

## 2.4 Θηκογράμματα

Εφόσον γνωρίσαμε τα τεταρτημόρια και το ενδοτεταρτημοριακό πλάτος μπορούμε να αναφερθούμε και σε ένα επιπλέον τύπο γραφημάτων, τα θηκογράμματα τα οποία μεταξύ άλλων μας βοηθούν να αναγνωρίσουμε τυχόν ακραίες τιμές στο δείγμα μας. Ακραίες ή ασυνήθιστες ή παράτυπες λέγονται οι τιμές οι οποίες είναι είτε πολύ μεγαλύτερες είτε πολύ μικρότερες από τον κύριο όγκο των δεδομένων. Μπορεί να οφείλονται είτε σε λανθασμένη καταχώρηση είτε σε ιδιαίτερες συνθήκες, οι οποίες επικρατούσαν κατά την παρατήρηση τους, οπότε θα πρέπει να επανεξεταστούν. Σε κάθε περίπτωση ο ερευνητής θα πρέπει κατ' αρχήν να τις εντοπίσει και στη συνέχεια να τις εξετάσει με προσοχή. Το ενδοτεταρτημοριακό εύρος χρησιμοποιείται για την ανίχνευση ακραίων τιμών σύμφωνα με τον ακόλουθο εμπειρικό κανόνα: Οποιαδήποτε τιμή είναι πιο μακριά από  $1.5Q$  από το πλησιέστερο τεταρτημόριο θεωρείται ύποπτη για ακρότητα και θα πρέπει να εξετάζεται με προσοχή ενώ αν βρίσκεται μακρύτερα από  $3Q$  είναι ακραία ή παράτυπη και η στατιστική ανάλυση θα πρέπει να επαναληφθεί χωρίς αυτήν.

Το θηκόγραμμα βασίζεται στη σύνοψη των 5 αριθμών και δίνει μια εικόνα των πληροφοριών που περιέχονται σ' αυτή. Το θηκόγραμμα ενός συνόλου δεδομένων αποτελείται από ένα ορθογώνιο κουτί, το οποίο έχει μήκος από το πρώτο ως το τρίτο τεταρτημόριο, δηλαδή ίσο με  $Q$ . Στο σημείο της διαμέσου σημειώνεται μια κάθετος ίση με το πλάτος του κουτιού. Από κάθε πλαϊνή πλευρά του κουτιού προεκτείνουμε γραμμή μέχρι τη μικρότερη και, αντίστοιχα, τη μεγαλύτερη τιμή των δεδομένων μόνον αν είναι σε απόσταση μικρότερη από 1.5 φορές το ενδοτεταρτημοριακό εύρος  $Q$ . Διαφορετικά, προεκτείνονται ως 1.5 φορές το  $Q$ . Οι γραμμές αυτές ονομάζονται μύστακες. Ονομάζεται εσωτερικός φράχτης το διάστημα  $\pm 1.5 Q$  από τις πλαϊνές πλευρές του κουτιού, ενώ το διάστημα  $\pm 3 Q$  ονομάζεται εξωτερικός φράχτης. Τα όρια των δύο φραχτών σημειώνονται στο θηκόγραμμα καθώς και οι τιμές που περιέχονται σ' αυτές και η πρακτική που συνήθως ακολουθείται είναι η εξής:

Οι τιμές που βρίσκονται έξω από τον εσωτερικό φράχτη, θεωρούνται ύποπτες για ακρότητα και εξετάζονται με προσοχή. Οι τιμές που βρίσκονται έξω από τον εξωτερικό φράχτη κρίνονται ως ακραίες ή παράτυπες τιμές και θα πρέπει η στατιστική ανάλυση να επαναλαμβάνεται οπωσδήποτε και χωρίς αυτές.

Το θηκόγραμμα είναι μια εξαιρετικά χρήσιμη γραφική παράσταση διότι προσδιορίζει με τρόπο άμεσο

- τη θέση των δεδομένων με τη διάμεσο
- τη διασπορά τους με το μήκος του κουτιού και των μυστάκων
- την ύπαρξη τιμών οι οποίες είναι ακραίες ή ύποπτες για ακρότητα

Τέλος, τα θηκογράμματα δύο ή περισσότερων μεταβλητών όταν δίνονται συγχρόνως, επιτρέπουν την μεταξύ τους σύγκριση.