ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ

Μη παραμετρικοί έλεγχοι για την ισότητα δύο κατανομών

Αναστασία Καραματσούκη

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
Υπεύθυνος
Ιωάννης Τριανταφύλλου
Επίκουρος Καθηγητής

Λαμία, 2022

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ

Μη παραμετρικοί έλεγχοι για την ισότητα δύο κατανομών

Αναστασία Καραματσούκη

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Επιβλέπων
Ιωάννης Τριανταφύλλου
Επίκουρος Καθηγητής

Λαμία, 2022

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις (1), που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

1. *Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί* ***χωρίς να τα περικλείω σε εισαγωγικά*** *και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάσθηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.*

2. *Δέχομαι ότι η αυτολεξεί* ***παράθεση χωρίς εισαγωγικά****, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική μου εργασία.*

3. *Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια*

4. *Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.*

Ημερομηνία: 11/9/2022

Η Δηλ.

Καραματσούκη

(Υπογραφή)

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.

Μη παραμετρικοί έλεγχοι για την ισότητα δύο κατανομών

Αναστασία Καραματσούκη

Τριμελής Επιτροπή:

Ονοματεπώνυμο, Βαθμίδα …..(επιβλέπων/σα)

Ονοματεπώνυμο, Βαθμίδα……

Ονοματεπώνυμο, Βαθμίδα…….

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα πτυχιακή εργασία πραγματοποιήθηκε στο πλαίσιο των σπουδών μου στο Τμήμα Πληροφορικής με Εφαρμογές στη Βιοιατρική του Πανεπιστημίου Θεσσαλίας. Στο σημείο αυτό αισθάνομαι την ανάγκη να εκφράσω τις ειλικρινείς και θερμές ευχαριστίες μου σε όσους συνέβαλλαν στην ολοκλήρωση αυτής της προσπάθειας. Θεωρώ υποχρέωσή μου να ευχαριστήσω τον επιβλέποντα καθηγητή μου Ιωάννη Τριανταφύλλου για την καθοδήγηση που μου προσέφερε και το χρόνο που διέθεσε, δίνοντάς μου χρήσιμες συμβουλές και οδηγίες για την ολοκλήρωση της πτυχιακής μου εργασίας και για την πολύτιμη καθοδήγησή του. Επιπρόσθετα, οφείλω να αφιερώσω την πτυχιακή μου εργασία στους γονείς μου που μου συμπαραστάθηκαν όλα τα χρόνια της φοίτησης μου.

Περιεχόμενα

ΛΙΣΤΑ ΠΙΝΑΚΩΝ

ΛΙΣΤΑ ΣΧΗΜΑΤΩΝ

ΕΙΣΑΓΩΓΗ

Αν και σε πολλές περιπτώσεις οι παραμετρικές υποθέσεις είναι λογικές, συχνά δεν έχουμε προηγούμενη γνώση των κατανομών που περιγράφουν τα δεδομένα μας. Σε τέτοιες περιπτώσεις η χρήση της παραμετρικής στατιστικής μπορεί να οδηγήσει σε λανθασμένα αποτελέσματα και συνεπώς συμπεράσματα. Με άλλα λόγια, υπάρχουν αρκετά προβλήματα όταν τα δείγματα που έχουν συλλεχθεί από τον υπό-μελέτη πληθυσμό είναι μικρά και τα δείγματα δεν κατανέμονται σύμφωνα με την κανονική κατανομή. Για αυτό το λόγο είναι πολύ σημαντικό να αναφερθούμε στην μη παραμετρική στατιστική, η οποία περιγράφεται αναλυτικά στο πρώτο κεφάλαιο. Το κυριότερο πλεονέκτημα των μη παραμετρικών τεχνικών είναι ότι δεν προϋποθέτουν γνώση της κατανομής του πληθυσμού από τον οποίο έχουν προέλθει τα υπό μελέτη στοιχεία. Σκοπός της πτυχιακής εργασίας είναι η παρουσίαση μη παραμετρικών ελέγχων. Στο δεύτερο κεφάλαιο παρουσιάζονται έλεγχοι για την ισότητα των κατανομών δύο ανεξάρτητων πληθυσμών. Πιο συγκεκριμένα, γίνεται αναφορά στον έλεγχο των Kolmogorov-Smirnov, στη μέθοδο Cramer-Von Mises, Anderson-Darling και Wald-Wolfowitz. Στο τρίτο κεφάλαιο παρουσιάζονται έλεγχοι για την ισότητα των παραμέτρων θέσης δύο ανεξάρτητων πληθυσμών. Πιο συγκεκριμένα, γίνεται αναφορά στη μέθοδο Mann-Whitney, στη μέθοδο Προσήμων και στη μέθοδο Προσημικής διάταξης του Wilcoxon. Τέλος, η πτυχιακή εργασία ολοκληρώνεται στο τέταρτο κεφάλαιο, όπου παρουσιάζονται έλεγχοι για την ισότητα των παραμέτρων κλίμακας δύο ανεξάρτητων πληθυσμών. Πιο συγκεκριμένα, αναφέρεται η μέθοδος Sieger-Tukey, η μέθοδος των Τετραγωνικών Βαθμών Μεγέθους, των Freund-Ansari- Bradley, των Barton-David, των Capon-Klotz και των Flinger-Killeen.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Μη Παραμετρική Στατιστική, Μέθοδοι, Υπόθεση, Ανεξάρτητοι Πληθυσμοί, Κατανομή, Μέση τιμή, Διασπορά

ABSTRACT

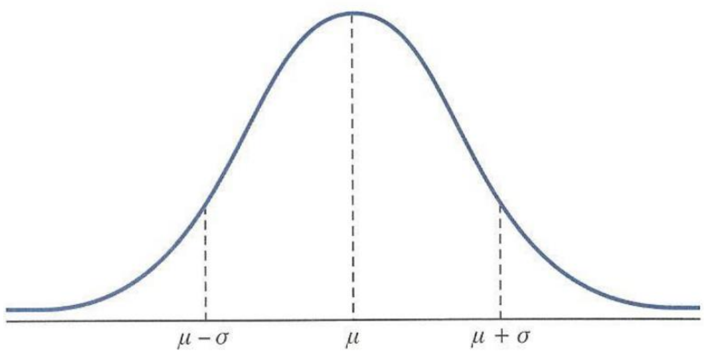
Although in many cases parametric assumptions are reasonable, we often have no prior knowledge of the distributions that describe our data. In such cases, the use of parametric statistics can lead to incorrect results and conclusions. In other words, there are several problems when the samples collected from the population under study are small and they don't follow the normal distribution. For this reason, it is very important to refer to non-parametric statistics, which is described in detail in the first chapter. The main advantage of non-parametric techniques is that they do not require knowledge of the distribution of the population from which the data under study have come. The purpose of the paper is the presentation of non-parametric methods. The second chapter presents methods for the equality of the distributions of two independent populations. More specifically, there is reference to the Kolmogorov-Smirnov method, the Cramer-Von Mises, Anderson-Darling and Wald-Wolfowitz. The third chapter presents methods for the equality of the means of two independent populations. More specifically, there is reference to the Mann-Whitney method, the sign method and the Wilcoxon sign method. Finally, the paper finishes in the fourth chapter, where methods for the equality of the scale parameters of two independent populations are presented. More specifically, the Sieger-Tukey method, the Squared marks method, Freund-Ansari-Bradley, Barton-David, Capon-Klotz and Flinger-Killeen methods are mentioned.

KEY WORDS: Non-Parametric Statistics, Methods, Hypothesis, Independent Populations, Distribution, Mean, Variance

ΚΕΦΑΛΑΙΟ 1 - Παραμετρική και μη παραμετρική στατιστική μεθοδολογία

Από τη μία πλευρά, οι παραμετρικές δοκιμές με ιδιαίτερη ισχύ και χαμηλή ανθεκτικότητα απαιτούν από τα δείγματα που εξετάζονται να ακολουθούν κανονική κατανομή. Από την άλλη πλευρά, οι μη παραμετρικές δοκιμές έχουν χαμηλότερη ισχύ από τις παραμετρικές, αλλά μεγαλύτερη ανθεκτικότητα και δεν απαιτούν την κανονική κατανομή των δεδομένων. Οι παραμετρικές μέθοδοι χρησιμοποιούνται συχνά για στατιστικές αναλύσεις και συχνά προτιμώνται έναντι των μη παραμετρικών ελέγχων επειδή τα αποτελέσματά τους είναι πιο αξιόπιστα. Ωστόσο, είναι πολύ συνηθισμένο ότι η κατανομή των δεδομένων δεν είναι γνωστή, επομένως δεν πληρούνται οι προϋποθέσεις για παραμετρικό έλεγχο. Εμφανίζονται επίσης περιπτώσεις, όπου ακόμα και αν υπάρχει κάποια ένδειξη κανονικότητας, ο ερευνητής που ενδεχομένως έχει βαθειά γνώση του πληθυσμού είναι επιφυλακτικός στο να κάνει μια τέτοια υπόθεση.

Οι στατιστικοί επινόησαν αρκετές εναλλακτικές τεχνικές για τους ερευνητές που "διστάζουν" να υποθέσουν κανονικότητα, Οι τεχνικές αυτές δεν απαιτούν συγκεκριμένες υποθέσεις για την μορφή του πληθυσμού από τον οποίο προέρχονται τα δεδομένα και μπορούν να χρησιμοποιηθούν τόσο για μικρά όσο και για μεγάλα δείγματα. Είναι δηλαδή σχεδιασμένες για να μπορούν να χρησιμοποιηθούν ανεξάρτητα από την κατανομή των δεδομένων. Επομένως, οι τεχνικές αυτές μπορούν να χρησιμοποιηθούν τόσο για κανονικούς πληθυσμούς όσο και για μη κανονικούς πληθυσμούς. Για το λόγο αυτό, οι τεχνικές αυτές ονομάζονται μη παραμετρικές τεχνικές, οι οποίες δεν έχουν περιορισμούς και είναι εξίσου, και κάποιες φορές, περισσότερο ισχυρές.



