της μεταβλητής τέτοια ώστε $\frac{kN}{10}$ από τις N τιμές της μεταβλητής να είναι μικρότερες ή ίσες της τιμής αυτής.

Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού δεκατημορίων μιάς ποσοτικής μεταβλητής.

[1] κ - Δεκατημόρια Διακριτής Ποσοτικής Μεταβλητής.

[2] κ - Δεκατημόριο Συνεχούς Ποσοτικής Μεταβλητής.

[4.11] κ - Δεκατημόριο Διακριτής Ποσοτικής Μεταβλητής.

Ο υπολογισμός του κ - δεκατημορίου D_κ, κ =1,2,...,9 μιάς διακριτής ποσοτικής μεταβολής X κάνει χρήση των ακόλουθων μεγεθών.

[1] N, το πλήθος των τιμών της μεταβλητής.

[2] F_{i-1}, F_i, οι διαδοχικές αθροιστικές συχνότητες μεταξύ των οποίων βρίσκεται ο αριθμός κN/10. Δηλαδή

$$F_{i-1} < \frac{kN}{10} < F_i \qquad (4.11.1)$$

Το κ - δεκατημόριο της διακριτής ποσοτικής μεταβλητής X είναι :

$$D_\kappa = x_i \qquad (4.11.2)$$

όπου x_i η τιμή της μεταβλητής η οποία αντιστοιχεί στην αθροιστική συχνότητα F_i.

Ο υπολογισμός του κ - δεκατημορίου D_κ, k = 1,2,...,9 της διακριτής ποσοτικής μεταβλητής του πίνακα 3.2 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε N = 4860, $\frac{N}{10} = 486$

και

$F_1 = 243 < 486 < F_2 = 1458.$

[2] Επομένως το πρώτο δεκατημόριο D_1 είναι $D_1 = 2$.

Επειδή N = 4860 είναι πολλαπλάσιο του 10 η τιμή 2 που κατέχει την $\frac{4860}{10} = 486$ θέση τάξης είναι το πρώτο δεκατημόριο της μεταβλητής X του πίνακα 3.2.

Όταν το N δεν είναι πολλαπλάσιο του 10 η θέση τάξης του κ - δεκατημορίου Dκ, κ=1,2,...,9 μπορεί να ληφθεί - συμβατικά - ίση με $\left[\frac{kN}{10}\right]$ ή $\left[\frac{kN}{10}\right] + 1$.

Αντίστοιχα υπολογίζονται τα άλλα δεκατημόρια.

[4.12] κ - Δεκατημόριο Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του κ - δεκατημορίου D_k, k=1,2,...,9 μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται γενικά κατά προσέγγιση με τη βοήθεια γραμμικής παρεμβολής και κάνει χρήση των ακόλουθων μεγεθών.

[1] α_i, α_{i-1}, τα άκρα του διαστήματος στο οποίο εντοπίζεται το κ - δεκατημόριο.

[2] δ_i, το μήκος του διαστήματος στο οποίο εντοπίζεται το κ - δεκατημόριο. Δηλαδή $\delta_i = \alpha_i - \alpha_{i-1}$.

[3] f_i, η απόλυτη συχνότητα του διαστήματος στο οποίο εντοπίζεται το κ - δεκατημόριο.

[4] N, το πλήθος των τιμών της μεταβλητής.

[5] F_{i-1}, η αθροιστική συχνότητα του διαστήματος που προηγείται εκείνου στο οποίο εντοπίζεται το κ - δεκατημόριο.

[6] F_i, η αθροιστική συχνότητα του διαστήματος στο οποίο εντοπίζεται το κ - δεκατημόριο.

Η χρήση γραμμικής παρεμβολής για τον υπολογισμό του κ - δεκατημορίου D_κ, κ=1,2,...,9 μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται κατά τρόπο ανάλογο της χρήσης γραμμικής παρεμβολής για τον υπολογισμό της διαμέσου M και των τεταρτημορίων της X.

Επομένως το κ - δεκατημόριο $D_κ$ της X είναι :

$$D_κ = α_{i-1} + \frac{δ_i}{f_i}\left(\frac{kN}{10} - F_{i-1}\right) \qquad (4.12.1)$$

Η εφαρμογή της (4.12.1) για τον υπολογισμό του πρώτου δεκατημορίου D_1 της κατανομής του πίνακα 3.3 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε $N = 1440$, $\frac{N}{10} = 144$

και

$F_0 = 0 < 144 < F_1 = 216.$

[2] Επομένως το πρώτο δεκατημόριο D_1 περιλαμβάνεται στο πρώτο διάστημα της κατανομής. Τότε έχουμε :

$i = 1$
$α_1 = 20$
$α_2 = 30$
$δ_1 = 10$
$f_1 = 216$
$F_1 = 0$

[3] Επομένως από τη σχέση (4.12.1) προκύπτει ότι το πρώτο δεκαρτημόριο D_1 είναι :

$$D_1 = 20 + \frac{10}{216}(144 - 0) = 26{,}6$$

Αντίστοιχα υπολογίζονται τα άλλα δεκατημόρια.

[4.13] Εκατοστημόρια Ποσοτικής Μεταβλητής.

Το k - εκατοστημόριο k = 1,2,...,99 μιάς ποσοτικής μεταβλητής λέγεται μία τιμή της μεταβλητής τέτοια ώστε $\frac{kN}{100}$ από τις N τιμές της μεταβλητής να είναι μικρότερες ή ίσες της τιμής αυτής.

Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού εκατοστημορίων μιάς ποσοτικής μεταβλητής.

[1] κ - Εκατοστημόριο Διακριτής Ποσοτικής Μεταβλητής.

[2] κ - Εκατοστημόριο Συνεχούς Ποσοτικής Μεταβλητής.

[4.14] κ - Εκατοστημόριο Διακριτής Ποσοτικής Μεταβλητής.

Ο υπολογισμός του κ - εκατοστημορίου $ρ_κ$, κ =1,2,...,99 μιάς διακριτής ποσοτικής μεταβολής X κάνει χρήση των ακόλουθων μεγεθών.

[1] N, το πλήθος των τιμών της μεταβλητής.

[2] F_{i-1}, F_i, οι διαδοχικές αθροιστικές συχνότητες μεταξύ των οποίων βρίσκεται ο αριθμός $\frac{kN}{100}$. Δηλαδή

$$F_{i-1} < \frac{kN}{100} < F_i \qquad (4.14.1)$$

Το κ - εκατοστημόριο της διακριτής ποσοτικής μεταβλητής X είναι :

$$P_κ = x_i \qquad (4.14.2)$$

όπου x_i η τιμή της μεταβλητής η οποία αντιστοιχεί στην αθροιστική συχνότητα F_i.

Ο υπολογισμός του κ - εκατοστημορίου P_κ, κ=1,2,...,99 της διακριτής ποσοτικής μεταβλητής του πίνακα 3.2 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε $N = 4860$, $\frac{N}{100} = 48{,}6$
και

$F_0 = 0 < 48{,}6 < F_1 = 243$.

[2] Επομένως το πρώτο εκατοστημόριο P_1 είναι $P_1 = 2$.

Επειδή N = 4860 είναι πολλαπλάσιο του 100 η θέση τάξης του πρώτου εκατοστημορίου P_1 μπορεί να ληφθεί - συμβατικά - ίση με 48 ή 49.

Αντίστοιχα υπολογίζονται τα άλλα εκατοστημόρια.

[4.15] <u>κ - Εκατοστημόριο Συνεχούς Ποσοτικής Μεταβλητής.</u>

Ο υπολογισμός του κ - εκατοστημορίου P_κ, κ=1,2,...,99 μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται γενικά κατά προσέγγιση με τη βοήθεια γραμμικής παρεμβολής και κάνει χρήση των ακόλουθων μεγεθών.

[1] α_i, α_{i-1}, τα άκρα του διαστήματος στο οποίο εντοπίζεται το κ - εκατοστημόριο.

[2] δ_i, το μήκος του διαστήματος στο οποίο εντοπίζεται το κ - εκατοστημόριο. Δηλαδή $\delta_i = \alpha_i - \alpha_{i-1}$.

[3] f_i, η απόλυτη συχνότητα του διαστήματος στο οποίο εντοπίζεται το κ - εκατοστημόριο.

[4] N, το πλήθος των τιμών της μεταβλητής.

[5] F_{i-1}, η αθροιστική συχνότητα του διαστήματος που προηγείται εκείνου στο οποίο εντοπίζεται το κ - εκατοστημόριο.

[6] F_i, η αθροιστική συχνότητα του διαστήματος στο οποίο εντοπίζεται το κ - εκατοστημόριο.

Η χρήση γραμμικής παρεμβολής για τον υπολογισμό του κ - εκατοστημορίου P_κ, κ=1,2,...,99 μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται κατά τρόπο ανάλογο της χρήσης γραμμικής παρεμβολής για τον υπολογισμό της διαμέσου M, των τεταρτημορίων και των δεκατημορίων της X.

Επομένως το κ - εκατοστημόριο P_κ της X είναι :

$$P_k = \alpha_{i-1} + \frac{\delta_i}{f_i}\left(\frac{kN}{100} - F_{i-1}\right) \qquad (4.15.1)$$

Η εφαρμογή της (4.15.1) για τον υπολογισμό του πρώτου εκατοστημορίου P_1 της κατανομής του πίνακα 3.3 γίνεται με τον ακόλουθο τρόπο.

[1] Έχουμε $N = 1440$, $\frac{N}{100} = 14{,}4$

και

$F_0 = 0 < 14{,}4 < F_1 = 216$.

[2] Επομένως το πρώτο εκατοστημόριο P_1 περιλαμβάνεται στο πρώτο διάστημα της κατανομής. Τότε έχουμε :

$i = 1$
$\alpha_1 = 20$
$\alpha_2 = 30$
$\delta_1 = 10$
$f_1 = 216$
$F_1 = 0$

[3] Επομένως από τη σχέση (4.15.1) προκύπτει ότι το πρώτο εκατοστημόριο P_1 είναι :

$$P_1 = 20 + \frac{10}{216}(14,4 - 0)$$
$$= 20,6$$

Αντίστοιχα υπολογίζονται τα άλλα εκατοστημόρια.

[4.16] Αριθμητικός Μέσος Όρος Ποσοτικής Μεταβλητής.

Το σπουδαιότερο και συνηθέστερα χρησιμοποιούμενο στη πράξη μέτρο θέσης μιάς ποσοτικής μεταβλητής είναι ο αριθμητικός μέσος όρος ή μέσος αριθμητικός.

Τα επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού του αριθμητικού μέσου όρου ή μέσου αριθμητικού μιάς ποσοτικής μεταβλητής.

[1] Αριθμητικός Μέσος Όρος Διακριτής Ποσοτικής Μεταβλητής.

[2] Αριθμητικός Μέσος Όρος Συνεχούς Ποσοτικής Μεταβλητής.

[4.17] Αριθμητικός Μέσος Όρος Διακριτής Ποσοτικής Μεταβλητής.

Αριθμητικός μέσος όρος ή μέσος αριθμητικός μιάς διακριτής ποσοτικής μεταβλητής X λέγεται το πηλίκο του αλγεβρικού αθροίσματος των τιμών της μεταβλητής δια του αθροίσματος των τιμών της μεταβλητής. Ο μέσος αριμθητικός συμβολίζεται συνήθως με το γράμμα μ. Επομένως

$$\mu = \frac{1}{N} \sum_{i=1}^{\kappa} f_i x_i \qquad (4.17.1)$$

όπου

[1] x_i, $i=1,2,\ldots,\kappa$ οι τιμές της μεταβλητής X,

[2] f_i, $i=1,2,\ldots,\kappa$ οι αντίσοιχες απόλυτες συχνότητας των τιμών της μεταβλητής X.

[3] $N = \sum_{i=1}^{\kappa} f_i$ το πλήθος των τιμών της μεταβλητής X.

Ένα παράδειγμα υπολογισμού του μέσου αριθμητικού μιάς διακριτής ποσοτικής μεταβλητής είναι το ακόλουθο.