Σχήμα 1: Η γραφική παράσταση της κανονικής κατανομής(μ: μέση τιμή, σ^2: διακύμανση)

Οι μη παραμετρικές τεχνικές είναι πολύ απλές στη χρήση τους. Εάν τα δεδομένα ακολουθούν την κανονική κατανομή, τότε οι μη παραμετρικοί έλεγχοι υποθέσεων δεν είναι το ίδιο ισχυροί όπως οι αντίστοιχοι παραμετρικοί έλεγχοι, οι οποίοι κάνουν χρήση της υπόθεσης της κανονικότητας, Ένας έλεγχος, ο οποίος αγνοεί πληροφορίες σχετικά με τα δεδομένα, όπως είναι η μορφή της κατανομής τους, δεν αναμένεται να είναι το ίδιο καλός όπως ένας έλεγχος ο οποίος κάνει χρήση αυτής της πληροφορίας. Έτσι, ένας έλεγχος, ο οποίος δεν λαμβάνει υπόψιν του ότι τα δεδομένα προέρχονται από μια κανονική κατανομή, δεν αναμένεται να είναι το ίδιο ισχυρός όπως ένας έλεγχος ο οποίος χρησιμοποιεί αυτή την υπόθεση. Από το άλλο μέρος, εάν τα δεδομένα προέρχονται από μη κανονικό πληθυσμό, τότε οι μη παραμετρικοί έλεγχοι έχουν ένα σαφές πλεονέκτημα έναντι των παραμετρικών ελέγχων, οι οποίοι στηρίζονται στην εσφαλμένη υπόθεση της κανονικότητας των δεδομένων.

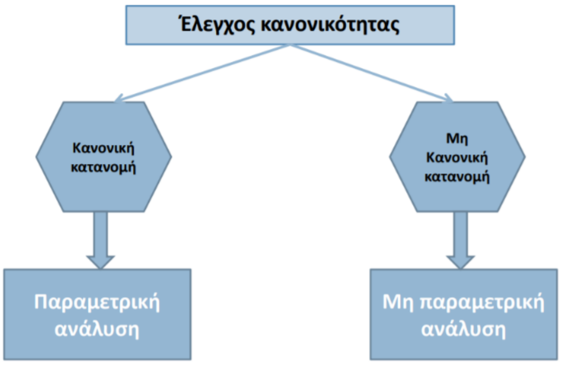
Οι μη παραμετρικές μέθοδοι ονομάζονται έτσι επειδή δεν χρησιμοποιούν παραμέτρους όπως ο μέσος όρος ή η μεταβλητότητα για να καταρρίψουν ή να επιβεβαιώσουν τη μηδενική υπόθεση, αλλά αντίθετα χρησιμοποιούν άλλες στατιστικές έννοιες όπως διάμεσος και εύρος. Ορισμένες από αυτές τις μεθόδους δεν χρησιμοποιούν τα ίδια τα δεδομένα, αλλά ταξινομημένες κλάσεις αυτών, γεγονός που τις ευνοεί να προσαρμοστούν σε μικρά μεγέθη δειγμάτων. Ωστόσο, μερικές φορές χάνεται ένα μέρος της πληροφορίας. Αν και ο μη παραμετρικός έλεγχος χαρακτηρίζεται από μεγαλύτερη αβεβαιότητα και χαμηλότερη απόδοση από τον παραμετρικό έλεγχο, πρέπει να επισημανθεί ότι ο μη παραμετρικός έχει αποδειχθεί αντάξιος και, σε ορισμένες περιπτώσεις, ακόμη πιο ισχυρός από τον παραμετρικό έλεγχο. Επιπλέον, η μεγάλη ανθεκτικότητά τους δεν τους επιτρέπει να εμφανίζουν αλλαγές στα αποτελέσματα λόγω έλλειψης τιμών. Όπως είναι φανερό οι μη παραμετρικοί έλεγχοι, θεωρητικά τουλάχιστον, αποτελούν τους καταλληλότερους στατιστικούς ελέγχους για τα δεδομένα που δεν ακολουθούν κανονική κατανομή ή όταν αυτή είναι απροσδιόριστη.

Συνοπτικά, οι μη παραμετρικές μέθοδοι:

1. Αποβλέπουν σε ευρύτερα πεδία εφαρμογής
2. Δεν είναι εξίσου ισχυρές με τις αντίστοιχες παραμετρικές μεθόδους
3. Είναι περισσότερο ευσταθείς επειδή δεν επηρεάζονται από την μορφή της κατανομής των δεδομένων
4. Είναι σχεδόν το ίδιο αποτελεσματικές
5. Μπορούν να εφαρμοστούν σε δεδομένα που είναι ταξινομημένα σε κατηγορίες(κατηγορικά δεδομένα) και τα οποία είναι σε κλίμακα διάταξης, ενώ οι παραμετρικές μέθοδοι προϋποθέτουν ακριβείς μετρήσεις
6. Εφαρμόζονται μέσω στατιστικών πακέτων που έχουν εξελιχθεί τα τελευταία χρόνια σε απαραίτητα εργαλεία

Πώς όμως επιλέγουμε το κατάλληλο στατιστικό έλεγχο;

- Βήμα 1: Θέτουμε το ερώτημα της έρευνας
- Βήμα 2: Ελέγχουμε εάν τα δεδομένα ακολουθούν κανονική κατανομή ή όχι
- Βήμα 3: Εξετάζουμε πόσες ομάδες δεδομένων έχουμε
- Βήμα 4: Ελέγχω αν τα δείγματα μου είναι εξαρτημένα ή ανεξάρτητα
- Βήμα 5: Επιλέγω το στατιστικό έλεγχο



Σχήμα 2: Έλεγχος κανονικότητας

Για την εφαρμογή ενός στατιστικού ελέγχου, διαδραματίζεται μια συγκεκριμένη διαδικασία. Αρχικά, πρέπει να οριστεί η μηδενική υπόθεση, με συμβολισμό H_0, και η εναλλακτική υπόθεση H_1 που ισχύει όταν απορρίπτεται η μηδενική υπόθεση. Η H_0 είναι αυτή για την οποία υπάρχει η αμφιβολία. Ουσιαστικά, πραγματοποιείται ο στατιστικός έλεγχος για να αποφασιστεί αν θα απορριφθεί η μηδενική υπόθεση και το να μην απορριφθεί δεν συνεπάγεται ότι επιβεβαιώνεται. Εάν δεν απορριφθεί, σημαίνει ότι τα δεδομένα που εξετάστηκαν μπορεί να μην είναι αρκετά για την απόρριψη της μηδενικής υπόθεσης.

Υπάρχουν δύο τύποι σφαλμάτων που μπορούν να προκύψουν από τους στατιστικούς ελέγχους.

Μια περίπτωση σφάλματος είναι το σφάλμα τύπου I(ή **σφάλμα του πρώτου είδους)**, όπου η μηδενική υπόθεση απορρίπτεται όταν ισχύει και δεν θα έπρεπε να απορριφθεί. Συνήθως ένα σφάλμα τύπου I οδηγεί στο συμπέρασμα μια υποτιθέμενη επίδραση ή σχέση υπάρχει όταν στην πραγματικότητα δεν υπάρχει. Παραδείγματα σφαλμάτων τύπου A περιλαμβάνουν έναν έλεγχο που δείχνει έναν ασθενή να έχει μια ασθένεια, ενώ στην πραγματικότητα ο ασθενής δεν έχει την ασθένεια, έναν συναγερμό για φωτιά να έχει ένδειξη για φωτιά όταν στην πραγματικότητα δεν υπάρχει φωτιά, ή ένα πείραμα που δείχνει ότι μια ιατρική θεραπεία θα θεραπεύσει μια ασθένεια, όταν στην πραγματικότητα δεν το κάνει.

Η άλλη περίπτωση σφάλματος είναι το σφάλμα τύπου II(ή **σφάλμα του δεύτερου είδους)**, όπου η μηδενική υπόθεση δεν απορρίπτεται όταν δεν ισχύει και θα έπρεπε να απορριφθεί. Παραδείγματα με τα σφάλματα τύπου B θα είναι μια εξέταση αίματος που αποτυγχάνει να ανιχνεύσει την ασθένεια που είναι σχεδιασμένο για να ανιχνεύει σε έναν ασθενή που έχει πραγματικά την ασθένεια, να ξεσπάσει φωτιά και ο συναγερμός πυρκαγιάς να μην παράγει ήχο ή μια κλινική δοκιμή μιας ιατρική θεραπεία αποτυγχάνει να δείξει ότι η θεραπεία δεν λειτουργεί όταν πραγματικά λειτουργεί.

Όλοι οι στατιστικοί έλεγχοι υποθέσεων έχουν μια πιθανότητα να κάνουν σφάλματα τύπου A και τύπου B. Όπως είναι αντιληπτό, το αν μια στατιστική μέθοδος υπόκειται σε αυτά τα σφάλματα παίζει σημαντικό ρόλο στο πόσο ισχυρή είναι σε σχέση με άλλες μεθόδους.

	Επιλέγω H_0	**Επιλέγω H_1**
H_0 Αληθινή	**Σωστή Απόφαση**	**Σφάλμα Τύπου Ι**
H_1 Αληθινή	**Σφάλμα Τύπου ΙΙ**	**Σωστή Απόφαση**

Σχήμα 3: Σφάλμα τύπου Ι και ΙΙ

Για να αξιολογηθεί το αποτέλεσμα του στατιστικού ελέγχου που εφαρμόστηκε, δηλαδή την αποδοχή ή την απόρριψη της μηδενικής υπόθεσης, πρέπει να ελεγχθεί η πιθανότητα το αποτέλεσμα να είναι τυχαίο. Αυτό γίνεται με τον υπολογισμό της στατιστικής σημαντικότητας, η οποία αντιπροσωπεύεται από μια τιμή p και δείχνει εάν το αποτέλεσμα είναι αντιπροσωπευτικό της πραγματικότητας, δηλαδή του πληθυσμού. Ο ερευνητής θέτει ένα τυχαίο επίπεδο σημαντικότητας α, συνήθως 95% (p = 0,05), το οποίο είναι το κατώτερο όριο στο οποίο ισχύει η μηδενική υπόθεση ή με άλλα λόγια είναι και το ανεκτό επίπεδο σφάλματος τύπου I που προκαθορίζεται. Εάν η υπολογισμένη τιμή p είναι μικρότερη από α, τότε η μηδενική υπόθεση απορρίπτεται. Επομένως, όσο πιο χαμηλή τιμή έχει το p, τόσο πιο αξιόπιστο είναι το αποτέλεσμα του στατιστικού ελέγχου.