Πίνακας 4.1

Κατανομή Συχνότητας του Αριθμού των Τυπογραφικών Λαθών στις Σελίδες Ενός Βιβλίου

Αριθμός Τυπογραφικών Λαθών	Αριθμός Σελίδων
0	60
1	20
2	12
3	9
4	5
5	6
6	8

Ο υπολογισμός του μέσου αριθμητικού μ του αριθμού των τυπογραφικών λαθών X γίνεται με τη βοήθεια του Πίνακα 4.2

Πίνακας 4.2

Τιμές της X	Απόλυτες Συχνότητες	Μερικά Γινόμενα
xi	fi	fi xi
0	60	0
1	20	20
2	12	24
3	9	27
4	5	20
5	6	30
6	8	48
	120	169

Από τον πίνακα 4.2 προκύπτει ότι

$\kappa = 7$

$$N = \sum_{i=1}^{\kappa} f_i$$

$$= 120$$

$$\sum_{i=1}^{\kappa} f_i x_i = 169$$

Επομένως ο μέσος αριθμητικός μ του αριθμού των τυπογραφικών λαθών X είναι :

$$\mu = \frac{1}{N} \sum_{i=1}^{\kappa} f_i x_i$$

$$= \frac{169}{120}$$

$$= 1{,}4$$

[4.18] Αριθμητικός Μέσος Όρος Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του μέσου αριθμητικού μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται με τη χρήση του τύπου

$$\mu = \frac{1}{N} \sum_{i=1}^{\kappa} f_i x_i \qquad (4.18.1)$$

όπου

[1] x_i, $i=1,2,\ldots,\kappa$ οι κεντρικές τιμές των τάξεων της μεταβλητής X,

[2] f_i, $i=1,2,\ldots,\kappa$ οι αντίστοιχες απόλυτες συχνότητας των τάξεων της μεταβλητής X.

[3] $N = \sum_{i=1}^{\kappa} f_i$ το πλήθος των τιμών της μεταβλητής X.

Ένα παράδειγμα υπολογισμού του μέσου αριθμητικού μιάς συνεχούς ποσοτικής μεταβλητής είναι το ακόλουθο.

Πίνακας 4.3

Κατανομή Συχνότητας του Βάρους των Μαθητών Ενός Σχολείου

Βάρος Μαθητών	Αριθμός Μαθητών
20 - 30	30
30 - 40	40
40 - 50	35
50 - 60	25
60 - 70	15

Ο υπολογισμός του μέσου αριθμητικού μ του βάρους των μαθητών X γίνεται με τη βοήθεια του πίνακα 4.3

Πίνακας 4.3

Τάξεις της X	Κεντρικές Τιμές των Τάξεων	Απόλυτες Συχνότητες	Μερικά Γινόμενα
	xi	fi	fi xi
20 - 30	25	30	750
30 - 40	35	40	1400
40 - 50	45	35	1575
50 - 60	55	25	1375
60 - 70	65	15	975
		145	6075

Από τον πίνακα 4.3 προκύπτει ότι

$κ = 5$

$$N = \sum_{i=1}^{κ} f_i$$

$$= 145$$

$$\sum_{i=1}^{κ} f_i x_i = 6075$$

Επομένως ο μέσος αριθμητικός μ του βάρους των μαθητών X είναι

$$μ = \frac{1}{N} \sum_{i=1}^{κ} f_i x_i$$

$$= \frac{6075}{145}$$

$$= 41{,}8$$

[4.19] Ιδιότητες Αριθμητικού Μέσου Όρου.

Το εδάφιο αυτό διαπραγματεύεται μερικές ιδιότητες του αριθμητικού μέσου όρου μιάς ποσοτικής - διακριτής ή συνεχούς - μεταβλητής. Οι ιδιότητες αυτές παρουσιάζουν θεωρητικό και πρακτικό ενδιαφέρον.

Ιδιότητα 1

Αν η ποσοτική μεταβλητή X είναι μία σταθερά, δηλαδή όλες οι τιμές της μεταβλητής είναι ίσες με α τότε ο μέσος αριθμητικός μ είναι :

$$μ = α \quad (4.19.1)$$

Απόδειξη

Από τον τύπο (4.17.1) ή τον τύπο (4.18.1) έχουμε

$$μ = \frac{1}{N} \sum_{i=1}^{κ} f_i x_i$$

$$= \frac{1}{N} \sum_{i=1}^{κ} f_i α$$

$$= \frac{1}{N} \sum_{i=1}^{κ} f_i \quad (4.19.2)$$

Επειδή

$$N = \sum_{i=1}^{κ} f_i$$

Τότε από την (4.19.2) έχουμε

$$\mu = \frac{aN}{N}$$
$$= a$$

Ιδιότητα 2

Έστω ℓ η ελάχιστη τιμή, L η μέγιστη τιμή και μ ο μέσος αριθμητικός μιάς ποσοτικής μεταβλητής X. Τότε

$$\ell \leq \mu \leq L \qquad (4.19.3)$$

Απόδειξη

Υποθέτουμε ότι η μεταβλητή X είναι διακριτή με τιμές x_i, i = 1,2,3,...,κ και αντίστοιχες συχνότητες f_i, i = 1,2,...,κ

Επειδή

$$\begin{array}{l} \ell \leq x_1 \leq L \\ \ell \leq x_2 \leq L \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \ell \leq x_\kappa \leq L \end{array} \qquad (4.19.4)$$

Τότε

$$\begin{array}{l} \ell f_1 \leq x_1 f_1 \leq L f_1 \\ \ell f_2 \leq x_2 f_2 \leq L f_2 \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \ell f_k \leq x_\kappa f_k \leq L f_k \end{array} \qquad (4.19.5)$$

Αν προσθέσουμε κατά μέλη τις ανισότητες (4.19.5) έχουμε

$$\ell\ (f_1+f_2+...+f_k) \leq x_1f_1 + x_2f_2 + ... + x_\kappa f_k \leq L\ (f_1+f_2+...+f_k)$$

ή ισοδύναμα

$$\ell \sum_{i=1}^{\kappa} f_i \leq \sum_{i=1}^{k} x_i f_i \leq L \sum_{i=1}^{k} f_i \qquad (4.19.6)$$

Αν διαιρέσουμε τις ανισότητες (4.19.6) με

$$N = \sum_{i=1}^{\kappa} f_i$$

Τότε προκύπτει

$$\ell \leq \frac{1}{N} \sum_{i=1}^{\kappa} f_i x_i \leq L$$

Επομένως

$$\ell \leq \mu \leq L$$

Αν η μεταβλητή είναι συνεχής τότε οι x_i, i = 1,2,3,...,k είναι οι κεντρικές τιμές των τάξεων της μεταβλητής και η απόδειξη της (4.19.3) γίνεται με τον ίδιο τρόπο.

Ιδιότητα 3.

Το άθροισμα των διαφορών των τιμών μιάς ποσοτικής μεταβλητής από τον μέσο αριθμητικό μ είναι ίσο με 0.

Απόδειξη.

Υποθέτουμε ότι η μεταβλητή X είναι διακριτή με τιμές x_i, i=1,2,...,κ και αντίστοιχες συχνότητες f_i, i = 1,2,,...,k και θα αποδείξουμε ότι :

$$\sum_{i=1}^{\kappa} f_i(x_i-\mu) = 0 \qquad (4.19.7)$$

Έχουμε

$$\sum_{i=1}^{\kappa} f_i (x_i-\mu) = \sum_{i=1}^{\kappa} f_i x_i - \mu \sum_{i=1}^{\kappa} f_i \qquad (4.19.8)$$

Επειδή

$$N = \sum_{i=1}^{\kappa} f_i$$

και

$$\mu = \frac{1}{N} \sum_{i=1}^{\kappa} f_i x_i$$

τότε από την (4.19.8) προκύπτει ότι

$$\sum_{i=1}^{\kappa} f_i (x_i-\mu) = \mu N - \mu N$$

$$= 0$$

Αν η μεταβλητή είναι συνεχής τότε οι x_i, $i = 1,2,\ldots,k$ είναι οι κεντρικές τιμές των τάξεων της μεταβλητής και η απόδειξη της (4.19.7) γίνεται με τον ίδιο τρόπο.

Ιδιότητα 4.

Έστω X ποσοτική μεταβλητή με μέσο αριθμητικό μ_x. Θεωρούμε την ποσοτική μεταβλητή

$$Y = \alpha + \beta X \qquad (4.19.9)$$

όπου α,β πραγματικοί αριθμοί. Ο μέσος αριθμητικός μ_y της Y είναι

$$\mu_y = \alpha + \beta\mu_x \qquad (4.19.10)$$

Απόδειξη.

Υποθέτουμε ότι η μεταβλητή X είναι διακριτή με τιμές x_i, $i=1,2,\ldots,k$ και αντίστοιχες συχνότητες f_i, $i=1,2,\ldots,k$. Οι τιμές της Y είναι $y_i = \alpha + \beta x_i$ με αντίστοιχες συχνότητες f_i, $i = 1,2,\ldots,k$. Έχουμε

$$\mu_y = \frac{\sum_{i=1}^{\kappa} f_i y_i}{\sum_{i=1}^{\kappa} f_i}$$

$$= \frac{\sum_{i=1}^{\kappa} f_i(\alpha + \beta x_i)}{\sum_{i=1}^{\kappa} f_i}$$

$$= \frac{\sum_{i=1}^{\kappa} f_i \alpha + \sum_{i=1}^{\kappa} f_i \beta x_i}{\sum_{i=1}^{\kappa} f_i}$$

$$= \frac{\alpha\sum_{i=1}^{\kappa} f_i + \beta\sum_{i=1}^{\kappa} f_i x_i}{\sum_{i=1}^{\kappa} f_i} \qquad (4.19.11)$$

Επειδή

$$N = \sum_{i=1}^{\kappa} f_i$$

και

$$\mu_X N = \sum_{i=1}^{\kappa} f_i x_i$$

τότε η (4.19.11) γράφεται

$$\mu_y = \frac{\alpha N + \beta\mu_x N}{N}$$

$$= \frac{\alpha N}{N} + \frac{\beta \mu x N}{N}$$

$$= \alpha + \beta\mu_x$$

Αν η μεταβλητή είναι συνεχής τότε οι x_i, $i = 1,2,...,k$ είναι οι κεντρικές τιμές των τάξεων της μεταβλητής και η απόδειξη της (4.19.9) γίνεται με τον ίδιο τρόπο.

[4.20] Σημείο Μέγιστης Συχνότητας μιας Ποσοτικής Μεταβλητής.

Σημείο μέγιστης συχνότητας μιας ποσοτικής μεταβλητής λέγεται η τιμή της μεταβλητής η οποία έχει την μέγιστη συχνότητα. Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού του σημείου μέγιστης συχνότητας μιας ποσοτικής μεταβλητής.

[1] Σημείο Μέγιστης Συχνότητας Διακριτής Ποσοτικής Μεταβλητής.

[2] Σημείο Μέγιστης Συχνότητας Συνεχούς Ποσοτικής Μεταβλητής.

[4.21] Σημείο Μέγιστης Συχνότητας Διακριτής Ποσοτικής Μεταβλητής.

Ο Προσδιορισμός του σημείου μέγιστης συχνότητας Τ μιας διακριτής ποσοτικής μεταβλητής είναι απλούστατος και γίνεται με τη βοήθεια της κατανομής συχνότητας της μεταβλητής. Ένα παράδειγμα προσδιορισμού του σημείου μεγίστης συχνότητας μιάς διακριτής ποσοτικής μεταβλητής είναι το ακόλουθο.

Πίνακας 4.4

Κατανομή Συχνότητας του Αριθμού των Ορόφων των Κατοικιών μιάς Πόλης.

Αριθμός Ορόφων xi	Αριθμός Κατοίκων fi
1	520
2	1662
3	2302
4	1299
5	950
	6733

Η μέγιστη συχνότητα της κατανομής συχνότητας του πίνακα 4.4 είναι 2302. Επομένως το σημείο μέγιστης συχνότητας του αριθμού των ορόφων Χ είναι Τ = 3.

Όταν η μέγιστή συχνότητα της κατανομής συχνότητας μιας ποσοτικής διακριτής μεταβλητής Χ είναι μοναδική τότε η κατανομμή έχει ένα σημείο μέγιστης συχνότητας και λέγεται μονοκόρυφη. Όταν η μέγιστη συχνότητα της κατανομής συχνότητας της Χ δεν είναι μοναδική τότε η κατανομή έχει πολλά σημεία μέγιστης συχνότητας και λέγεται πολυκόρυφη.

[4.22] Σημείο Μέγιστης Συχνότητας Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του σημείου μέγιστης συχνότητας Τ μιας συνεχούς ποσοτικής μεταβλητής Χ γίνεται κατά προσέγγιση και κάνει χρήση των ακόλουθων μεγεθών.

[1] α_{i-1}, α_i τα άκρα του διαστήματος το οποίο έχει τη μέγιστη συχνότητα.

[2] δ_i, το μήκος του διαστήματος το οποίο έχει τη μέγιστη συχνότητα.