Εάν για παράδειγμα έχουμε επίπεδο σημαντικότητας 0,05, τότε αυτό σημαίνει ότι υπάρχει λιγότερο από 5% πιθανότητα να επιλεγεί μια υπόθεση λανθασμένα.

Επομένως, η διαδρομή ενός στατιστικού ελέγχου είναι:

- Περιγραφή του πληθυσμού που μελετάται και διατύπωση των περιορισμών και συνθηκών που είναι αναγκαίοι για την μελέτη
- Προσδιορισμός της παραμέτρου και διατύπωση στατιστικών υποθέσεων
- Καθορισμός της μηδενικής υπόθεσης H_0
- Καθορισμός της εναλλακτικής υπόθεσης H_1
- Επιλογή του επιπέδου σημαντικότητας α και της κατάλληλης στατιστικής τεχνικής που θα χρησιμοποιηθεί
- Προσδιορισμός στατιστικής συνάρτησης ελέγχου και της κατανομής της
- Προσδιορισμός περιοχών απόρριψης και συνθηκών αποδοχής της μηδενικής υπόθεσης
- Υπολογισμός της τιμής στατιστικής ελέγχου
- Λήψη της κατάλληλης στατιστικής απόφασης, σύνδεση με την υπόθεση εργασίας
- Διατύπωση της στατιστικής αυτής απόφασης στην ορολογία του προβλήματος
- Συμπέρασμα

Πιο συγκεκριμένα, η στατιστική συνάρτηση ελέγχου είναι πολύ σημαντική γιατί μάς δείχνει μια ένδειξη της απόλυτης διαφοράς της δειγματικής εκτίμησης από την υποθετική τιμή της πληθυσμιακής παραμέτρου. Ο υπολογισμός της πιθανότητας η στατιστική συνάρτηση ελέγχου να πάρει τιμές μεγαλύτερες ή μικρότερες από το επίπεδο σημαντικότητας α είναι δυνατή αφού γνωρίζουμε την κατανομή της στατιστικής συνάρτησης ελέγχου. Επομένως, με την στατιστική συνάρτηση ελέγχου και τα όρια που θέτουμε για αυτή, ο παραμετρικός χώρος του πληθυσμού χωρίζεται σε δύο περιοχές, την περιοχή αποδοχής και την περιοχή απόρριψης.

- Περιοχή αποδοχής: είναι εκείνη η περιοχή τιμών στην οποία αν βρεθεί η τιμή της στατιστικής συνάρτησης ελέγχου δεν απορρίπτουμε την μηδενική υπόθεση H_0
- Περιοχή απόρριψης: είναι εκείνη η περιοχή τιμών στην οποία αν βρεθεί η τιμή της στατιστικής συνάρτησης ελέγχου απορρίπτουμε την μηδενική υπόθεση H_0

Αξίζει να αναφερθεί ότι ο όρος "αποδεχόμαστε την μηδενική υπόθεση H_0" αποφεύγεται στην στατιστική. Προτιμούνται και χρησιμοποιούνται οι όροι "απορρίπτουμε" και "δεν απορρίπτουμε".

Όσον αφορά το επίπεδο σημαντικότητας, η τιμή του εξαρτάται από τη φύση του προβλήματος που μελετάμε αλλά και από την γνώση και την πληροφόρηση που έχουμε για αυτό. Η σωστή τιμή για το α θα πρέπει να μας ελαχιστοποιεί την πιθανότητα για σφάλμα τύπου II. Στην στατιστική έχει επικρατήσει το επίπεδο σημαντικότητας $\alpha = 5\%$ =0,05. Πολλοί αναρωτιούνται όμως γιατί στην προσέγγιση του ελέγχου σημαντικότητας χρησιμοποιούνται κυρίως οι τιμές 5% και 1%. Ο λόγος είναι ότι για τις τιμές αυτές έχουν κατασκευασθεί πίνακες που επιτρέπουν την λήψη απόφασης. Ο R. A. Fisher ήταν ο πρώτος που δημοσίευσε τέτοιους πίνακες και τους κατασκεύασε με αυτό τον τρόπο. Δεδομένου ότι υπάρχει περιορισμένος χώρος σε μια σελίδα χαρτιού, οι πίνακες θα έπρεπε να περιορισθούν σε ορισμένες μόνο τιμές του α. Το 5% και 1% θεωρήθηκαν "καλές" τιμές και από τότε χρησιμοποιούνται παραδοσιακά, χωρίς να υπάρχει ειδικός λόγος. Με την διάδοση των υπολογιστών, τέτοιοι πίνακες είναι πλέον ελάχιστα χρήσιμοι. Το ίδιο συμβαίνει με το 5% και 1% ως τιμές του α αλλά και με την προσέγγιση του ελέγχου σημαντικότητας. Η προσέγγιση της p-τιμής έχει πια κυριαρχήσει στην Στατιστική Συμπερασματολογία.

Εξίσου σημαντική όμως, είναι και η εκτίμηση ενός διαστήματος εμπιστοσύνης, μιας και βοηθάει στον καθορισμό των υποθέσεων που πρέπει να ελεγχτούν. Αυτό συμβαίνει διότι η τιμή που ελέγχουμε μπορεί να περιέχεται ή και όχι στο αντίστοιχο $(1-\alpha)\%$ διάστημα εμπιστοσύνης και έτσι μπορούμε να οδηγηθούμε στο συμπέρασμα της απόρριψης ή της μη απόρριψης των υποθέσεων μας.

ΚΕΦΑΛΑΙΟ 2- Έλεγχοι για την ισότητα των κατανομών δύο ανεξάρτητων πληθυσμών

Στην έρευνα, στην περίπτωση που έχουμε δύο ανεξάρτητα δείγματα παρατηρήσεων, πολλές φορές τίθεται το ερώτημα αν αυτά προέρχονται από τον ίδιο πληθυσμό(ή αλλιώς αν έχουν την ίδια κατανομική θέση). Συγκεκριμένα, ο ερευνητής ενδέχεται να ενδιαφέρεται να προσδιορίσει αν οι δύο συναρτήσεις κατανομής που σχετίζονται με τους δύο πληθυσμούς ταυτίζονται ή όχι. Οι έλεγχοι καλής προσαρμογής υποδεικνύουν πόσο καλά ένα σύνολο δεδομένων περιγράφεται, προσαρμόζεται από ή σε ένα συγκεκριμένο μοντέλο. Με άλλα λόγια, έχουν αναπτυχθεί διάφορες στατιστικές μεθοδολογίες για τον έλεγχο αν δύο σύνολα δεδομένων μπορούν να θεωρηθούν ότι προέρχονται από τον ίδιο πληθυσμό. Οι έλεγχοι που θα μελετηθούν παρακάτω είναι κατάλληλοι στις περιπτώσεις αυτές γιατί είναι συνεπείς έναντι όλων των τύπων διαφορών που μπορεί να υπάρχουν μεταξύ δύο συναρτήσεων κατανομών(για παράδειγμα διαφορές στις μέσες τιμές τους ή στις διαμέσους τους).

Μέθοδος Kolmogorov-Smirnov

Ο έλεγχος Kolmogorov – Smirnov για δύο δείγματα εξετάσθηκε από τον Smirnov το 1939 και είναι μια μορφή ελέγχου Kolmogorov για την περίπτωση δύο ανεξάρτητων δειγμάτων. Βασίστηκε - επέκτεινε την πρόταση του Kolmogorov, ο οποίος πρότεινε το 1933 μια στατιστική συνάρτηση που ουσιαστικά αποτελεί ένα μέτρο της εγγύτητας των συναρτήσεων κατανομής που μελετιούνται στη συνέχεια. Έχει πάρει το όνομά του από τους Andrey Kolmogorov (1903-1987) και Nikolai Smirnov(1900-1966).

Έστω δύο ανεξάρτητα τυχαία δείγματα $X_1, X_2, \dots, X_n$ και $Y_1, Y_2, \dots, Y_m$, όπου n και m ο αριθμός των παρατηρήσεων στις τυχαίες μεταβλητές X και Y αντίστοιχα. Ορίζουμε $F_x(x)$, όπου $x \in (-\infty, +\infty)$ και $F_y(y)$, όπου $y \in (-\infty, +\infty)$, τις αντίστοιχες άγνωστες συναρτήσεις κατανομής των X και Y.

Οι υποθέσεις που θα ελεγχθούν είναι οι εξής:

- H_0: $F_x(x) = F_y(x)$, για κάθε $x \in (-\infty, +\infty)$
- H_1: $F_x(x) \neq F_y(x)$, για τουλάχιστον ένα $x \in (-\infty, +\infty)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν την ίδια κατανομή, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν την ίδια κατανομή.

Δηλαδή:

- H_0: Έχουν την ίδια κατανομή
- H_1: Δεν έχουν την ίδια κατανομή

Η στατιστική συνάρτηση με την οποία θα ελέγξουμε τις συγκεκριμένες υποθέσεις στηρίζεται στις εμπειρικές κατανομές των δύο δειγμάτων παρατηρήσεων. Για το τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_x(x)$, ενώ για το τυχαίο δείγμα $Y_1, Y_2, \ldots, Y_m$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_y(y)$. Ας μην ξεχνάμε τον ορισμό της $S(x)$: Η $S(x)$ είναι το ποσοστό των παρατηρήσεων του δείγματος οι οποίες είναι το πολύ ίσες με την τιμή x, για κάθε x στο διάστημα $(-\infty, +\infty)$.

Επομένως, έχουμε την στατιστική συνάρτηση ελέγχου: $T_1 = sup_x \, |S_x(x) - S_y(x)|$ και διαβάζεται "η T_1 είναι ίση με το supremum, για όλα τα x της απόλυτης τιμής της διαφοράς $S_x(x) - S_y(x)$, δηλαδή είναι ίση με το μεγαλύτερο στοιχείο του υποσυνόλου".