Δηλαδή $\delta_i = \alpha_i - \alpha_{i-1}$.

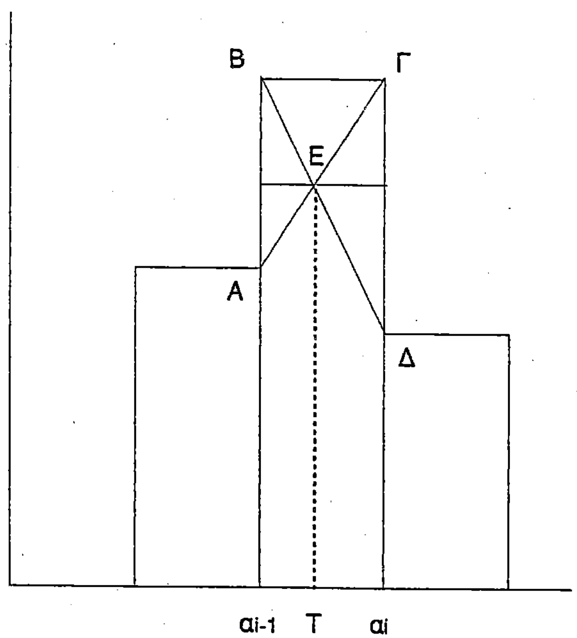
[3] Δ_1, η διαφορά μεταξύ της μέγιστης συχνότητας και της προηγούμενης συχνότητας. Δηλαδή $\Delta_1 = f_i - f_{i-1}$.

[4] Δ_2, η διαφορά μεταξύ της μέγιστης συχνότητας και της επόμενης συχνότητας. Δηλαδή $\Delta_2 = f_i - f_{i+1}$.

Ο προσεγγιστικός προσδιορισμός του σημείου μέγιστης συχνότητας Τ μιας συνεχούς ποσοτικής μεταβλητής Χ βασίζεται στην υπόθεση ότι οι αποστάσεις του σημείου Τ από τα όρια α_{i-1} και α_i είναι ανάλογες των διαφορών Δ_1 και Δ_2.

Ο προσεγγιστικός προσδιορισμός του σημείου Τ παρουσιάζεται γραφικά στο σχήμα 4.2.

Σχήμα 4.2



Τα τρίγωνα ΑΒΕ, ΓΕΔ είναι όμοια και ΕΖ και ΕΗ είναι ύψη των

τριγώνων αυτών.

Έχουμε

$$\frac{AB}{\Gamma\Delta} = \frac{EZ}{EH} \qquad (4.22.1)$$

Επειδή

$$EZ = T - \alpha_{i-1}$$
$$EH = \alpha_i - T$$
$$AB = \Delta_1$$
$$\Gamma\Delta = \Delta_2$$
$$\delta_i = \alpha_i - \alpha_{i-1}$$

Τότε η σχέση γράφεται

$$\frac{\Delta_1}{\Delta_2} = \frac{T - \alpha_{i-1}}{\alpha_2 - T}$$

ή ισοδύναμα

$$\frac{\Delta_1}{\Delta_2} = \frac{T - \alpha_{i-1}}{\alpha_i + \delta_i - T} \qquad (4.22.2)$$

Επομένως

$$T = \alpha_{i-1} + \delta_i \frac{\Delta_1}{\Delta_1 + \Delta_2} \qquad (4.22.3)$$

Ο τύπος (4.22.3) εφαρμόζεται μόνο όταν η συνεχής ποσοτική μεταβλητή Χ έχει ένα σημείο μέγιστης συχνότητας.

[4.23] <u>Χαρακτηριστικά των Μέτρων Θέσης</u>

Βασικά χαρακτηριστικά των μέτρων θέσης είναι τα ακόλουθα :

[1] Η διάμεσος, τα τεταρτημόρια, δεκατημόρια και εκατοστημόρια το σημείο μέγιστης συχνότητας δεν επηρεάζονται από τις ακραίες τιμές της ποσοτικής μεταβλητής.

[2] Η σημασία της διαμέσου, των τεταρτημορίων, των δεκατημορίων και των εκατοστημορίων είναι πολύ σημαντική σε θέματα της οικονομικής επιστήμης.

[3] Η διάμεσος, τα τερτημόρια, δεκατημόρια και τα εκατοστημόρια μιάς ποσοτικής μεταβλητής μπορούν να υπολογιστούν και γραφικά με τη χρήση της γραφικής απεικόνισης της αντίστοιχης αθροιστικής κατανομής.

[4] Ο αριθμητικός μέσος επηρεάζεται από κάθε τιμή της ποσοτικής μεταβλητής.

ΚΕΦΑΛΑΙΟ 5

ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ ΜΟΝΟΜΕΤΑΒΛΗΤΩΝ ΚΑΤΑΝΟΜΩΝ

[5.1] <u>Κατηγορίες Μέτρων Διασποράς.</u>

Τα μέτρα διασποράς μιας μονομεταβλητής κατανομής συχνότητας χαρακτηρίζουν την διασπορά των τιμών της αντίστοιχης ποσοτικής στατιστικής μεταβλητής. Το κεφάλαιο αυτό διαπραγματεύεται τα ακόλουθα μέτρα διασποράς μονομεταβλητών ποσοτικών κατανομών συχνότητας.

[1] Εύρος Μεταβολής.

[2] Ημιενδοτεταρτημοριακό Εύρος.

[3] Μέση Απόλυτη Απόκλιση

[4] Διακύμανση.

[5] Τυπική Απόκλιση.

[6] Συντελεστής Μεταβλικότητας.

[5.2] <u>Εύρος Μεταβολής Ποσοτικής Μεταβλητής.</u>

Το εύρος μεταβολής είναι το απλούστερο μέτρο διασποράς των τιμών μιάς ποσοτικής μεταβλητής και ορίζεται ως η διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής της μεταβλητής. Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού του εύρους μεταβολής μιας ποσοτικής μεταβλητής.

[1] Εύρος Μεταβολής Διακριτής Ποσοτικής Μεταβλητής.

[2] Εύρος Μεταβολής Συνεχούς Ποσοτικής Μεταβλητής.

[5.3] Εύρος Μεταβολής Διακριτής Ποσοτικής Μεταβλητής.

Έστω X ποσοτική διακριτή μεταβλητή με τιμές x_i, $i=1,2,...,k$. Το εύρος μεταβολής R της X είναι :

$$R = \max\{x_1,x_2,...,x_k\} - \min\{x_1,x_2,...,x_k\} \quad (5.3.1)$$

Το εύρος μεταβολής R της ποσοτικής διακριτής μεταβλητής X του πίνακα 4.1 είναι :

$$R = \max\{0,1,2,3,4,5,6\} - \min\{0,1,2,3,4,5,6\}$$
$$= 6 - 0$$
$$= 6.$$

[5.4] Εύρος Μεταβολής Συνεχούς Ποσοτικής Μεταβλητής.

Έστω X συνεχής ποσοτική μεταβλητή με πρώτο άκρο του πρώτου διαστήματος α_0 και δεύτερο άκρο του τελευταίου διαστήματος α_κ. Το εύρος μεταβολής R της X είναι

$$R = \alpha_\kappa - \alpha_0 \quad (5.4.1)$$

Το εύρος μεταβολής R της συνεχούς ποσοτικής μεταβλητής X του

πίνακα 4.3 είναι

$$R = 70 - 20$$
$$= 50.$$

[5.5] Ημιενδοτεταρτημοριακό Εύρος Ποσοτικής Μεταβλητής.

Το ημιενδοτεταρτημοριακό εύρος Q μιας ποσοτικής μεταβλητής ορίζεται ως το μισό της διαφοράς μεταξύ του τρίτου τεταρτημορίου Q_3 και του πρώτου τεταρτημορίου Q_1 της μεταβλητής.

Επομένως

$$Q = 1/2\,(Q_3 - Q_1) \quad (5.5.1)$$

Τα επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού του ημιενδοτεταρτημοριακού εύρους μιας ποσοτικής μεταβλητής.

[1] Ημιενδοτεταρτημοριακό Εύρος Διακριτής Ποσοτικής Μεταβλητής.

[2] Ημιενδοτεταρτημοριακό Εύρος Συνεχούς Ποσοτικής Μεταβλητής.

[5.6] Ημιενδοτεταρτημοριακό Εύρος Διακριτής Ποσοτικής Μεταβλητής.

Έστω X διακριτή ποσοτική μεταβλητή με πρώτο τεταρτημόριο Q_1 και τρίτο τεταρτημόρια Q_3. Ο τρόπος υπολογισμού του Q_1 παρουσιάζεται στο εδάφιο [4.6] και ο τρόπος υπολογισμού του Q_3 παρουσιάζεται στο εδάφιο [4.7]. Επομένως ο υπολογισμός του ημιενδοτεταρτημοριακού εύρους Q της X γίνεται με την βοήθεια της σχέσης (5.5.1) και των εδαφίων [4.6], [4.7].

Ένα παράδειγμα υπολογισμού του ημιενδοτεταρτημοριακού εύρους Q μιας διακριτής ποσοτικής μεταβλητής είναι το ακόλουθο.

Από το εδάφιο [4.6] το πρώτο τεταρτημόριο Q_1 της διακριτής ποσοτικής μεταβλητής του πίνακα 3.2 είναι

$Q_1 = 2$

Από το εδάφιο [4.7] το τρίτο τεταρτημόριο Q_3 της διακριτής ποσοτικής μεταβλητής του πίνακα 3.2 είναι

$Q_3 = 4$

Επομένως το ημιενδοτεταρτημοριακό εύρος Q της διακριτής ποσοτικής μεταβλητής του πίνακα 3.2 είναι

$$Q = 1/2\ (4-2) = 1.$$

[5.7] <u>Ημιενδοτεταρτημοριακό Εύρος Συνεχούς Ποσοτικής Μεταβλητής.</u>

Έστω X συνεχής ποσοτική μεταβλητή με πρώτο τεταρτημόριο Q_1 και τρίτο τεταρτημόρια Q_3. Ο τρόπος υπολογισμού του Q_1 παρουσιάζεται στο εδάφιο [4.8] και ο τρόπος υπολογισμού του Q_3 παρουσιάζεται στο εδάφιο [4.9]. Επομένως ο υπολογισμός του ημιενδοτεταρτημοριακού εύρους Q της X γίνεται με την βοήθεια της σχέσης (5.5.1) και των εδαφίων [4.8], [4.9].

Ένα παράδειγμα υπολογισμού του ημιενδοτεταρτημοριακού εύρους Q μιας συνεχούς ποσοτικής μεταβλητής είναι το ακόλουθο.

Από το εδάφιο [4.8] το πρώτο τεταρτημόριο Q_1 της συνεχούς ποσοτικής μεταβλητής του πίνακα 3.3 είναι

$Q_1 = 32{,}8$

Από το εδάφιο [4.9] το τρίτο τεταρτημόριο Q_3 της συνεχούς ποσοτικής μεταβλητής του πίνακα 3.3 είναι

$Q_3 = 50$

Επομένως το ημιενδοτεταρτημοριακό εύρος Q της συνεχούς ποσοτικής μεταβλητής του πίνακα 3.3 είναι

$$Q = 1/2\ (50 - 32{,}8) = 8{,}6.$$

[5.8] <u>Μέση Απόλυτη Απόκλιση Ποσοτικής Μεταβλητής.</u>

Ένα άλλο μέτρο διασποράς μιας ποσοτικής μεταβλητής είναι η μέση απόλυτη απόκλιση. Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού της μέσης απόλυτης απόκλισης μιας ποσοτικής μεταβλητής.

[1] Μέση Απόλυτη Απόκλιση Διακριτής Ποσοτικής Μεταβλητής.

[2] Μέση Απόλυτη Απόκλιση Συνεχούς Ποσοτικής Μεταβλητής.

[5.9] <u>Μέση Απόλυτη Απόκλιση Διακριτής Ποσοτικής Μεταβλητής.</u>

Μέση απόλυτη απόκλιση μιας διακριτής ποσοτικής μεταβλητής X λέγεται ο μέσος αριθμητικός των απόλυτων διαφορών του μέσου αριθμητικού μ από τις τιμές της μεταβλητής. Η μέση απόλυτη απόκλιση συμβολίζεται με το γράμμα d.

Επομένως

$$d = \frac{1}{N} \sum_{i=1}^{\kappa} f_i \, |x_i - \mu| \qquad (5.9.1)$$

όπου

[1] x_i, i=1,2,...,κ, οι τιμές της μεταβλητής Χ,

[2] f_i, i=1,2,...,κ, οι αντίστοιχες απόλυτες συχνότητες των τιμών της μεταβλητής Χ,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής Χ.

Ο υπολογισμός της μέσης απόλυτης απόκλισης d της κατανομής συχνότητας του πίνακα 4.1 γίνεται με τη βοήθεια του πίνακα 5.1. Ο μέσος αριθμητικός μ της κατανομής του πίνακα 4.1 είναι μ = 1,4.

Πίνακας 5.1

Τιμές της Χ	Απόλυτες Συχνότητες	Απόλυτες Διαφορές	Μερικά Γινόμενα
x_i	f_i	$\|x_i - \mu\|$	$f_i \|x_i - \mu\|$
0	60	1,4	84,0
1	20	0,4	8,0
2	12	0,6	7,2
3	9	1,6	14,4
4	5	2,6	13,0
5	6	3,6	21,6
6	8	4,6	36,8
	120		185,0

Από τον πίνακα 5.1 προκύπτει

κ = 7

$N = \sum_{i=1}^{\kappa} f_i,$

= 120

$\sum_{i=1}^{\kappa} f_i \, |x_i - \mu| = 185$

Επομένως η μέση απόλυτη απόκλιση d της κατανομής του πίνακα 4.1 είναι :

$$d = \frac{1}{N} \sum_{i=1}^{\kappa} f_i \, |x_i - \mu| = \frac{185}{120} = 1,5$$

[5.10] Μέση Απόλυτη Απόκλιση Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός της μέσης απόλυτης απόκλισης μιας συνεχούς ποσοτικής μεταβλητής Χ γίνεται με τη χρήση του τύπου.

$$d = \frac{1}{N} \sum_{i=1}^{\kappa} f_i \, |x_i - \mu| \qquad (5.10.1)$$

όπου

[1] x_i, i=1,2,...,κ, οι κεντρικές τιμές των τάξεων της μεταβλητής Χ,

[2] f_i, i=1,2,...,κ, οι αντίστοιχες απόλυτες συχνότητες των τιμών των τάξεων της μεταβλητής Χ,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής Χ.

Ο τύπος (5.10.1) κάνει χρήση των κεντρικών τιμών των τάξεων της συνεχούς ποσοτικής μεταβλητής Χ διότι είναι άγνωστες οι ακριβείς τιμές της μεταβλητής Χ.

Ο υπολογισμός της μέσης απόλυτης απόκλισης d της κατανομής συχνότητας του πίνακα 4.3. γίνεται με τη βοήθεια του πίνακα 5.2. Ο μέσος αριθμητικός μ της κατανομής του πίνακα 4.3 είναι μ = 41,8.

Πίνακας 5.2

Κεντρικές Τιμές των Τάξεων	Απόλυτες Συχνότητες	Απόλυτες Αποκλίσεις	Μερικά Γινόμενα
x_i	f_i	$\|x_i - \mu\|$	$f_i \|x_i - \mu\|$
25	30	16,8	504
35	40	6,8	272
45	35	3,2	112
55	25	13,2	330
65	15	23,2	348
	145		1566

Από τον πίνακα 5.2 προκύπτει

$\kappa = 5$

$N = \sum_{i=1}^{\kappa} f_i,$

$= 145$

$\sum_{i=1}^{\kappa} f_i \, |x_i - \mu| = 1566$

Επομένως η μέση απόλυτη απόκλιση d της κατανομής του πίνακα 4.3 είναι

$$d = \frac{1}{N} \sum_{i=1}^{\kappa} f_i \, |x_i - \mu|$$

$$= \frac{1566}{145}$$

$$= 10{,}8.$$

[5.11] Διακύμανση Ποσοτικής Μεταβλητής.

Το σπουδαιότερο και συνηθέστερο χρησιμοποιούμενο στη πράξη μέτρο διασποράς είναι η διακύμανση. Τα επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού της διακύμανσης μιας ποσοτικής μεταβλητής.

[1] Διακύμανση Διακριτής Ποσοτικής Μεταβλητής.

[2] Διακύμανση Συνεχούς Ποσοτικής Μεταβλητής.

[5.12] Διακύμανση Διακριτής Ποσοτικής Μεταβλητής.

Διακύμανση μιας διακριτής ποσοτικής μεταβλητής Χ λέγεται ο μέσος αριθμητικός των τετραγώνων των διαφορών του μέσου αριθμητικού μ από τις τιμές της μεταβλητής. Η διακύμανση συμβολίζεται με σ^2 ή V(X). Επομένως

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i \, (x_i - \mu)^2 \qquad (5.12.1)$$

όπου

[1] x_i, i=1,2,...,κ, οι τιμές της μεταβλητής Χ,

[2] f_i, i=1,2,...,κ, οι αντίστοιχες απόλυτες συχνότητες των τιμών της μεταβλητής Χ,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής Χ.

Ο υπολογισμός της διακύμανσης σ^2 της κατανομής συχνότητας του πίνακα 4.1 γίνεται με τη βοήθεια του πίνακα 5.3. Ο μέσος αριθμητικός μ της κατανομής του πίνακα 4.1 είναι $\mu = 1{,}4$.

Πίνακας 5.3

Τιμές της Χ·	Απόλυτες Συχνότητες	Τετράγωνα Διαφορών	Μερικά Γινόμενα
x_i	f_i	$(x_i - \mu)^2$	$f_i\,(x_i - \mu)^2$
0	60	1,96	117,60
1	20	0,16	3,20
2	12	0,36	4,32
3	9	2,56	23,04
4	5	6,76	33,80
5	6	12,96	77,76
6	8	21,16	169,28
	120		429,00

Από τον πίνακα 5.3 προκύπτει

$\kappa = 7$

$$N = \sum_{i=1}^{\kappa} f_i$$

$$= 120$$

$$\sum_{i=1}^{\kappa} f_i\,(x_i - \mu)^2 = 429$$

Επομένως η διακύμανση σ^2 της κατανομής του πίνακα 4.1 είναι

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i\,(x_i - \mu)^2$$

$$= \frac{429}{120}$$

$$= 3{,}57.$$

[5.13] Διακύμανση Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός της διακύμανσης μιας συνεχούς ποσοτικής μεταβλητής X γίνεται με τη χρήση του τύπου

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i\,(x_i - \mu)^2 \qquad (5.13.1)$$

όπου

[1] x_i, $i=1,2,\ldots,\kappa$, οι κεντρικές τιμές των τάξεων της μεταβλητής X,

[2] f_i, $i=1,2,\ldots,\kappa$, οι αντίστοιχες απόλυτες συχνότητες των τάξεων της μεταβλητής X,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής X.

Ο τύπος (5.13.1) κάνει χρήση των κεντρικών τιμών των τάξεων της συνεχούς ποσοτικής μεταβλητής X διότι είναι άγνωστες οι ακριβείς τιμές της μεταβλητής X.

Ο υπολογισμός της διακύμανσης σ^2 της κατανομής συχνότητας του πίνακα 4.3 γίνεται με τη βοήθεια του πίνακα 5.4. Ο μέσος αριθμητικός μ της κατανομής του πίνακα 4.3 είναι

$$\mu = 41{,}8.$$

Πίνακας 5.4

Κεντρικές Τιμές των Τάξεων	Απόλυτες Συχνότητες	Τετράγωνα Διαφορών	Μερικά Γινόμενα
x_i	f_i	$(x_i - \mu)^2$	$f_i (x_i - \mu)^2$
25	30	282,24	8467,2
35	40	46,24	1849,6
45	35	10,24	358,4
55	25	174,24	4356,0
65	15	538,24	8073,6
	145		23104,80

Από τον πίνακα 5.4 προκύπτει

$$\kappa = 5$$

$$N = \sum_{i=1}^{\kappa} f_i$$

$$= 145$$

$$\sum_{i=1}^{\kappa} f_i (x_i - \mu)^2 = 23104,8$$

Επομένως η διακύμανση σ^2 της κατανομής του πίνακα 4.3 είναι

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i (x_i - \mu)^2$$

$$= \frac{23104,8}{145}$$

$$= 159,34$$

[5.14] Ιδιότητες Διακύμανσης

Το εδάφιο αυτό διαπραγματεύεται μερικές ιδιότητες της διακύμανσης μιας ποσοτικής - διακριτής ή συνεχούς - μεταβλητής. Οι ιδιότητες αυτές παρουσιάζουν θεωρητικό και πρακτικό ενδιαφέρον.

Ιδιότητα 1.

Έστω Χ ποσοτική μεταβλητή με μέσο αριθμητικό μ. Η διακύμανση σ^2 της Χ γράφεται

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i x_i^2 - \mu^2 \qquad (5.14.1)$$

Απόδειξη.

Υποθέτουμε ότι η μεταβλητή Χ είναι διακριτή με τιμές x_i, i=1,2,...,κ, και αντίστοιχες συχνότητες f_i, i=1,2,...,κ. Έχουμε

$$\sum_{i=1}^{\kappa} f_i (x_i - \mu)^2 = \sum_{i=1}^{\kappa} f_i (x_i^2 - 2\mu x_i + \mu^2) =$$

$$= \sum_{i=1}^{k} f_i x_i^2 - 2\mu \sum_{i=1}^{k} f_i x_i + \mu^2 \sum_{i=1}^{k} f_i \qquad (5.14.2)$$

Επειδή

$$N = \sum_{i=1}^{\kappa} f_i$$

και

$$\mu N = \sum_{i=1}^{\kappa} f_i x_i$$

τότε η (5.14.2) γράφεται

$$\sum_{i=1}^{\kappa} f_i (x_i - \mu)^2 = \sum_{i=1}^{\kappa} f_i x_i^2 - 2N\mu^2 + N\mu^2 =$$

$$= \sum_{i=1}^{\kappa} f_i\, x_i^2 - N\mu^2 \qquad (5.14.2)$$

Από την (5.14.2) έχουμε

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i\, (x_i - \mu)^2$$

$$= \frac{1}{N} \sum_{i=1}^{\kappa} f_i x_i^2 - \mu^2$$

Αν η μεταβλητή είναι συνεχής τότε οι x_i, i=1,2,...,κ είναι οι κεντρικές τιμές των τάξεων της μεταβλητής και η απόδειξη της (5.14.1) γίνεται με τον ίδιο τρόπο.

Ιδιότητα 2.

Αν η ποσοτική μεταβλητή X είναι μια σταθερά δηλαδή όλες οι τιμές της μεταβλητής είναι ίσες με α τότε η διακύμανση σ^2 είναι :

$$\sigma^2 = 0 \qquad (5.14.3)$$

Απόδειξη.

Επειδή

$$x_i = \alpha,\ i=1,2,...,\kappa \qquad (5.14.4)$$

τότε από τη σχέση (4.19.1) έχουμε ότι ο μέσος αριθμητικός μ της X είναι

$$\mu = \alpha \qquad (5.14.5)$$

Από την (5.14..4) και τη (5.14.5) έχουμε ότι η διακύμανση σ^2 της X είναι :

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i\, (x_i - \mu)^2$$

$$= \frac{1}{N} \sum_{i=1}^{\kappa} f_i\, (\alpha - \alpha)^2$$

$$= 0.$$

Ιδιότητα 3.

Έστω X ποσοτική μεταβλητή με διακύμανση V(X).Θεωρούμε την ποσοτική μεταβλητή

$$Y = \alpha + \beta X \qquad (5.14.6)$$

όπου α, β πραγματικοί αριθμοί. Η διακύμανση V(Y) της Y είναι

$$V(Y) = \beta^2 V(X) \qquad (5.14.7)$$

Απόδειξη.

Υποθέτουμε ότι η μεταβλητή X είναι διακριτή με τιμές x_i, i=1,2,...,κ και αντίστοιχες συχνότητες f_i, i=1,2,...,κ. Οι τιμές της Y είναι :

$$y_i = \alpha + \beta_{Xi} \qquad (5.14.8)$$

με αντίστοιχες συχνότητες f_i, i=1,2,...,κ. Αν μ_X είναι ο αριθμητικός μέσος τότε από την (4.19.10) προκύπτει ότι ο αριθμητικός μέσος μ_y της Y είναι :

$$\mu_y = \alpha + \beta\mu_X \qquad (5.14.9)$$

Από την (5.14.8) και (5.14.9) προκύπτει ότι η διακύμανση V(X) της Y είναι :

$$V(Y) = \frac{1}{N} \sum_{i=1}^{\kappa} f_i\, (y_i - \mu_y)^2$$

$$= \frac{1}{N} \sum_{i=1}^{\kappa} f_i\, [(\alpha + \beta x_i) - (\alpha + \beta\mu_X)]^2$$

$$= \beta^2 \frac{1}{N} \sum_{i=1}^{\kappa} f_i\, (x_i - \mu_x)^2$$

$$= \beta^2\, V(X).$$

Ιδιότητα 4.