Παρατηρείται ότι μια ένδειξη εναντίον της μηδενικής υπόθεσης αποτελούν οι μεγάλες τιμές στην ελεγχοσυνάρτηση.

Κανόνας απόφασης: Η μηδενική απόφαση απορρίπτεται σε επίπεδο σημαντικότητας α, αν η τιμή της κατάλληλης στατιστικής συνάρτησης υπερβαίνει το $(1-\alpha)$-ποσοστιαίο σημείο της κατανομής της, όπως αυτό δίνεται στον πίνακα 1 του παραρτήματος για $n = m$ και στον πίνακα 2 του παραρτήματος για $n \neq m$. Δηλαδή, αν $T_1 > w_{1-a}$.

Παρατήρηση: Στη βιβλιογραφία έχουν μελετηθεί αρκετές παραλλαγές του ελέγχου Kolmogorov-Smirnov, οι οποίες επιτρέπουν τη χρήση του σε περιπτώσεις όπου οι παράμετροι εκτιμώνται από τα δεδομένα. Οι σημαντικότερες από αυτές είναι οι Cramer- Von Mises(1928), Kuiper(1960), Lilliefors(1967), Watson(1957) και Anderson-Darling(1952).

Παράδειγμα 1

Σε ένα τυχαίο δείγμα εννιά μαθητών της έκτης τάξης ενός Δημοτικού σχολείου που ανήκε σε μία περιοχή κάποιας πόλης, δόθηκε ένα διαγώνισμα γνώσεων. Σε ένα άλλο τυχαίο δείγμα δεκαπέντε μαθητών της έκτης τάξης ενός άλλου Δημοτικού σχολείου που ανήκε σε μία διαφορετική περιοχή της πόλης, δόθηκε το ίδιο διαγώνισμα γνώσεων. Τα αποτελέσματα βρίσκονται στον παρακάτω πίνακα:

	Αποτελέσματα
1ο δείγμα	7.6, 9.9, 8.4, 10.1, 8.6, 10.6, 8.7, 11.2, 9.3
2ο δείγμα	5.2, 8.2, 11.5, 5.7, 9.1, 12.3, 5.9, 9.8, 12.5, 6.5, 10.8, 13.4, 6.8, 14.6, 11.3

Πίνακας 1: Αποτελέσματα μαθητών

Εφαρμόζοντας τον έλεγχο Smirnov θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν υπάρχει διαφορά στις γνώσεις, όπως αυτές μετριούνται με αυτό το διαγώνισμα, των δύο πληθυσμών μαθητών της έκτης τάξης.

Ορίζεται η μηδενική υπόθεση: Οι δύο πληθυσμοί έχουν την ίδια συνάρτηση κατανομής και η εναλλακτική υπόθεση: Οι δύο πληθυσμοί δεν έχουν την ίδια συνάρτηση κατανομής.

Επομένως, σύμφωνα με αυτά που έχουν προηγηθεί, η μηδενική και η εναλλακτική υπόθεση μπορούν να γραφτούν με τον ακόλουθο τρόπο:

- H_0: $F_x(x) = F_y(x)$, για κάθε $x \in (-\infty, +\infty)$
- H_1: $F_x(x) \neq F_y(x)$, για τουλάχιστον ένα $x \in (-\infty, +\infty)$

Όπου $F_x(x)$ και $F_y(y)$ οι αντίστοιχες συναρτήσεις κατανομής των X και Y. Για το τυχαίο δείγμα $X_1, X_2, \dots, X_9$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_x(x)$, ενώ για το τυχαίο δείγμα $Y_1, Y_2, \dots, Y_{15}$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_y(y)$.

Στον παρακάτω πίνακα βρίσκονται τα δύο δείγματα διατεταγμένα κατά αύξουσα σειρά μεγέθους, σαν να αποτελούν ένα ενιαίο δείγμα, αλλά και οι τιμές τις διαφοράς $S_x(x) - S_y(x)$ των εμπειρικών συναρτήσεων κατανομής.

X_i	Y_i	$S_x(x) - S_y(x)$	X_i	Y_i	$S_x(x) - S_y(x)$
	5.2	0-1/15= -1/15		9.8	5/9-8/15=1/45
	5.7	0-2/15= -2/15	9.9		6/9-8/15= 2/15
	5.9	0-3/15= -1/5	10.1		7/9-8/15= 11/45
	6.5	0-4/15= -4/15	10.6		8/9-8/15=16/45
	6.8	0-5/15= -1/3		10.8	8/9-9/15= 13/45
7.6		1/9-5/15= -2/9	11.2		1-9/15= 2/5
	8.2	1/9-6/15= -13/45		11.3	1-10/15= 1/3
8.4		2/9-6/15= -8/45		11.5	1-11/15=4/15
8.6		3/9-6/15= -1/15		12.3	1-12/15= 1/5
8.7		4/9-6/15= 2/45		12.5	1-13/15= 2/15
	9.1	4/9-7/15= -1/45		13.4	1-14/15= 1/15
9.3		5/9-7/15=4/45		14.6	1-1= 0

Πίνακας 2: Τα δύο δείγματα και η διαφορά $S_x(x) - S_y(x)$

Η κατάλληλη στατιστική συνάρτηση με την οποία θα ελέγξουμε τις συγκεκριμένες υποθέσεις στηρίζεται στις εμπειρικές κατανομές των δύο δειγμάτων παρατηρήσεων και είναι η: $T_1 = sup_x \,|S_x(x) - S_y(x)|$.

Σύμφωνα με τον παραπάνω πίνακα, η μεγαλύτερη τιμή της στατιστικής συνάρτησης T_1 βρίσκεται στην τιμή 11.2, δηλαδή:

$$T_1 = sup_x \,|S_x(x) - S_y(x)|.$$

$$= |S_x(11.2) - S_y(11.2)|$$

$= 2/5$

$= 0.4$

Θεωρούμε επίπεδο σημαντικότητας $\alpha = 0.05$.

Κάνοντας χρήση του πίνακα 2, παρατηρούμε ότι το 0.95-ποσοστιαίο σημείο της κατανομής της στατιστικής συνάρτησης T_1 για αμφίπλευρο έλεγχο, για $N_1 = 9$ και N_2 =15, είναι $w_{0.95} = 8/15$.

Συγκρίνοντας την T_1 με το $w_{0.95}$ έχουμε: $T_1 = 2/5 < w_{0.95} = 8/15$.
Για να αποφασίσουμε αν θα απορρίψουμε ή όχι την μηδενική υπόθεση δεν πρέπει να ξεχνάμε τον κανόνα απόφασης.

Κανόνας απόφασης: Η μηδενική απόφαση απορρίπτεται σε επίπεδο σημαντικότητας α, αν η τιμή της κατάλληλης στατιστικής συνάρτησης υπερβαίνει το $(1-\alpha)$-ποσοστιαίο σημείο της κατανομής της, όπως αυτό δίνεται στον πίνακα 1 του παραρτήματος για $n = m$ και στον πίνακα 2 του παραρτήματος για $n \neq m$. Δηλαδή, αν $T_1 > w_{1-a}$.

Επομένως, σύμφωνα με τα προηγούμενα, η μηδενική υπόθεση H_0 δεν απορρίπτεται σε επίπεδο σημαντικότητας 5%, δηλαδή δεν απορρίπτεται ότι οι δύο πληθυσμοί έχουν την ίδια συνάρτηση κατανομής.

Παράδειγμα 2

Ένας φοιτητής του Πανεπιστημίου Θεσσαλίας έχει στη βιβλιοθήκη του 257 βιβλία, από τα οποία τα 114 είναι βιβλία στατιστικής και τα 143 είναι βιβλία πληροφορικής. Ο φοιτητής επιλέγει δύο ανεξάρτητα τυχαία δείγματα βιβλίων, ένα από κάθε κατηγορία, μεγέθους 12 και 16 αντίστοιχα και καταγράφει τους αριθμούς σελίδων τους.

	Αριθμός σελίδων
Βιβλία στατιστικής	126, 142, 156, 228, 245, 246, 370, 419, 433, 454, 478, 503
Βιβλία πληροφορικής	29, 39, 60, 78, 82, 112, 125, 170, 192, 224, 263, 275, 276, 286, 369, 756

Πίνακας 3: Τα δύο δείγματα βιβλίων και οι αριθμοί σελίδων τους

Εφαρμόζοντας τον έλεγχο Smirnov θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν ο πληθυσμός των βιβλίων στατιστικής διαφέρει από τον πληθυσμό των βιβλίων πληροφορικής ως προς τον αριθμό σελίδων, με βάση τα δύο ανεξάρτητα δείγματα που επέλεξε ο φοιτητής.

Ορίζεται η μηδενική υπόθεση: Οι δύο πληθυσμοί έχουν την ίδια συνάρτηση κατανομής και η εναλλακτική υπόθεση: Οι δύο πληθυσμοί δεν έχουν την ίδια συνάρτηση κατανομής.

Επομένως, η μηδενική και η εναλλακτική υπόθεση μπορούν να γραφτούν με τον ακόλουθο τρόπο:

- H_0: $F_x(x) = F_y(x)$, για κάθε $x \in (-\infty, +\infty)$
- H_1: $F_x(x) \neq F_y(x)$, για τουλάχιστον ένα $x \in (-\infty, +\infty)$

Όπου $F_x(x)$ και $F_y(y)$ οι αντίστοιχες συναρτήσεις κατανομής των X και Y. Για το τυχαίο δείγμα $X_1, X_2, \ldots, X_{12}$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_x(x)$, ενώ για το τυχαίο δείγμα $Y_1, Y_2, \ldots, Y_{16}$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_y(y)$.

Στον παρακάτω πίνακα βρίσκονται τα δύο δείγματα διατεταγμένα κατά αύξουσα σειρά μεγέθους, σαν να αποτελούν ένα ενιαίο δείγμα, αλλά και οι τιμές τις διαφοράς $S_x(x) - S_y(x)$ των εμπειρικών συναρτήσεων κατανομής.

X_i	Y_i	$S_x(x) - S_y(x)$	X_i	Y_i	$S_x(x) - S_y(x)$
29		0.0625		245	0.2083
39		0.1250		246	0.1250
60		0.1875	263		0.1875
78		0.2500	275		0.2500
82		0.3125	276		0.3125
112		0.3750	286		0.3750
125		0.4375	369		0.4375
	126	0.3542		370	0.3542
	142	0.2708		419	0.2708
	156	0.1875		433	0.1875
170		0.2500		454	0.1042
192		0.3125		478	0.0208
224		0.3750		503	0.0625
	228	0.2917	756		0.0000

Πίνακας 4: Τα δύο δείγματα και τη διαφορά $S_x(x) - S_y(x)$

Η κατάλληλη στατιστική συνάρτηση με την οποία θα ελέγξουμε τις συγκεκριμένες υποθέσεις στηρίζεται στις εμπειρικές κατανομές των δύο δειγμάτων παρατηρήσεων και είναι η: $T_1 = sup_x \, |S_x(x) - S_y(x)|$.

Σύμφωνα με τον παραπάνω πίνακα, η μεγαλύτερη τιμή της στατιστικής συνάρτησης T_1 βρίσκεται στην τιμή 369, δηλαδή:

$$T_1 = sup_x \, |S_x(x) - S_y(x)|$$

$$= |S_x(369) - S_y(369)|$$

$$= 0.4375$$

Θεωρούμε επίπεδο σημαντικότητας $\alpha = 0.05$.

Κάνοντας χρήση του πίνακα 2, παρατηρούμε ότι το 0.95-ποσοστιαίο σημείο της κατανομής της στατιστικής συνάρτησης T_1 για αμφίπλευρο έλεγχο, για $N_1 = 12$ και $N_2 = 16$, είναι $w_{0.95} =$ 23/48 = 0.4792.

Συγκρίνοντας την T_1 με το $w_{0.95}$ έχουμε: $T_1 = 0.4375 < w_{0.95} = 0.4792$.

Επομένως, σύμφωνα με τον κανόνα απόφασης, η μηδενική υπόθεση H_0 δεν απορρίπτεται σε επίπεδο σημαντικότητας 5%, δηλαδή δεν απορρίπτεται ότι οι δύο πληθυσμοί έχουν την ίδια συνάρτηση κατανομής.

Μέθοδος Cramer- Von Mises

Ένας εναλλακτικός αμφίπλευρος έλεγχος της υπόθεσης ισότητας δύο κατανομών είναι η μέθοδος Cramer – Von Mises. Η διαφορά αυτού του ελέγχου με τον έλεγχο Smirnov είναι ότι υπάρχει μια δυσκολία στο να υπολογισθεί η ελεγχοσυνάρτηση του. Παρόλα αυτά, χρησιμοποιείται συχνά μιας και κάνει περισσότερο αποτελεσματική χρήση των δεδομένων. Στην πραγματικότητα, όμως, η διαφορά στην ισχύ των δύο ελέγχων είναι ελάχιστη. Ο έλεγχος αυτός πήρε το όνομά του από τον Harald Cramer (1893-1985) και τον Richard von Mises (1883-1953), οι οποίοι το πρότειναν για πρώτη φορά το 1928-1930.

Έστω δύο ανεξάρτητα τυχαία δείγματα $X_1, X_2, \dots, X_n$ και $Y_1, Y_2, \dots, Y_m$, όπου n και m ο αριθμός των παρατηρήσεων στις τυχαίες μεταβλητές X και Y αντίστοιχα. Ορίζουμε $F_x(x)$, όπου $x \in (-\infty, +\infty)$ και $F_y(y)$όπου $y \in (-\infty, +\infty)$τις αντίστοιχες άγνωστες συναρτήσεις κατανομής των X και Y.

Οι υποθέσεις που θα ελεγχθούν είναι οι εξής:

- H_0: $F_x(x) = F_y(x)$, για κάθε $x \in (-\infty, +\infty)$
- H_1: $F_x(x) \neq F_y(x)$, για τουλάχιστον ένα $x \in (-\infty, +\infty)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν την ίδια κατανομή, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν την ίδια κατανομή.

Δηλαδή:

- H_0: Έχουν την ίδια κατανομή
- H_1: Δεν έχουν την ίδια κατανομή

Η στατιστική συνάρτηση με την οποία θα ελέγξουμε τις συγκεκριμένες υποθέσεις στηρίζεται στις εμπειρικές κατανομές των δύο δειγμάτων παρατηρήσεων. Για το τυχαίο δείγμα $X_1, X_2, \dots, X_n$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_x(x)$, ενώ για το τυχαίο δείγμα $Y_1, Y_2, \dots, Y_m$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_y(y)$. Επιπλέον, η ελεγχοσυνάρτηση εκφράζει ένα μέτρο της μέσης τετραγωνικής απόκλισής τους.

Επομένως, έχουμε την στατιστική συνάρτηση ελέγχου:

$$T_2 = \frac{mn}{(m+n)^2}\left\{\sum_{i=1}^{n}[S_X(Xi) - S_Y(Xi)]^2 + \sum_{j=1}^{n}[S_X(Yi) - S_Y(Yi)]^2\right\}$$

Παρατηρείται ότι μια ένδειξη εναντίον της μηδενικής υπόθεσης αποτελούν οι μεγάλες τιμές στην ελεγχοσυνάρτηση.

Κανόνας απόφασης: Η μηδενική απόφαση απορρίπτεται σε επίπεδο σημαντικότητας α, αν η τιμή της κατάλληλης στατιστικής συνάρτησης υπερβαίνει το $(1-\alpha)$-ποσοστιαίο σημείο της κατανομής της, όπως αυτό δίνεται στον πίνακα 3 του παραρτήματος. Δηλαδή, αν $T_2 > w_{1-a}$.

Ιστορική αναδρομή: Τα ποσοστιαία σημεία της κατανομής της ελεγχοσυνάρτησης ορίστηκαν το 1962 από τον Anderson. Η κατανομή της στατιστικής συνάρτησης T_2 είχε προσδιοριστεί το 1952 από τους Darling και Anderson, οι οποίοι αναφέρθηκαν στην ελεγχοσυνάρτηση που πρότειναν οι Cramer και Von Mises. Δηλαδή, παρόλο που οι Cramer και Von Mises δεν επινόησαν οι ίδιοι την στατιστική συνάρτηση T_2, αυτή πήρε το όνομά τους και ονομάστηκε ελεγχοσυνάρτηση των Cramer και Von Mises.

Παράδειγμα

Στο συγκεκριμένο παράδειγμα θα βασιστούμε στην εκφώνηση και στα στοιχεία του παραδείγματος 1 του στατιστικού ελέγχου Smirnov. Εφαρμόζοντας τον έλεγχο Cramer–Von Mises θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν υπάρχει διαφορά στις γνώσεις, όπως αυτές μετριούνται με αυτό το διαγώνισμα, των δύο πληθυσμών μαθητών της έκτης τάξης.

Ορίζεται η μηδενική υπόθεση: Οι δύο πληθυσμοί έχουν την ίδια συνάρτηση κατανομής και η εναλλακτική υπόθεση: Οι δύο πληθυσμοί δεν έχουν την ίδια συνάρτηση κατανομής.

Επομένως, σύμφωνα με αυτά που έχουν προηγηθεί, η μηδενική και η εναλλακτική υπόθεση μπορούν να γραφτούν με τον ακόλουθο τρόπο:

- H_0: $F_x(x) = F_y(x)$, για κάθε $x \in (-\infty, +\infty)$
- H_1: $F_x(x) \neq F_y(x)$, για τουλάχιστον ένα $x \in (-\infty, +\infty)$

Όπου $F_x(x)$ και $F_y(y)$ οι αντίστοιχες συναρτήσεις κατανομής των X και Y.

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν την ίδια κατανομή, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν την ίδια κατανομή.

Δηλαδή:

- H_0: Έχουν την ίδια κατανομή
- H_1: Δεν έχουν την ίδια κατανομή

Για το τυχαίο δείγμα $X_1, X_2, \dots, X_9$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_x(x)$, ενώ για το τυχαίο δείγμα $Y_1, Y_2, \dots, Y_{15}$ θέτουμε την εμπειρική συνάρτηση κατανομής $S_y(y)$.

Στον παρακάτω πίνακα βρίσκονται τα δύο δείγματα διατεταγμένα κατά αύξουσα σειρά μεγέθους, σαν να αποτελούν ένα ενιαίο δείγμα, αλλά και οι τιμές τις διαφοράς $S_x(x) - S_y(x)$ των εμπειρικών συναρτήσεων κατανομής.

X_i	Y_i	$S_x(x) - S_y(x)$	X_i	Y_i	$S_x(x) - S_y(x)$
	5.2	0-1/15= -1/15		9.8	5/9-8/15=1/45
	5.7	0-2/15= -2/15	9.9		6/9-8/15= 2/15
	5.9	0-3/15= -1/5	10.1		7/9-8/15= 11/45
	6.5	0-4/15= -4/15	10.6		8/9-8/15=16/45
	6.8	0-5/15= -1/3		10.8	8/9-9/15= 13/45
7.6		1/9-5/15= -2/9	11.2		1-9/15= 2/5

	8.2	1/9-6/15= -13/45		11.3	1-10/15= 1/3
8.4		2/9-6/15= -8/45		11.5	1-11/15=4/15
8.6		3/9-6/15= -1/15		12.3	1-12/15= 1/5
8.7		4/9-6/15= 2/45		12.5	1-13/15= 2/15
	9.1	4/9-7/15= -1/45		13.4	1-14/15= 1/15
9.3		5/9-7/15=4/45		14.6	1-1= 0

Πίνακας 5: Τα δύο δείγματα και η διαφορά $S_x(x) - S_y(x)$

Η κατάλληλη στατιστική συνάρτηση με την οποία θα ελέγξουμε τις συγκεκριμένες υποθέσεις στηρίζεται στις εμπειρικές κατανομές των δύο δειγμάτων παρατηρήσεων και είναι η:

$$T_2 = \frac{mn}{(m+n)^2}\left\{\sum_{i=1}^{n}[S_X(X_i) - S_Y(X_i)]^2 + \sum_{j=1}^{n}[S_X(Y_i) - S_Y(Y_i)]^2 \right\}$$

Υπολογίζουμε

$$\sum_{e'=1}^{9}[S_X(X_i) - S_Y(X_i)]^2 = 0.459$$

και

$$\sum_{j=1}^{15}[S_X(Y_j) - S_Y(Y_j)]^2 = 0.657$$

Επομένως, η στατιστική συνάρτηση T_2 γίνεται:

$$T_2 = \frac{(15)(9)}{(24)^2}(0.459 + 0.657)$$

$$= 0.262$$

Θεωρούμε επίπεδο σημαντικότητας $\alpha = 0.05$.