Έστω X ποσοτική μεταβλητή με αριθμητικό μέσο μ και λ πραγματικός αριθμός. Η ποσότητα

$$f(\lambda) = \frac{1}{N} \sum_{i=1}^{\kappa} f_i \, (y_i - \lambda)^2 \qquad (5.14.10)$$

ελαχιστοποιείται όταν

$$\lambda = \mu$$

Απόδειξη.

Υποθέτουμε ότι η X είναι διακριτή ποσοτική μεταβλητή με τιμές x_i, i = 1,2,...,k και αντίστοιχες συχνότητες f_i, i = 1,2,...,k. Οι ικανές συνθήκες για την τιμή του λ που ελχιστοποιεί την (5.14.10) είναι :

$$\frac{df\ (\lambda)}{d\lambda} = 0 \qquad (5.14.11)$$

και

$$\frac{d^2f\ (\lambda)}{d\lambda^2} > 0$$

Επειδή

$$\frac{df\ (\lambda)}{d\lambda} = - \frac{2}{N} \sum_{i=1}^{k} f_i(x_i - \lambda) \qquad (5.14.12)$$

και

$$\frac{d^2f\ (\lambda)}{d\lambda^2} = 2 > 0$$

τότε από την (5.14.11) και την (5.14.12) έχουμε :

$$- \frac{2}{N} \sum_{i=1}^{k} f_i(x_i - \lambda) = 0$$

ισοδύναμα

$$\lambda = \frac{1}{N} \sum_{i=1}^{k} f_i x_i \qquad (5.14.13)$$

Επειδή

$$N = \sum_{i=1}^{k} f_i$$

τότε από την (5.14.13) έχουμε :

$$\lambda = \mu$$

Αν η μεταβλητή είναι συνεχής τότε οι x_i, i=1,2,...,k είναι οι κεντρικές τιμές των τάξεων της μεταβλητής και η απόδειξη της (5.14.10) γίνεται με τον ίδιο τρόπο.

[5.15] Τυπική Απόκλιση Ποσοτικής Μεταβλητής.

Η τυπική απόκλιση είναι ένα βασικό και συχνά χρησιμοποιούμενο στη πράξη μέτρο διασποράς. Η τυπική απόκλιση μιας ποσοτικής μεταβλητής X ορίζεται ως η θετική τετραγωνική ρίζα της διακύμανσης της μεταβλητής και συμβολίζεται με σ ή $\sqrt{V(X)}$. Διακρίνουμε τις ακόλουθες περιπτώσεις υπολογισμού της τυπικής απόκλισης μιας ποσοτικής μεταβλητής.

[1] Τυπική Απόκλιση Διακριτής Ποσοτικής Μεταβλητής.

[2] Τυπική Απόκλιση Συνεχούς Ποσοτικής Μεταβλητής.

[5.16] Τυπική Απόκλιση Διακριτής Ποσοτικής Μεταβλητής.

Έστω X διακριτή ποσοτική μεταβλητή. Ο υπολογισμός της διακύμανσης σ^2 ή V(X) της X γίνεται με χρήση του τύπου (5.12.1). Η τυπική απόκλιση σ ή $\sqrt{V(X)}$ της X είναι η θετική τετραγωνική ρίζα της (5.12.1).

Η διακύμανση της διακριτής ποσοτικής μεταβλητής του πίνακα 4.1 είναι :

$$\sigma^2 = 3{,}57$$

Επομένως η τυπική απόκλιση της διακριτής ποσοτικής μεταβλητής του πίνακα 4.1 είναι :

$$\sigma = \sqrt{3{,}57}$$

$$= 1{,}88$$

[5.17] Τυπική Απόκλιση Συνεχούς Ποσοτικής Μεταβλητής.

Έστω X συνεχής ποσοτική μεταβλητή. Ο υπολογισμός της διακύμανσης σ^2 ή V(X) της X γίνεται με χρήση του τύπου (5.13.1). Η τυπική απόκλιση σ ή $\sqrt{V(X)}$ της X είναι η θετική τετραγωνική ρίζα της (5.13.1).

Η διακύμανση της συνεχούς ποσοτικής μεταβλητής του πίνακα 4.3 είναι :

$$\sigma^2 = 159{,}34$$

Επομένως η τυπική απόκλιση της συνεχούς ποσοτικής μεταβλητής

του πίνακα 4.3 είναι :

$$\sigma = \sqrt{159{,}34}$$

$$= 12{,}62.$$

[5.18] Συντελεστής Μεταβλητικότητας Ποσοτικής Μεταβλητής.

Ο συντελεστής μεταβλητικότητας μιας ποσοτικής μεταβλητής ορίζεται ως το πηλίκο της τυπικής απόκλισης της μεταβλητής δια του αριθμητικού μέσου της μεταβλητής. Διακρίνουμε τις ακόλουθες περιπτώσεις του υπολογισμού του συντελεστή μεταβλητικότητας μιας ποσοτικής μεταβλητής.

[1] Συντελεστής Μεταβλητικότητας Διακριτής Ποσοτικής Μεταβλητής.

[2] Συντελεστής Μεταβλητικότητας Συνεχούς Ποσοτικής Μεταβλητής.

[5.19] Συντελεστής Μεταβλητικότητας Διακριτής Ποσοτικής Μεταβλητής.

Έστω X διακριτή ποσοτική μεταβλητή με μέσο αριθμητικό μ και τυπική απόκλιση σ. Ο αριθμητικός μέσος μ της X υπολογίζεται με χρήση του τύπου (4.17.1). Η τυπική απόκλιση σ της X είναι η τετραγωνική ρίζα της (5.12.1). Ο συντελεστής μεταβλητικότητας CV(X) της X είναι :

$$CV(X) = \frac{\sigma}{\mu}$$

$$= \frac{\sqrt{V(X)}}{\mu} \qquad (5.19.1)$$

Ο αριθμητικός μέσος μ και η τυπική απόκλιση σ της διακριτής ποσοτικής μεταβλητής του πίνακα 4.1 είναι :

$\mu = 1,4$

$\sigma = 1,88$

Επομένως ο συντελεστής μεταβλητικότητας CV(X) της διακριτής ποσοτικής μεταβλητής του πίνακα 4.1 είναι

$$CV(X) = \frac{1,88}{1,4}$$

$$= 1,34$$

[5.20] Συντελεστής Μεταβλητικότητας Συνεχούς Ποσοτικής Μεταβλητής.

Έστω X συνεχής ποσοτική μεταβλητή με μέσο αριθμητικό μ και τυπική απόκλιση σ. Ο αριθμητικός μέσος μ της X υπολογίζεται με χρήση του τύπου (4.18.1). Η τυπική απόκλιση σ της X είναι η τετραγωνική ρίζα της (5.13.1). Ο συντελεστής μεταβλητικότητας CV(X) της X είναι :

$$CV(X) = \frac{\sigma}{\mu} = \frac{\sqrt{V(X)}}{\mu} \qquad (5.20.1)$$

Ο αριθμητικός μέσος μ και η τυπική απόκλιση σ της συνεχούς ποσοτικής μεταβλητής του πίνακα 4.3 είναι :

$\mu = 41,8$

$\sigma = 12,62$

Επομένως ο συντελεστής μεταβλητικότητας CV(X) της συνεχούς ποσοτικής μεταβλητής του πίνακα 4.3 είναι

$$CV(X) = \frac{12,62}{41,8}$$

$$= 0,30$$

[5.21] Χαρακτηριστικά των Μέτρων Διασποράς.

Βασικά χαρακτηριστικά των μέτρων διασποράς είναι τα ακόλουθα :

[1] Το εύρος μεταβολής επηρεάζεται μόνο από τις ακραίες τιμές της μεταβλητής.

[2] Το ημιενδοτεταρτημοριακό εύρος επηρεάζεται από το πρώτο και το τρίτο τεταρτημόριο.

[3] Η διακύμανση, η τυπική απόκλιση και η μέση απόλυτη απόκλιση επηρεάζονται από όλες τις τιμές της μεταβλητής.

[4] Ένα σημαντικό μειονέκτημα του συντελεστή μεταβλητικότητας είναι ότι επηρεάζεται έντονα από το μέσο αριθμητικό.

ΚΕΦΑΛΑΙΟ 6

ΡΟΠΕΣ ΜΟΝΟΜΕΤΑΒΛΗΤΩΝ ΚΑΤΑΝΟΜΩΝ

[6.1] Κατηγορίες Ροπών.

Τα σημαντικότερα μέτρα θέσης και διασποράς καθώς και τα βασικά μέτρα ασυμμετρίας και κυρτότητας, τα οποία παρουσιάζονται στο επόμενο κεφάλαιο, είναι δυνατόν να εκφραστούν με τη βοήθεια ορισμένων χαρακτηριτικών μεγεθών γνωστών, ως ροπών. Το κεφάλαιο αυτό διαπραγματεύεται τις ακόλουθες κατηγορίες ροπών.

[1] Ροπές περί την Αρχή.

[2] Ροπές περί την Αριθμητικό Μέσο.

[6.2] Ροπές περί την Αρχή Ποσοτικής Μεταβλητής.

Τα επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού των ροπών περί την αρχή μιας ποσοτικής μεταβλητής.

[1] Ροπές περί την Αρχή Διακριτής Ποσοτικής Μεταβλητής.

[2] Ροπές περί την Αρχή Συνεχούς Ποσοτικής Μεταβλητής.

[6.3] Ροπές περί την Αρχή Διακριτής Ποσοτικής Μεταβλητής.

Η ροπή τ-τάξης περί την αρχή μιας διακριτής ποσοτικής μεταβλητής Χ λέγεται ο αριθμητικός μέσος των τ-δυνάμεων, $\tau = 0,1,2,\ldots$, των τιμών της μεταβλητής. Η ροπή τ-τάξης συμβολίζεται συνήθως με το γράμμα v_τ. Επομένως

$$v_\tau = \frac{1}{N}\sum_{i=1}^{k} f_i\, x_i^{\tau} \qquad (6.3.1)$$

όπου

[1] x_i, $i=1,2,3,\ldots,\kappa$, οι τιμές της μεταβλητής Χ,

[2] f_i, $i=1,2,3,\ldots,\kappa$, οι αντίστοιχες απόλυτες συχνότητες των τιμών της μεταβλητής Χ,

[3] $N = \sum_{i=1}^{k} f_i$, το πλήθος των τιμών της μεταβλητής Χ.

Από τη σχέση (6.3.1) για $\tau = 0$ προκύπτει

$$v_0 = 1 \qquad (6.3.2)$$

δηλαδή η μηδενική ροπή περί την αρχή είναι πάντοτε ίση προς τη μονάδα.

Από τη σχέση (6.3.1) για $\tau = 1$ προκύπτει

$$v_1 = \mu \qquad (6.3.3)$$

δηλαδή η πρώτη ροπή περί την αρχή είναι πάντοτε ίση προς τον αριθμητικό μέσο μ.

Ο υπολογισμός της τρίτης ροπής περί την αρχή v_3 της κατανομής του πίνακα 4.1 γίνεται με την βοήθεια του πίνακα 6.1.

Πίνακας 6.1

Τιμές της X	Απόλυτες Συχνότητες	Μερικά Γινόμενα
xi	fi	fi xi^3
0	60	0
1	20	20
2	12	96
3	9	243
4	5	320
5	6	750
6	8	1728
	120	3157

Από τον πίνακα 6.1 προκύπτει ότι :

$\kappa = 7$

$$N = \sum_{i=1}^{\kappa} f_i,$$

$$= 120$$

$$\sum_{i=1}^{\kappa} f_i x_i^{\,3} = 3157$$

Επομένως η τρίτη ροπή περί την αρχή v_3 της κατανομής του πίνακα 4.1 είναι

$$v_3 = \frac{1}{N}\sum_{i=1}^{k} f_i\, x_i^3$$

$$= \frac{3157}{120}$$

$$= 26{,}30.$$

[6.4] Ροπές περί την Αρχή Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός της ροπή τ-τάξης περί την αρχή μιας συνεχούς ποσοτικής μεταβλητής X γίνεται με τη χρήση του τύπου :

$$v_\tau = \frac{1}{N}\sum_{i=1}^{k} f_i\, x_i^\tau \qquad (6.4.1)$$

όπου

[1] x_i, $i=1,2,3,\ldots,\kappa$, οι κεντρικές τιμές των τάξεων της μεταβλητής X,

[2] f_i, $i=1,2,3,\ldots,\kappa$, οι αντίτσοιχες απόλυτες συχνότητες των τάξεων της μεταβλητής X,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής X.