Κάνοντας χρήση του πίνακα 3, παρατηρούμε ότι το 0.95-ποσοστιαίο σημείο της κατανομής της στατιστικής συνάρτησης T_2 είναι $w_{0.95} = 0.461$.

Συγκρίνοντας την T_2 με το $w_{0.95}$ έχουμε: $T_2 = 0.262 < w_{0.95} = 0.461$.

Για να αποφασίσουμε αν θα απορρίψουμε ή όχι την μηδενική υπόθεση δεν πρέπει να ξεχνάμε τον κανόνα απόφασης.

Κανόνας απόφασης: Η μηδενική απόφαση απορρίπτεται σε επίπεδο σημαντικότητας α, αν η τιμή της κατάλληλης στατιστικής συνάρτησης υπερβαίνει το $(1-\alpha)$-ποσοστιαίο σημείο της κατανομής της, όπως αυτό δίνεται στον πίνακα 3 του παραρτήματος. Δηλαδή, αν $T_2 > w_{1-\alpha}$.

Επομένως, σύμφωνα με τα προηγούμενα, η μηδενική υπόθεση H_0 δεν απορρίπτεται σε επίπεδο σημαντικότητας 5%, δηλαδή δεν απορρίπτεται ότι οι δύο πληθυσμοί έχουν την ίδια συνάρτηση κατανομής.

Μέθοδος Anderson – Darling

Ο συγκεκριμένος έλεγχος προτάθηκε το 1952 από τους Theodore Anderson και Donald Darling και αποτελεί μια παραλλαγή του Kolmogorov – Smirnov. Παράλληλα, στηρίζεται στην εμπειρική αθροιστική συνάρτηση κατανομής. Μπορεί να έχει τον ίδιο "σκοπό" με τον έλεγχο του Smirnov αλλά αποτελεί έναν πιο ισχυρό έλεγχο. Χρησιμοποιείται όταν ο ερευνητής ενδέχεται να ενδιαφέρεται να προσδιορίσει αν οι δύο συναρτήσεις κατανομής που σχετίζονται με τους δύο πληθυσμούς ταυτίζονται ή όχι.

Έστω το τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ μεγέθους n από τον πληθυσμό 1 με κατανομή F και έστω ένα άλλο τυχαίο δείγμα $Y_1, Y_2, \ldots, Y_m$ μεγέθους m από τον πληθυσμό 2 με κατανομή G, τα οποία είναι ανεξάρτητα μεταξύ τους.

Οι υποθέσεις που θέλουμε να ελέγξουμε είναι οι εξής:

- $H_0: F = G$
- $H_1: F \neq G$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν την ίδια κατανομή, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν την ίδια κατανομή.

Δηλαδή:

- H_0: Έχουν την ίδια κατανομή
- H_1: Δεν έχουν την ίδια κατανομή

Το δείγμα που προκύπτει από την συνένωση των δυο δειγμάτων $Z = \{z_1, \ldots, z_\kappa\}$ έχει μέγεθος $k = n + m$ και οι παρατηρήσεις ταξινομούνται σε αύξουσα σειρά μεγέθους.

Η στατιστική συνάρτηση που θα χρησιμοποιηθεί είναι η:

$$AD = \frac{1}{mn} \sum_{i=1}^{k-1} \frac{(kc_i - m_i)^2}{i(k-i)}$$

όπου $1 \leq i \leq k$ και $c_i =$ πλήθος των στοιχείων στο $C_i = \{x: x \in X \; \kappa\alpha\iota \; x \leq z_i\}$.

Η μηδενική υπόθεση απορρίπτεται όταν $AD \geq ADcrit$, όπου το $ADcrit$ είναι η κρίσιμη τιμή που μπορεί να βρεθεί στον πίνακα 8 του παραρτήματος.

<u>Παράδειγμα</u>

Ένας ταλαντούχος σεφ της Νέας Υόρκης επιθυμεί να μαγειρέψει. Στην κουζίνα του μπορεί να βρει ποικιλίες από κάθε είδους φρούτα και λαχανικά. Συγκεκριμένα, έχει 20 φρούτα και 30 λαχανικά. Ο σεφ επιλέγει δύο ανεξάρτητα τυχαία δείγματα, ένα από κάθε κατηγορία, μεγέθους 8 και 7 αντίστοιχα και καταγράφει τον αριθμό των θερμίδων τους.

Φρούτα(X)	Λαχανικά(Y)
3	2
2	8
3	2
5	4
8	4
9	3
8	6
8	

Πίνακας 6: Αριθμός θερμίδων στα φρούτα και στα λαχανικά

Εφαρμόζοντας τον έλεγχο Smirnov θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν ο πληθυσμός των φρούτων διαφέρει από τον πληθυσμό των λαχανικών ως προς τον αριθμό θερμίδων, με βάση τα δύο ανεξάρτητα δείγματα που επέλεξε ο σεφ.

Οι υποθέσεις που θέλουμε να ελέγξουμε είναι οι εξής:

- $H_0 : F = G$
- $H_1 : F \neq G$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν την ίδια κατανομή, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν την ίδια κατανομή.

Δηλαδή:

- H_0: Έχουν την ίδια κατανομή
- H_1: Δεν έχουν την ίδια κατανομή

Το δείγμα που προκύπτει από την συνένωση των δυο δειγμάτων $Z = \{z_1, \ldots, z_\kappa\}$ έχει μέγεθος $k = n + m$ και οι παρατηρήσεις ταξινομούνται σε αύξουσα σειρά μεγέθους.

z	c	z	c
2	1	5	4
2	1	6	4
2	1	8	7
3	3	8	7
3	3	8	7
3	3	8	7
4	3	9	8
4	3		

Πίνακας 7: Συνενωμένο δείγμα και οι τιμές c

Η στατιστική συνάρτηση που θα χρησιμοποιηθεί είναι η:

$$AD = \frac{1}{mn}\sum_{i=1}^{k-1}\frac{(kc_i-m_i)^2}{i(k-i)} = 0.748687$$

όπου $1 \leq i \leq k$ και $c_i =$ (πλήθος εμφανίσεων του στοιχείου z_i στο $X_i +$ την προηγούμενη τιμή c_i). Η τιμή c_i παραμένει ίδια όταν δεν αλλάζει το z_i.

Η μηδενική υπόθεση απορρίπτεται όταν $AD \geq ADcrit,$ όπου το $ADcrit$ είναι η κρίσιμη τιμή που μπορεί να βρεθεί στον πίνακα 8 του παραρτήματος. Στην περίπτωση μας, $ADcrit$ = 2.381, σύμφωνα με τον πίνακα 8 του παραρτήματος για m = 7 και n = 8. Επομένως, σύμφωνα με όσα έχουν προηγηθεί δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση.

Μέθοδος Wald-Wolfowitz –Έλεγχος των ροών

Η συγκεκριμένη μη παραμετρική μέθοδος, που πήρε το όνομά της από τους στατιστικολόγους Abraham Wald και Jacob Wolfowitz, στηρίζεται στο πλήθος των ροών ενός δείγματος παρατηρήσεων. Πιο συγκεκριμένα, ροή θεωρείται μια διαδοχή πανόμοιων συμβόλων, της οποίας προηγούνται ή/και έπονται διαφορετικά σύμβολα. Για παράδειγμα, έχουμε μια σειρά-ακολουθία από άνδρες και γυναίκες, $A\ A\ A\ \Gamma\ \Gamma\ A\ \Gamma\ \Gamma\ \Gamma\ A\ A\ A\ \Gamma\ A$, η οποία εμφανίζει 7 ροές, 4 του συμβόλου A και 3 του συμβόλου Γ.

Αξίζει να σημειωθεί ότι ο συνολικός αριθμός των ροών σε μια ακολουθία συμβόλων δύο τύπων μπορεί να χρησιμοποιηθεί για να ελεγχθεί το αν δύο ανεξάρτητα τυχαία δείγματα προέρχονται από τον ίδιο πληθυσμό ή όχι.

Έστω δύο ανεξάρτητα τυχαία δείγματα $X_1, X_2, \dots, X_n$ και $Y_1, Y_2, \dots, Y_m$, όπου $n = n_1$ και $m = n_2$ ο αριθμός των παρατηρήσεων στις τυχαίες μεταβλητές X και Y αντίστοιχα.

Οι υποθέσεις που θα ελεγχθούν είναι οι εξής:

- H_0: $F_x(x) = F_y(x)$, για κάθε $x \in (-\infty, +\infty)$
- H_1: $F_x(x) \neq F_y(x)$, για τουλάχιστον ένα $x \in (-\infty, +\infty)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν την ίδια κατανομή, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν την ίδια κατανομή.

Δηλαδή:

- H_0: Έχουν την ίδια κατανομή
- H_1: Δεν έχουν την ίδια κατανομή

Για να εκτελέσουμε τον έλεγχο αρχικά διατάσσουμε τις $n_1 + n_2$ παρατηρήσεις σε αύξουσα σειρά και αντικαθιστούμε τις παρατηρήσεις που αντιστοιχούν στο πρώτο δείγμα με a και τις παρατηρήσεις που αντιστοιχούν στο δεύτερο δείγμα με b, ή με άλλα σύμβολα της επιλογής μας. Στη συνέχεια, είναι απαραίτητος ο υπολογισμός του πλήθος των ροών.

Η στατιστική συνάρτηση με την οποία θα ελέγξουμε τις συγκεκριμένες υποθέσεις είναι η:

$$T \; = \text{αριθμός των ροών}$$

Το εύρος των τιμής της παραμέτρου T είναι $2 \leq T \leq (n_1 + n_2)$. Στην ακραία περίπτωση που $T = 2$, οι δύο πληθυσμοί X_i και Y_i δεν προέρχονται από την ίδια κατανομή, πράγμα που σημαίνει ότι τα συγκεκριμένα χαρακτηριστικά που εξετάζονται είναι εντελώς ανόμοια. Καθώς το πλήθος των εναλλαγών αυξάνει, οι δύο πληθυσμοί αρχίζουν να "εισχωρούν" ο ένας στον άλλον, γεγονός που οδηγεί με μμεγαλύτερη βεβαιότητα ότι οι δύο υπό εξέταση πληθυσμοί προέρχονται από την ίδια κατανομή.

Ως εκ τούτου, η στατιστική παράμετρος T, η οποία είναι απλώς ένας θετικός ακέραιος αριθμός, μπορεί να χρησιμοποιηθεί ως μέτρο ομοιότητας, θεωρώντας ότι όσο πιο μεγάλη είναι η τιμή της, τόσο πιο όμοιες είναι οι κατανομές από τις οποίες υπολογίζεται. Με άλλα λόγια, τα δείγματα είναι "μπλεγμένα" μεταξύ τους. Παρατηρείται ότι μια ένδειξη εναντίον της μηδενικής υπόθεσης αποτελούν οι μικρές τιμές στην ελεγχοσυνάρτηση.

Για παράδειγμα, η πρώτη ομάδα εικόνων που παρουσιάζεται ακριβώς από κάτω έχει μεγάλο αριθμό T ενώ η δεύτερη ομάδα έχει μικρό T:

1η ομάδα:

2η ομάδα:

Σχήμα 4: 2 ομάδες εικόνων με διαφορετικό αριθμό Τ

Η κρίσιμη περιοχή μεγέθους α του ελέγχου των παραπάνω υποθέσεων ορίζεται από τις ανισότητες $T \leq w_{a/2}$ και $T \geq 1 - w_{a/2}$, όπου w_p συμβολίζει το p-ποσοστιαίο σημείο της κατανομής της στατιστικής συνάρτησης T. Ποσοστιαία σημεία της κατανομής αυτής για διάφορες τιμές των n_1 και n_2 δίνονται για την περίπτωση αμφίπλευρων ελέγχων μεγέθους $\alpha = 0.05$ στον πίνακα 4 του παραρτήματος. Στην περίπτωση μεγάλων δειγμάτων ($n_1 > 20$ ή $n_2 >$20), η κατανομή της στατιστικής συνάρτησης T προσεγγίζεται ικανοποιητικά από την κανονική κατανομή με μέση τιμή:

$$\frac{2n_1 n_2}{n_1 + n_2} + 1$$

και διασπορά:

$$\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}$$

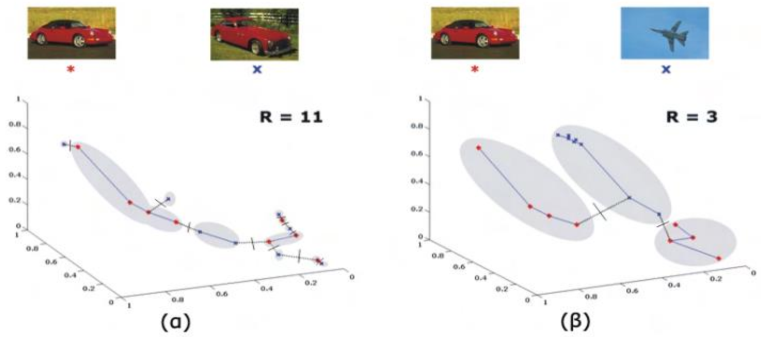
Επομένως, για τον έλεγχο υποθέσεων τυχαιότητας, στην περίπτωση αυτή, μπορεί να χρησιμοποιηθεί ως ελεγχοσυνάρτηση, η τυποποιημένη μορφή της στατιστικής συνάρτησης T, δηλαδή, η στατιστική συνάρτηση:

$$T_1 \;=\frac{T-\left(\frac{2n_1n_2}{n_1+n_2}+1\right)}{\sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}}$$

Η κρίσιμη περιοχή μεγέθους α του ελέγχου, θα ορίζεται, επομένως, από την ανισότητα $|T_1| \;\geq\; z_{1-a/2}$.

Παρατήρηση: Μερικές φορές στο πλαίσιο του συγκεκριμένου στατιστικού ελέγχου γίνεται ο υπολογισμός του ζειγνύοντος δέντρου ελαχίστου μήκους(MST). Ο ρόλος του MST στην συγκεκριμένη προτεινόμενη μεθοδολογία, έγκειται στον σχηματισμό μιας διατεταγμένης ακολουθίας των δειγμάτων και των δύο κατανομών στο χώρο των χαρακτηριστικών. Με το τρόπο αυτό γίνεται εύκολη η καταμέτρηση των εναλλαγών μεταξύ των κατανομών αυτών μέσα στο συνολικό δέντρο και ο υπολογισμός της στατιστικής παραμέτρου T. Για την εύρεση του δέντρου ελάχιστου μήκους-MST έχουν προταθεί διάφοροι αλγόριθμοι, οι πιο σημαντικοί από τους οποίους είναι ο αλγόριθμος Kruskal και ο αλγόριθμος Prim.

Για παράδειγμα, με βάση τις εικόνες που αναφέραμε προηγουμένως έχουμε:



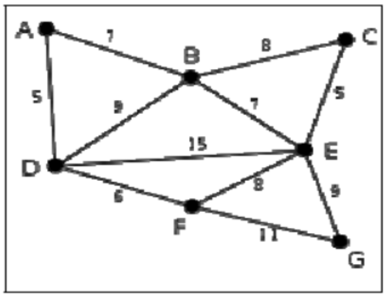
Σχήμα 5: 2 διαφορετικά MST

Αλγόριθμος Kruskal

Ο αλγόριθμός Kruskal είναι ο πιο παλαιός και απλός αλγόριθμος κατασκευής του MST και επιτυγχάνεται ξεκινώντας από κάποιο γράφο και επιλέγουμε διαδοχικά ακμές. Η υλοποίησή του είναι ιδιαίτερα κατάλληλη για αραιούς γράφους, εκείνους δηλαδή τους γράφους που έχουν μικρό αριθμό από ακμές. Τα βήματα που ακολουθεί ο αλγόριθμος για την κατασκευή του MST είναι τα ακόλουθα:

1. Ελέγχουμε ότι ο γράφος είναι ενωμένος. Αν δεν είναι, διακόπτουμε.
2. Ταξινομούμε τις ακμές του γράφου G κατά αύξουσα σειρά μήκους.
3. Επιλέγουμε τις ακμές με αυτήν τη σειρά, ελέγχοντας παράλληλα σε κάθε στάδιο να μη δημιουργείται κύκλος. Αν η πρόσφατα προστιθέμενη ακμή σχηματίζει κύκλο με ακμές που έχουν ήδη επιλεγεί, την απορρίπτουμε και δοκιμάζουμε την επόμενη.
4. Αν βρεθούν $N-1$ ακμές, τότε το υπάρχον σύνολο των ακμών αποτελεί το MST, συνεπώς σταματάμε.

Επομένως υπό την προϋπόθεση ότι έχουμε ένα συνδεδεμένο σταθμισμένο γράφο $G(V,E)$ και αντίστοιχα βάρη (Ευκλείδειες αποστάσεις) lij ως εισόδους, ο αλγόριθμος Kruskal επιλέγει πάντοτε τη μικρότερη διαθέσιμη ακμή από το σύνολο E των ακμών και την προσθέτει στο σύνολο των ήδη επιλεγμένων ακμών. Ας υποθέσουμε ότι έχουμε τον παρακάτω γράφο με τα βάρη των ακμών του.



Σχήμα 6: Αλγόριθμος Kruskal

Οι ακμές AD και CE έχουν το μικρότερο μήκος (5). Αρχικά αυθαίρετα επιλέγουμε την AD και μετά την CE. Η επόμενη ακμή που επιλέγεται είναι η DF μήκους (6) και αμέσως μετά οι AB και BE μήκους (7). Η ακμή BD απορρίπτεται διότι υπάρχει ήδη ένα μονοπάτι μεταξύ των B και D, αλλιώς θα σχηματιζόταν κύκλος (ABD). Για αποφυγή σχηματισμού του (BCE) κύκλου απορρίπτουμε την ακμή BC. Επίσης απορρίπτουμε τις ακμές DE και FE για να μην σχηματιστούν οι κύκλοι $(DEBA)$ και $(FEBAD)$ αντίστοιχα. Τέλος επιλέγοντας την EG καταλήγουμε στο MST.

Παράδειγμα

Κάθε χρόνο πραγματοποιείται ο πανελλήνιος διαγωνισμός μαθηματικών, στον οποίο μπορούν να συμμετέχουν μαθητές και μαθήτριες από όλα τα σχολεία της Ελλάδας. Σε ένα τυχαίο δείγμα πέντε μαθητών της πρώτης τάξης του Λυκείου ενός σχολείου που ανήκε σε μία περιοχή κάποιας πόλης, δόθηκε το συγκεκριμένο μαθηματικό διαγώνισμα. Σε ένα άλλο τυχαίο δείγμα έξι μαθητών της πρώτης τάξης του Λυκείου ενός άλλου σχολείου που ανήκε σε μία διαφορετική περιοχή της πόλης, δόθηκε το ίδιο διαγώνισμα. Τα αποτελέσματα βρίσκονται στον παρακάτω πίνακα:

	Αποτελέσματα
1^o δείγμα	1.5, 2.4, 1.2, 4.6, 0.8
2^o δείγμα	2.5, 3.4, 0.2, 1.6, 1.8, 5.1

Πίνακας 8: Αποτελέσματα μαθητών

Εφαρμόζοντας τον έλεγχο ροών θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν υπάρχει διαφορά στις γνώσεις, όπως αυτές μετριούνται με αυτό το διαγώνισμα, των δύο πληθυσμών μαθητών της πρώτης τάξης.

Ορίζεται η μηδενική υπόθεση: Οι δύο πληθυσμοί έχουν την ίδια συνάρτηση κατανομής και η εναλλακτική υπόθεση: Οι δύο πληθυσμοί δεν έχουν την ίδια συνάρτηση κατανομής.