Ο υπολογισμός της τρίτης ροπής περί την αρχή v_3 της κατανομής του πίνακα 4.3 γίνεται με την βοήθεια του πίνακα 6.2.

Πίνακας 6.2

Τάξεις της X	Κεντρικές Τιμές των Τάξεων	Απόλυτες Συχνότητες	Μερικά Γινόμενα
	xi	fi	fi xi^3
20 - 30	25	30	468750
30 - 40	35	40	1715000
40 - 50	45	35	3189375
50 - 60	55	25	4159375
60 - 70	65	15	4119375
		145	13651875

Από τον πίνακα 6.2 προκύπτει ότι :

$\kappa = 5$

$$N = \sum_{i=1}^{\kappa} f_i,$$
$$= 145$$

$$\sum_{i=1}^{\kappa} f_i x_i^{\,3} = 13651875$$

Επομένως η τρίτη ροπή περί την αρχή v_3 της κατανομής του πίνακα 4.3 είναι

$$v_3 = \frac{1}{N}\sum_{i=1}^{k} f_i\, x_i^3$$
$$= \frac{13651875}{145}$$
$$= 94150{,}86.$$

[6.5] Ροπές περί την Αριθμητικό Μέσο Ποσοτικής Μεταβλητής.

Τα επόμενα εδάφη διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού των ροπών περί τον μέσο αριθμητικό μιας ποσοτικής μεταβλητής.

[1] Ροπές περί τον Αριθμητικό Μέσο Διακριτής Ποσοτικής Μεταβλητής.

[2] Ροπές περί τον Αριθμητικό Μέσο Συνεχούς Ποσοτικής Μεταβλητής.

[6.6] Ροπές περί τον Αριθμητικό Μέσο Διακριτής Ποσοτικής Μεταβλητής.

Ο υπολογισμός της ροπής τ-τάξης περί τον μέσο μ μιας διακριτής ποσοτικής μεταβλητής γίνεται με τη χρήση του τύπου :

$$\mu_\tau = \frac{1}{N}\sum_{i=1}^{k} f_i\,(x_i - \mu)^\tau \qquad (6.6.1)$$

όπου

[1] x_i, i=1,2,...,κ, οι τιμές της μεταβλητής X,

[2] f_i, i=1,2,...,κ, οι αντίστοιχες απόλυτες συχνότητες των τιμών της μεταβλητής X,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής X.

Από τη σχέση (6.6.1) για τ = 0 προκύπτει

$$\mu_0 = 1 \qquad (6.6.2)$$

δηλαδή η μηδενική τάξης ροπή περί τον αριθμητικό μέσο είναι πάντοτε ίση προς τη μονάδα.

Από τη σχέση (6.6.1) για τ = 1 και τη σχέση (4.19.7) προκύπτει

$$\mu_1 = 0 \qquad (6.6.3)$$

δηλαδή η πρώτη τάξης ροπή περί τον αριθμητικό μέσο είναι ίση προς το μηδέν. Επίσης από την σχέση (6.6.1) για τ=2 και τη σχέση (5.12.1) προκύπτει $\mu_2 = \sigma^2$.

Ο υπολογισμός της τρίτης τάξης ροπής περί τον αριθμητικό μέσο μ_3 της κατανομής του πίνακα 4.1 γίνεται με την βοήθεια του πίνακα 6.3.

Πίνακας 6.3

Τιμές της X	Απόλυτες Συχνότητες	Μερικά Γινόμενα
xi	fi	$f_i (x_i - \mu)^3$
0	60	- 164,640
1	20	- 1,280
2	12	2,592
3	9	36,864
4	5	87,880
5	6	279,936
6	8	778,688
	120	1020,040

Από τον πίνακα 6.3 προκύπτει ότι :

$\kappa = 7$

$$N = \sum_{i=1}^{\kappa} f_i,$$

$$= 120$$

$$\sum_{i=1}^{\kappa} f_i(x_i - \mu)^3 = 1020{,}04$$

Επομένως η τρίτη τάξης ροπή περί τον αριθμητικό μέσο μ_3 της κατανομής του πίνακα 4.1 είναι

$$v_3 = \frac{1}{N} \sum_{i=1}^{k} f_i (x_i - \mu)^3$$

$$= \frac{1020{,}04}{120}$$

$$= 8{,}50.$$

[6.7] Ροπές περί τον Αριθμητικό Μέσο Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός της ροπής τ-τάξης περί τον μέσο μ_3 μιας συνεχούς ποσοτικής μεταβλητής X γίνεται με την χρήση του τύπου :

$$\mu_\tau = \frac{1}{N} \sum_{i=1}^{k} f_i (x_i - \mu)^\tau \qquad (6.7.1)$$

όπου

[1] x_i, i=1,2,...,κ, οι κεντρικές τιμές των τάξεων της μεταβλητής X,

[2] f_i, i=1,2,...,κ, οι αντίσοιχες απόλυτες συχνότητες των τάξεων της μεταβλητής X,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής X.

Ο υπολογισμός της τρίτης τάξης ροπής περί τον αριθμητικό μέσο μ_3 της κατανομής του πίνακα 4.3 γίνεται με την βοήθεια του πίνακα 6.4.

Πίνακας 6.4

Τάξεις της X	Κεντρικές Τιμές των Τάξεων	Απόλυτες Συχνότητες	Μερικά Γινόμενα
	xi	fi	fi (xi - μ)^3
20 - 30	25	30	- 142248,96
30 - 40	35	40	- 12577,28
40 - 50	45	35	1146,88
50 - 60	55	25	57499,20
60 - 70	65	15	187307,52
		145	91127,36

Από τον πίνακα 6.4 προκύπτει ότι :

κ = 5

$$N = \sum_{i=1}^{\kappa} f_i,$$

$$= 145$$

$$\sum_{i=1}^{5} f_i (x_i - \mu)^3 = 91127{,}36$$

Επομένως η τρίτη τάξης ροπή περί τον μέσο μ_3 της κατανομής του πίνακα 4.3 είναι

$$\mu_3 = \frac{1}{N} \sum_{i=1}^{k} f_i (x_i - \mu)^3$$

$$= \frac{91127{,}36}{145}$$

$$= 628{,}46.$$

[6.8] Σχέση Ροπών περί τον Αριθμητικό Μέσο και Ροπών περί την Αρχή.

Σκοπός του εδάφιου αυτού είναι να θεμελιώσει τη σχέση μεταξύ ροπών περί τον αριθμητικό μέσο και ροπών περί την αρχή.

Έχουμε

$$\mu_\tau = \frac{1}{N} \sum_{i=1}^{k} f_i (x_i - \mu)^\tau$$

Επειδή

$$(x_i - \mu)^\tau = \sum_{i=1}^{\kappa} (-1)^\lambda \binom{\tau}{\lambda} x_i^{\tau-\lambda} \mu^\lambda$$

όπου

$$\binom{\tau}{\lambda} = \frac{\tau\,!.}{\lambda!\ (\tau - \lambda)!}$$

είναι το πλήθος των συνδυασμών των τ στοιχείων ανά λ, τότε η ροπή μ_τ γράφεται

$$\mu_\tau = \frac{1}{N} \sum_{i=1}^{\kappa} f_i \sum_{\lambda=0}^{\tau} (-1)^\lambda \binom{\tau}{\lambda} x_i^{\tau-\lambda} \mu^\lambda$$

ή ισοδύναμα

$$\mu_\tau = \sum_{\lambda=0}^{\tau} (-1)^\lambda \binom{\tau}{\lambda} \mu^\lambda \frac{1}{N} \sum_{i=1}^{\kappa} f_i\, x_i^{\tau-\lambda} \qquad (6.8.1)$$

Επειδή

$$v_{\tau-\lambda} = \frac{1}{N} \sum_{i=1}^{k} f_i x_i^{\tau-\lambda}$$

είναι η (τ-λ)-τάξης ροπή περί την αρχή η (6.8.1) γράφεται

$$\mu_\tau = \sum_{\lambda=0}^{\tau} (-1)^\lambda \binom{\tau}{\lambda} \nu_{\tau-\lambda}\, \mu^\lambda \qquad (6.8.2)$$

Επομένως η (6.8.2) είναι η σχέση μεταξύ ροπών περί τον αριθμητικό μέσο και ροπών περί την αρχή.

[6.9] Σφάλματα Ομαδοποίησης.

Η συνοπτική παρουσίαση των τιμών των ποσοτικών στατιστικών μεταβλητών - ιδιαίτερα των συνεχών ποσοτικών μεταβλητών - υπό μορφή κατανομής συχνότητας βασίζεται στην υπόθεση ότι οι τιμές κάθε τάξης συμπίπτουν με την κεντρική τιμή της τάξης. Η υπόθεση αυτή είναι συμβατική και δεν ανταποκρίνεται στην πραγματικότητα. Επομένως οι ροπές οι οποίες υπολογίζονται με βάση την υπόθεση αυτή διαφέρουν από τις πραγματικές ροπές οι οποίες υπολογίζονται από τις απλές τιμές μιας ποσοτικής στατιστικής μεταβλητής Χ, οι διαφορές αυτές λέγονται σφάλματα ομαδοποίησης.

Ένας τρόπος περιορισμού των σφαλμάτων ομαδοποίησης είναι η χρήση των διορθώσεων κατά Sheppard. Οι διορθώσεις αυτές εφαρμόζονται κυρίως στην διακύμανση σ^2 και στην τέταρτης τάξης ροπή περί τον αριθμητικό μέσο μ_4 ταξινομημένων κατανομών συχνότητας.

Έχουμε

$$\sigma_*^2 = \sigma^2 - \frac{\delta^2}{12} \qquad (6.9.1)$$

και

$$\mu_4^* = \mu_4 - \frac{\sigma^2\delta^2}{2} + \frac{7\delta^4}{240} \qquad (6.9.2)$$

όπου

[1] σ_*^2, η διορθωμένη διακύμανση της μεταβλητής Χ,

[2] σ^2, η διακύμανση της μεταβλητής Χ,

[3] δ, το πλάτος των τάξεων της μεταβλητής Χ,

[4] μ_4^*, η διορθωμένη τέταρτης τάξης ροπή περί τον αριθμητικό μέσο της μεταβλητής Χ,

[5] μ_4, η τέταρτη τάξης ροπή περί τον αριθμητικό μέσο της μεταβλητής Χ.

ΚΕΦΑΛΑΙΟ 7

ΜΕΤΡΑ ΑΣΥΜΜΕΤΡΙΑΣ ΚΑΙ ΜΕΤΡΑ ΚΥΡΤΟΤΗΤΑΣ ΜΟΝΟΜΕΤΑΒΛΗΤΩΝ ΚΑΤΑΝΟΜΩΝ

[7.1] Κατηγορίες Μέτρων Ασυμμετρίας.

Τα μέτρα ασυμμετρίας μιας μονομεταβλητής κατανομής συχνότητας χαρακτηρίζουν την ασυμμετρία των τιμών της αντίστοιχης ποσοτικής στατιστικής μεταβλητής, δηλαδή το βαθμό απόκλισης της μονομεταβλητής κατανομής συχνότητας από μία πρότυπο συμμετρική μονομεταβλητή κατανομή συχνότητας. Το κεφάλαιο αυτό διαπραγματεύεται τα ακόλουθα μέτρα ασυμμετρίας μονομεταβλητών κατανομών συχνότητας.

[1] Συντελεστής Ασυμμετρίας του Pearson.

[2] Συντελεστής Ασυμμετρίας με Βάση την Διάμεσο και τα Τεταρτημόρια.

[3] Συντελεστής Ασυμμετρίας με Βάση τον Αριθμητικό Μέσο και την Διάμεσο.

[7.2] Συντελεστής Ασυμετρίας του Pearson Ποσοτικής Μεταβλητής.

Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού του συντελεστή ασυμμετρίας του Pearson μιας ποσοτικής μεταβλητής.

[1] Συντελεστής Ασυμμετρίας του Pearson Διακριτής Ποσοτικής Μεταβλητής.

[2] Συντελεστής Ασυμμετρίας του Pearson Συνεχούς Ποσοτικής Μεταβλητής.

[7.3] Συντελεστής Ασυμμετρίας του Pearson Διακριτής Ποσοτικής Μεταβλητής.

Ο υπολογισμός του συντελεστή ασυμμετρίας του Pearson β_1 μιας διακριτής ποσοτικής μεταβλητής X γίνεται με τη χρήση του τύπου

$$\beta_1 = \frac{\mu_3}{\sigma^3} \qquad (7.3.1)$$

όπου
[1] μ_3, της τρίτης τάξης ροπή περί τον αριθμητικό μέσο της μεταβλητής X,

[2] σ, η τυπική απόκλιση της μεταβλητής X.

Ο υπολογισμός της μ_3 και της σ γίνεται με τη χρήση των τύπων.

$$\mu_3 = \frac{1}{N}\sum_{i=1}^{k} f_i(x_i - \mu)^3 \qquad (7.3.2)$$

και

$$\sigma = \left[\frac{1}{N}\sum_{i=1}^{k} f_i(x_i - \mu)^2\right]^{1/2} \qquad (7.3.3)$$

όπου
[1] x_i, i = 1,2,...,κ, οι τιμές της μεταβλητής X,
[2] f_i, i = 1,2,...,k, οι αντίστοιχες απόλυτες συχνότητες των τιμών της μεταβλητής X,
[3] $N = \sum_{i=1}^{k} f_i$, το πλήθος των τιμών της μεταβλητής X,
[4] μ, ο μέσος αριθμητικός της μεταβλητής X.

Αν η κατανομή συχνότητας της διακριτής ποσοτικής μεταβλητής X είναι συμμετρική περί τον μέσο αριθμητικό μ δηλαδή, αν οι ισαπέχουσες από τον αριθμητικό μέσο μ τιμές της X παρουσιάζουν την ίδια συχνότητα, τότε η τρίτης τάξης ροπή περί τον μέσο αριθμητικό μ_3 και ο συντελεστής ασυμμετρίας του Pearson β_1 είναι μηδέν. Αν η κατανομή συχνότητας της διακριτής ποσοτικής μεταβλητής X παρουσιάζει ουρά προς τα δεξιά του αριθμητικού μέσου μ - δεξιά η ισοδύναμα θετική ασυμμετρία - τότε η τρίτη τάξης ροπή περί τον μέσο αριθμητικό μ_3 και ο συντελεστής ασυμμετρίας του Pearson β_1 έχουν θετική τιμή. Αντίθετα, αν η κατανομή συχνότητας της X παρουσιάζει ουρά προς τα αριστερά του μ - αριστερή ή ισοδύναμα αρνητική ασυμμετρία - τότε η τρίτης τάξης ροπή περί τον αριθμητικό μέσο μ_3 και ο συντελεστής ασυμμετρίας του Pearson β_1 έχουν αρνητική τιμή.

Επομένως και η τρίτης τάξης ροπή περί τον μέσον αριθμητικό μ_3 μπορεί να χρησιμοποιηθεί ως μέτρο ασυμμετρίας μιας κατανομής συχνότητας. Ένα βασικό πλεονέκτημα της ποσότητας β_1 ως μέτρου ασυμμετρίας είναι ότι η ποσότητα β_1 είναι καθαρός αριθμός.

Ο υπολογισμός της τυπικής απόκλισης σ και της τρίτης τάξης ροπής περί τον αριθμητικό μέσο μ_3 της κατανομής του πίνακα 4.1 έγιναν στα εδάφια [5.16] και [6.6] αντίστοιχα. Έχουμε

$$\sigma = 1,88$$

και

$$\mu_3 = 8,50$$

ο συντελεστής ασυμμετρίας του Pearson β_1 της κατανομής του πίνακα 4.1 είναι :

$$\beta_1 = \frac{\mu_3}{\sigma^3}$$

$$= \frac{8,50}{6,64}$$

$$= 1,28$$

Επομένως η κατανομμή του πίνακα 4.1 παρουσιάζει θετική ασυμμετρία περί την διάμεσο.

[7.4] Συντελεστής Ασυμμετρίας του Pearson Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του συντελεστή ασυμμετρίας του Pearson β_1 μιας συνεχούς ποσοτικής μεταβλητής X γίνεται με τη χρήση του τύπου

$$\beta_1 = \frac{\mu_3}{\sigma^3} \qquad (7.4.1)$$

όπου

[1] μ_3, η τρίτης τάξης ροπή περί τον μέσο αριθμητικό της μεταβλητής X,

[2] σ, η τυπική απόκλιση της μεταβλητής X.

Ο υπολογισμός της μ_3 και της σ γίνεται με τη χρήση των τύπων.

$$\mu_3 = \frac{1}{N}\sum_{i=1}^{k} f_i(x_i - \mu)^3 \qquad (7.4.2)$$

και

$$\sigma = [\frac{1}{N}\sum_{i=1}^{k} f_i(x_i - \mu)^2]^{1/2} \qquad (7.4.3)$$

όπου

[1] x_i, i = 1,2,...,κ, οι κεντρικές τιμές των τάξεων της μεταβλητής X,

[2] f_i, i = 1,2,...,k, οι αντίστοιχες απόλυτες συχνότητες των τάξεων της μεταβλητής X,

[3] $N = \sum_{i=1}^{k} f_i$, το πλήθος των τιμών της μεταβλήτής X,

[4] μ, ο αριθμητικός μέσος της μεταβλητής X.

Ο υπολογισμός της τυπικής απόκλισης σ και της τρίτης τάξης ροπής περί τον αριθμητικό μέσο μ_3 της κατανομής του πίνακα 4.3 έγιναν στα εδάφια [5.17] και [6.7] αντίστοιχα. Έχουμε

$$\sigma = 12,62$$

και

$$\mu_3 = 628,46$$

ο συντελεστής ασυμμετρίας του Pearson β_1 της κατανομής του πίνακα 4.3 είναι :

$$\beta_1 = \frac{\mu_3}{\sigma^3}$$

$$= \frac{628,46}{2009,91}$$

$$= 0,31$$

Επομένως η κατανομή του πίνακα 4.3 παρουσιάζει θετική ασυμμετρία περί την διάμεσο.

[7.5] Συντελεστής Ασυμμετρίας με Βάση την Διάμεσο και τα Τεταρτημόρια

Ποσοτικής Μεταβλητής.

Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού του συντελεστή ασυμμετρίας με βάση την διάμεσο και τα τεταρτημόρια μιας ποσοτικής μεταβλητής.

[1] Συντελεστής Ασυμμετρίας με Βάση την Διάμεσο και τα Τεταρτημόρια Διακριτής Ποσοτικής Μεταβλητής.

[2] Συντελεστής Ασυμμετρίας με Βάση την Διάμεσο και τα Τεταρτημόρια Συνεχούς Ποσοτικής Μεταβλητής.

[7.6] Συντελεστής Ασυμμετρίας με Βάση την Διάμεσο και τα Τεταρτημόρια Διακριτής Ποσοτικής Μεταβλητής.

Ο υπολογισμός του συντελεστή ασυμμετρίας με βάση την διάμεσο και τα τεταρτημόρια S_1 μιάς διακριτής ποσοτικής μεταβλητής X, γίνεται με τη χρήση του τύπου :

$$S_1 = \frac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)}$$

όπου :

[1] Q_1, το πρώτο τεταρτημόριο της μεταβλητής X,

[2] Q_3, το τρίτο τεταρτημόριο της μεταβλητής X,

[3] M, η διάμεσος της μεταβλητής X.

Ο υπολογισμός της διαμέσου M και των τεταρτημορίων Q_1 και Q_3 διακριτής ποσοτικής μεταβλητής έγιναν στα εδάφια [4.3], [4.6] και [4.7] αντίστοιχα.

Αν η κατανομή συχνότητας της διακριτής ποσοτικής μεταβλητής X είναι συμμετρική περί την διάμεσο M τότε η διαφορά M - Q_1 είναι ίση με την διαφορά Q_3 - M και επομένως ο συντελεστής ασυμμετρίας S_1 είναι μηδέν. Αν η κατανομή συχνότητας της διακριτής ποσοτικής μεταβλητής X παρουσιάζει ουρά προς τα δεξιά της διαμέσου M - δεξιά ή ισοδύναμα θετική ασυμμετρία - η διαφορά Q_3 - M είναι μεγαλύτερη από την διαφορά M - Q_1 και επομένως ο συντελεστής ασυμμετρίας S_1 είναι θετική τιμή. Αντίθετα αν η κατανομή συχνότητας της X παρουσιάζει ουρά προς τα αριστερά της διαμέσου M - αριστερή ή ισοδύναμα αρνητική ασυμμετρία - τότε η διαφορά Q_3 - M είναι μικρότερη από την διαφορά M - Q_1 και επομένως ο συτελεστής ασυμμετρίας S_1 έχει αρνητική τιμή.

Η διάμεσος M και τα τεταρτημόρια Q_1 και Q_3 του πίναακα 3.2 είναι

$M = 3$
$Q_1 = 2$
και
$Q_3 = 4$

Ο συντελεστής ασυμμετρίας με βάση την διάμεσο και τα τεταρτημόρια S_1 της κατανομής του πίνακα 3.2 είναι :

$$S_1 = \frac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)}$$

$$= \frac{(4 - 3) - (3 - 2)}{(4 - 3) + (3 - 2)}$$

$$= 0$$

Επομένως η κατανομή του πίνακα 3.2 είναι συμμετρική περί την διάμεσο.

[7.7] Συντελεστής Ασυμμετρίας με Βάση την Διάμεσο και τα Τεταρτημόρια

Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του συντελεστή ασυμμετρίας με βάση την διάμεσο και τα τεταρτημόρια S_1 μιάς συνεχούς ποσοτικής μεταβλητής X, γίνεται με τη χρήση του τύπου :

$$S_1 = \frac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)}$$

όπου :

[1] Q_1, το πρώτο τεταρτημόριο της μεταβλητής X,

[2] Q_3, το τρίτο τεταρτημόριο της μεταβλητής X,

[3] M, η διάμεσος της μεταβλητής X.

Ο υπολογισμός της διαμέσου M και των τεταρτημορίων Q_1 και Q_3 συνεχούς ποσοτικής μεταβλητής έγιναν στα εδάφια [4.4], [4.8] και [4.9] αντίστοιχα.

Η διάμεσος M και τα τεταρτημόρια Q_1 και Q_3 της κατανομής του πίνακα 3.3 είναι

$M = 40$
$Q_1 = 32,8$
και
$Q_3 = 50$

Ο συντελεστής ασυμμετρίας με βάση την διάμεσο και τα τεταρτημόρια S_1 της κατανομής του πίνακα 3.3 είναι :

$$S_1 = \frac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)}$$

$$= \frac{(50 - 40) - (40 - 32,8)}{(50 - 40) + (40 - 32,8)}$$

$= 0,16$

Επομένως η κατανομή του πίνακα 3.3 παρουσιάζει θετική ασυμμετρία περί την διάμεσο.

[7.8] Συντελεστής Ασυμμετρίας με Βάση τον Μέσο Αριθμητικό και την Διάμεσο Ποσοτικής Μεταβλητής.

Τα δύο επόμενα εδάφη διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού του συντελεστού ασυμμετρίας με βάση τον αριθμητικό μέσο και την διάμεσο μιάς ποσοτικής μεταβλητής.

[1] Συντελεστής Ασυμμετρίας με Βάση τον Μέσο Αριθμητικό και την Διάμεσο μιας Διακριτής Ποσοτικής Μεταβλητής.

[2] Συντελεστής Ασυμμετρίας με Βάση τον Μέσο Αριθμητικό και την Διάμεσο μιας Συνεχούς Ποσοτικής Μεταβλητής.

[7.9] Συντελεστής Ασυμμετρίας με Βάση τον Μέσο Αριθμητικό και την Διάμεσο μιας Διακριτής Ποσοτικής Μεταβλητής.

Ο υπολογισμός του συντελεστή ασυμμετρίας με βάση τον αριθμητικό μέσο και την διάμεσο S_2 μιάς διακριτής ποσοτικής μεταβλητής X γίνεται με τη χρήση του τύπου :

$$S_2 = \frac{\mu - M}{\sigma} \qquad (7.9.1)$$

όπου

[1] μ, ο αριθμητικός μέσος της μεταβλητής X,

[2] M, η διάμεσος της μεταβλητής X,

[3] σ, η τυπική απόκλιση της μεταβλητής X.

Ο υπολογισμός του αριθμητικού μέσου μ της διαμέσου M και της τυπικής απόκλισης σ διακριτής ποσοτικής μεταβλητής έγιναν στα εδάφια [4.17], [4.3] και [5.16] αντίστοιχα.

Αν η κατανομή συχνότητας της διακριτής ποσοτικής μεταβλητής X είναι συμμετρική περί την διάμεσο M τότε είναι συμμετρική και προς τον μέσο μ και επομένως η διαφορά μ - M και ο συντελεστής ασυμμετρίας S_2 είναι μηδέν. Αν η κατανομή συχνότητας της διακριτής ποσοτικής μεταβλητής X παρουσιάζει ουρά προς τα δεξιά της διαμέσου M - δεξιά ή ισοδύναμα θετική ασυμμετρία - η διαφορά μ - M είναι θετική και επομένως ο συντελεστής ασυμμετρίας S_2 έχει θετική τιμή. Αν η κατανομή συχνότητας της X παρουσιάζει ουρά προς τα αριστερά της διαμέσου M - αριστερή ή ισοδύναμα αρνητική ασυμμετρία - τότε η διαφορά μ - M είναι θετική και επομένως ο συντελεστής ασυμμετρίας S_2 έχει αρνητική τιμή.

Ο υπολογισμός της διαμέσου M της διακριτής ποσοτικής μεταβλητής του πίνακα 4.1 γίνεται με χρήση του εδαφίου [4.3]

Έχουμε $N = 120$, $N / 2 = 60$

και $F_1 = 60 \le 60 < F_2 = 80$

Επομένως η διάμεσος M είναι $M = 1$. Ο υπολογισμός του αριθμητικού μέσου μ και της τυπικής απόκλισης σ της κατανομής του πίνακα 4.1, έγιναν στα εδάφια [4.17] και [5.16] αντίστοιχα.

Επομένως

$\mu = 1,4$

$\sigma = 1{,}88$

Ο συντελεστής ασυμμετρίας με βάση τον αριθμητικό μέσο και την διάμεσο S_2 της κατανομής του πίνακα 4.1 είναι :

$$S_2 = \frac{\mu - M}{\sigma}$$

$$= \frac{1{,}4 - 1}{1{,}88}$$

$$= 0{,}2$$

Επομένως η κατανομή του πίνακα 4.1 παρουσιάζει θετική ασυμμετρία περί την διάμεσο και τον μέσο αριθμητικό.

[7.10] Συντελεστής Ασυμμετρίας με Βάση τον Μέσο Αριθμητικό και την Διάμεσο μιας Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του συντελεστή ασυμμετρίας με βάση τον αριθμητικό μέσο και την διάμεσο S_2 μιάς συνεχούς ποσοτικής μεταβλητής X γίνεται με τη χρήση του τύπου :

$$S_2 = \frac{\mu - M}{\sigma} \qquad (7.10.1)$$

όπου

[1] μ, ο αριθμητικός μέσος της μεταβλητής X,

[2] M, η διάμεσος της μεταβλητής X,

[3] σ, η τυπική απόκλιση της μεταβλητής X.

Ο υπολογισμός του αριθμητικού μέσου μ της διαμέσου M και της

τυπικής απόκλισης σ συνεχούς ποσοτικής μεταβλητής έγιναν στα εδάφια [4.4], [5.17] και [4.18] αντίστοιχα.

Ο υπολογισμός της διαμέσου M της συνεχούς ποσοτικής μεταβλητής του πίνακα 4.3 γίνεται με χρήση του εδάφιου [4.4]

Έχουμε

$$N = 145, \quad \frac{N}{2} = 72{,}5$$

και

$$F_2 = 70 < 72{,}5 < F_3 = 105$$

Η διάμεσος M περιλαμβάνει στο τρίτο διάστημα της κατανομής. Επομένως η διάμεσος M είναι

$$M = \alpha_2 + \frac{\delta_2}{f_2}\left(\frac{N}{2} - F_2\right)$$

$$= 40 + \frac{10}{35}(72{,}5 - 70)$$

$$= 40{,}7$$

Ο υπολογισμός του αριθμητικού μέσου μ και της τυπικής απόκλισης σ της κατανομής του πίνακα 4.3 έγιναν στα εδάφια [4.18] και [5.17] αντίστοιχα.

Έχουμε

$\mu = 41{,}8$

$\sigma = 12{,}62$

Ο συντελεστής ασυμμετρίας με βάση τον αριθμητικό μέσο και την διάμεσο S_2 της κατανομής του πίνακα 4.3 είναι :

$$S_2 = \frac{\mu - M}{\sigma}$$

$$= \frac{41,8 - 40,7}{12,62}$$
$$= 0,08$$

Επομένως η κατανομή του πίνακα 4.3 παρουσιάζει θετική ασυμμετρία περί τον αριθμητικό μέσο και την διάμεσο.

[7.11] Μέτρα Κυρτότητας

Τα μέτρα κυρτότητας μιας μονομεταβλητής κατανομής συχνότητας χαρακτηρίζουν την κυρτότητα, δηλαδή το πεπλατυσμένο ή την αιχμηρότητα της κατανομής συχνότητας. Το κεφάλαιο αυτό διαπραγματεύεται το ακόλουθο μέτρο κυρτότητας μονομεταβλητών κατανομών συχνότητας.

[1] Συντελεστής κυρτότητας του Pearson.

[7.12] Συντελεστής κυρτότητας του Pearson Ποσοτικής Μεταβλητής.

Τα δύο επόμενα εδάφια διαπραγματεύονται τις ακόλουθες περιπτώσεις υπολογισμού του συντελεστή κυρτότητας του Pearson μιας ποσοτικής μεταβλητής.

[1] Συντελεστής Κυρτότητας τού Pearson Διακριτής Ποσοτικής Μεταβλητής.

[2] Συντελεστής Κυρτότητας του Pearson Συνεχούς Ποσοτικής Μεταβλητής.

[7.13] Συντελεστής Κυρτότητας του Pearson Διακριτής Ποσοτικής Μεταβλητής.

Ο υπολογισμός του συντελεστή κυρτότητας του Pearson μιας διακριτής ποσοτικής μεταβλητής Χ γίνεται με τη χρήση του τύπου :

$$\beta_2 = \frac{\mu_4}{\sigma^4} \tag{7.13.1}$$

ή του τύπου

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 \tag{7.13.2}$$

όπου

[1] μ_4, η τέταρτης τάξης ροπή περί τον μέσο αριθμητικό της μεταβλητής Χ,

[2] σ, η τυπική απόκλιση της μεταβλητής Χ.

Ο υπολογισμός της μ_4 γίνεται με την χρήση του τύπου

$$\mu_4 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i (x_i - \mu)^4 \tag{7.13.3}$$

όπου

[1] x_i, i=1,2,...,κ, οι τιμές της μεταβλητής Χ,

[2] f_i, i=1,2,...,κ, οι αντίστοιχες απόλυτες συχνότητες της μεταβλητής Χ,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής Χ.

[4] μ, ο αριθμητικός μέσος της μεταβλητής Χ.

Ο χαρακτηρισμός της κυρτότητας - αιχμηρότητας ή πεπλατυσμένου - μιάς κατανομής συχνότητας γίνεται με τη σύγκριση της κατανομής συχνότητας με μία ειδική κατανομή συνχότητας η οποία λέγεται κανονική κατανομή συχνότητας. Αν η κατανομή συχνότητας παρουσιάζει αιχμηρότητα μεγαλύτερη, ίση ή μικρότερη της αιχμηρότητας της κανονικής κατανομής συχνότητας τότε χαρακτηρίζεται αντίστοιχα ως λεπτόκυρτη, μεσόκυρτη ή πλατύκυρτη. Ο συντελεστής κυρτότητας του Pearson β_2 της κανονικής κατανομής συχνότητας έχει πάντοτε την τιμή 3. Επομένως μια κατανομή συχνότητας χαρακτηρίζεται ως :

[1] λεπτόκυρτη αν $\beta_2 > 3$ ή ισοδύναμα $\gamma_2 > 0$,

[2] μεσόκυρτη αν $\beta_2 = 3$ ή ισοδύναμα $\gamma_2 = 0$,

[3] πλατύκυρτη αν $\beta_2 < 3$ ή ισοδύναμα $\gamma_2 < 0$.

Ο υπολογισμός της τετάρτης τάξης ροπής περί τον αριθμητικό μέσο μ_4 της κατανομής του πίνακα 4.1 γίνεται με την βοήθεια του πίνακα 7.1.

Πίνακας 7.1

Τιμές της X	Απόλυτες Συχνότητες	Μερικά Γινόμενα
x_i	f_i	$f_i(x_i - \mu)^4$
0	60	230,4960
1	20	0,5120
2	12	1,5552
3	9	58,9824
4	5	228,4880
5	6	1007,7696
6	8	3581,9648
	120	5109,7680

Από τον πίνακα 7.1 προκύπτει ότι :

$$\kappa = 7$$

$$N = \sum_{i=1}^{\kappa} f_i$$

$$= 120$$

$$\sum_{i=1}^{\kappa} f_i(x_i - \mu)^4 = 5109{,}768$$

Επομένως η τέταρτης τάξης ροπή περί τον αριθμητικό μέσο μ_4 της κατανομής του πίνακα 4.1 είναι :

$$\mu_4 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i(x_i - \mu)^4$$

$$= \frac{5109{,}768}{120}$$

$$= 42{,}58$$

Ο υπολογισμός της τυπικής απόκλισης σ έγινε στο εδάφιο [5.16]. Έχουμε

$$\sigma = 1{,}88$$

και

$$\mu_4 = 42{,}58$$

Ο συντελεστής κυρτότητας του Pearson β_2 της κατανομής του πίνακα 4.1 είναι :

$$\beta_2 = \frac{\mu_4}{\sigma^4}$$

$$= \frac{42{,}58}{12{,}49}$$

$$= 3{,}40.$$

Επομένως η κατανομή του πίνακα 4.1 είναι λεπτόκυρτη.

[7.14] Συντελεστής Κυρτότητας του Pearson Συνεχούς Ποσοτικής Μεταβλητής.

Ο υπολογισμός του συντελεστή κυρτότητας του Pearson μιας συνεχούς ποσοτικής μεταβλητής X γίνεται με τη χρήση του τύπου :

$$\beta_2 = \frac{\mu_4}{\sigma^4} \qquad (7.14.1)$$

ή του τύπου

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 \qquad (7.14.2)$$

όπου

[1] μ_4, η τέταρτης τάξης ροπή περί τον μέσο αριθμητικό της μεταβλητής X,

[2] σ, η τυπική απόκλιση της μεταβλητής X.

Ο υπολογισμός της μ_4 γίνεται με την χρήση του τύπου

$$\mu_4 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i (x_i - \mu)^4 \qquad (7.14.3)$$

όπου

[1] x_i, i=1,2,...,κ, οι κεντρικές τιμές της μεταβλητής X,

[2] f_i, i=1,2,...,κ, οι αντίστοιχες απόλυτες συχνότητες των τάξεων της μεταβλητής X,

[3] $N = \sum_{i=1}^{\kappa} f_i$, το πλήθος των τιμών της μεταβλητής X.

[4] μ, ο αριθμητικός μέσος της μεταβλητής X.

Ο υπολογισμός της τετάρτης τάξης ροπής περί τον αριθμητικό μέσο μ_4 της κατανομής του πίνακα 4.3 γίνεται με την βοήθεια του πίνακα 7.2.

Πίνακας 7.2

Τάξης της X	Κεντρικές Τιμές των Τάξεων x_i	Απόλυτες Συχνότητες f_i	Μερικά Γινόμενα $f_i(x_i - \mu)^4$
20-30	25	30	2389782,528
30-40	35	40	85525,504
40-50	45	35	3670,016
50-60	55	25	758989,440
60-70	65	15	4345534,464
		145	7583501,952

Από τον πίνακα 7.2 προκύπτει ότι :

κ = 5

$$N = \sum_{i=1}^{\kappa} f_i$$
$$= 145$$

$$\sum_{i=1}^{\kappa} f_i (x_i - \mu)^4 = 7583501,952$$

Επομένως η τέταρτης τάξης ροπή περί τον αριθμητικό μέσο μ_4 της κατανομής του πίνακα 4.3 είναι :

$$\mu_4 = \frac{1}{N} \sum_{i=1}^{\kappa} f_i (x_i - \mu)^4$$

$$= \frac{7583501,952}{145}$$

$$= 52300,01$$

Ο υπολογισμός της τυπικής απόκλισης σ έγινε στο εδάφιο [5.17]. Έχουμε

$$\sigma = 12,62$$

και

$$\mu_4 = 52300,01$$

Ο συντελεστής κυρτότητας του Pearson β_2 της κατανομής του πίνακα 4.3 είναι :

$$\beta_2 = \frac{\mu_4}{\sigma^4}$$

$$= \frac{52300,01}{25365,14}$$

$$= 2,06.$$

Επομένως η κατανομή του πίνακα 4.3 είναι πλατύκυρτη.