Επομένως, σύμφωνα με αυτά που έχουν προηγηθεί, η μηδενική και η εναλλακτική υπόθεση μπορούν να γραφτούν με τον ακόλουθο τρόπο:

- H_0: $F_x(x) = F_y(x)$, για κάθε $x \in (-\infty, +\infty)$
- H_1: $F_x(x) \neq F_y(x)$, για τουλάχιστον ένα $x \in (-\infty, +\infty)$

Λαμβάνουμε το διατεταγμένο δείγμα:

0.2, 0.8, 1.2, 1.5, **1.6**, **1.8**, 2.4, **2.5**, **3.4**, 4.6, **5.1**,

όπου οι παρατηρήσεις του δεύτερου δείγματος επισημαίνονται με έντονη γραφή.

Αν στην παραπάνω ακολουθία συμβολίσουμε με X τις παρατηρήσεις από το 1ο δείγμα και με Y τις παρατηρήσεις από το 2ο δείγμα τότε λαμβάνουμε την ακολουθία συμβόλων: $\boldsymbol{Y}, X, X, X, \boldsymbol{Y}, \boldsymbol{Y}, X, \boldsymbol{Y}, \boldsymbol{Y}, X, \boldsymbol{Y}$

Ο αριθμός των ροών στην παραπάνω ακολουθία είναι

$$T = 7\text{: } (Y), (XXX), (YY), (X), (YY), (X), (Y).$$

Σύμφωνα με τον πίνακα του παραρτήματος, προκύπτει ότι η κρίσιμη περιοχή μεγέθους 0.05 του ελέγχου ορίζεται από τις ανισότητες $T \leq w_{0.025} = 3$ και $T \geq w_{0.025} = 10$. Η παρατηρούμενη τιμή της T είναι 7 και βρίσκεται εντός της κρίσιμης περιοχής και συνεπώς η υπόθεση ότι οι δύο πληθυσμοί προέρχονται από την ίδια κατανομή απορρίπτεται σε επίπεδο σημαντικότητας $\alpha = 5\%$.

ΚΕΦΑΛΑΙΟ 3 - Έλεγχοι για την ισότητα των παραμέτρων θέσης δύο ανεξάρτητων πληθυσμών

Ως μέτρα θέσης εννοούμε κυρίως τα μέτρα κεντρικής τάσης που προσδιορίζουν ένα κεντρικό σημείο γύρω από το οποίο τείνουν να συγκεντρώνονται τα δεδομένα. Ένα από τα κυριότερα μέτρα κεντρικής τάσης είναι η δειγματική μέση τιμή ή αριθμητικός μέσος ή αλλιώς μέσος όρος. Ο ορισμός που έχει επικρατήσει στην στατιστική για τη μέση τιμή είναι: Η δειγματική μέση τιμή είναι το πιο γνωστό και χρήσιμο μέτρο του κέντρου των δεδομένων. Έστω $X_1, X_2, \ldots, X_n$, οι τιμές των παρατηρήσεων του δείγματος για μία τυχαία μεταβλητή X που μελετάμε. Η δειγματική μέση τιμή συμβολίζεται $\bar{X}$ και ορίζεται ως το άθροισμα των παρατηρήσεων διά του πλήθους των παρατηρήσεων.

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Διαβάζεται "άθροισμα των x_i από $i = 1$ έως n". Από τον ορισμό της μέσης τιμής, είναι φανερό ότι αν οι τιμές $X_1, X_2, \ldots, X_n$ είναι όλες μεταξύ τους ίσες, θα είναι ίσες με τη μέση τιμή τους.

Η μέση τιμή είναι το "κέντρο ισορροπίας" των δεδομένων. Για να καταλάβουμε τη φυσική της σημασία μπορούμε να φανταστούμε μία σανίδα πάνω στην οποία σκορπίζουμε ένα αριθμό n ίδιων βαριδιών. Το σημείο στήριξης της σανίδας, ώστε να ισορροπεί σε οριζόντια θέση, είναι η μέση τιμή της θέσης των βαριδιών πάνω στη σανίδα.

Στην έρευνα, στην περίπτωση που έχουμε δύο ανεξάρτητα δείγματα παρατηρήσεων, πολλές φορές τίθεται ένα ερώτημα για τις μέσες τιμές τους. Συγκεκριμένα, ο ερευνητής ενδέχεται να ενδιαφέρεται να προσδιορίσει αν τα δύο μέσα που σχετίζονται με τους δύο πληθυσμούς ισούνται ή όχι. Για αυτό, έχουν αναπτυχθεί διάφορες στατιστικές μεθοδολογίες για τον έλεγχο αν δύο σύνολα δεδομένων μπορούν να θεωρηθούν ότι έχουν ίδια μέσα. Οι έλεγχοι που θα μελετηθούν παρακάτω είναι κατάλληλοι στις περιπτώσεις αυτές.

Μέθοδος Mann-Whitney

Ο έλεγχος των Mann-Whitney είναι επίσης γνωστός ως έλεγχος του Wilcoxon. O Wilcoxon το 1945 τον μελέτησε πρώτος για ισομεγέθη δείγματα. Η περίπτωση που τα δείγματα έχουν διαφορετικό μέγεθος μελετήθηκε από τους Mann και Whitney το 1947 και από τον Wilcoxon το 1949. Για αυτό ακριβώς ο έλεγχος αυτός εμφανίζεται στην βιβλιογραφία με διαφορετικά ονόματα. Μπορεί συχνά να αναφέρεται και ως έλεγχος Wilcoxon-MannWhitney. Ο συγκεκριμένος έλεγχος χρησιμοποιείται κυρίως στην περίπτωση δύο ανεξάρτητων δειγμάτων, αλλά και σε πολλές άλλες περιπτώσεις. Συνήθως ο ερευνητής έχει δύο ανεξάρτητα δείγματα τα οποία προέρχονται από δύο διαφορετικούς πληθυσμούς και επιθυμεί να χρησιμοποιήσει έναν στατιστικό έλεγχο για να ελέγξει αν τα μέσα των δύο πληθυσμών ταυτίζονται ή δεν ταυτίζονται. Επομένως, Ο Mann-Whitney ελέγχει αν τα 2 ανεξάρτητα δείγματα (ομάδες) διαφέρουν ως προς τις μέσες σειρές τους (mean ranks) στη μεταβλητή ή αν οι 2 πληθυσμοί από τους οποίους προέρχονται τα 2 δείγματα είναι ίδιοι στην κεντρική τους θέση και αποτελεί το μέγιστο μη παραμετρικό ισοδύναμο του ελέγχου t για ανεξάρτητα δείγματα.

Η μεθοδολογία αντιμετώπισης του προβλήματος των δύο δειγμάτων στηρίζεται στην συνένωση των δύο δειγμάτων σε ένα ενιαίο δείγμα, του οποίου οι τιμές διατάσσονται κατ’ αύξουσα σειρά μεγέθους. Στις τιμές του νέου δείγματος αντιστοιχίζονται βαθμοί από την μικρότερη στη μεγαλύτερη, ανεξάρτητα από τον πληθυσμό από τον οποίο προήλθαν. Επομένως, η στατιστική συνάρτηση-ελεγχοσυνάρτηση που θα μπορούσε να χρησιμοποιηθεί, θα μπορούσε να ορισθεί ως το άθροισμα των βαθμών που αντιστοιχούν σε εκείνες τις τιμές που προέρχονται από τον ένα από τους δύο πληθυσμούς.

Έστω το τυχαίο δείγμα $X_1, X_2, \dots, X_n$ μεγέθους n από τον πληθυσμό 1 και έστω ένα άλλο τυχαίο δείγμα $Y_1, Y_2, \dots, Y_m$ μεγέθους m από τον πληθυσμό 2. Το δείγμα που προκύπτει από την συνένωση των δυο δειγμάτων έχει μέγεθος $N = n + m$ και αντιστοιχίζουμε στις τιμές του τους βαθμούς 1 έως $n + m$.

Έστω, επίσης, ότι $R(X_i)$ και $R(Y_i)$ οι βαθμοί των μεταβλητών X_i και Y_i για κάθε i και j. Σε περίπτωση που αρκετές τιμές του δείγματος ταυτίζονται, αντιστοιχίζουμε σε κάθε μία από αυτές τον μέσο των βαθμών που θα είχαν αν δεν ταυτίζονταν.

Έστω $E(X)$ και $E(Y)$ οι μέσες τιμές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα. Οι υποθέσεις που θα ελεγχτούν για την ισότητα των μέσων τιμών των δύο πληθυσμών, είναι οι εξής:

- $H_0: E(X) = E(Y)$
- $H_1: E(X) \neq E(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες μέσες τιμές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες μέσες τιμές.

Δηλαδή:

- H_0: Ίσες μέσες τιμές
- H_1: Δεν έχουν ίσες μέσες τιμές

Η στατιστική συνάρτηση T για τον έλεγχο των συγκεκριμένων υποθέσεων ισούται με το άθροισμα των βαθμών που οι παρατηρήσεις $X_1, X_2, \dots, X_n$ έχουν στο τελικό δείγμα των $n + m$ παρατηρήσεων, δηλαδή

$$T = \sum_{i=1}^{n} R(X_i)$$

Αν υπάρχουν πολλές περιπτώσεις ταύτισης τιμών στο δείγμα, αντί της στατιστικής συνάρτησης T, χρησιμοποιείται η τυποποιημένη μορφή της:

$$T_1 = \frac{T - E(T)}{\sqrt{V(T)}}$$

ή ισοδύναμα

$$T_1 = \frac{T - \frac{n(N+1)}{2}}{\sqrt{\frac{nm}{N(N-1)} \sum_{i=1}^{N} R_i^{\,2} - \frac{nm(N+1)^2}{4(N-1)}}}$$

όπου το συγκεκριμένο άθροισμα αναφέρεται στο άθροισμα των τετραγώνων των N μέσων βαθμών που στην πραγματικότητα χρησιμοποιούνται στο συνενωμένο δείγμα.

Αν S είναι το άθροισμα n ακεραίων, οι οποίοι έχουν επιλεγεί τυχαία και χωρίς επανάθεση από τους πρώτους N ακεραίους, τότε η μέση τιμή και η διασπορά του S δίνονται από τις σχέσεις

$$E(S) = \frac{n(N+1)}{2} \qquad \text{και} \qquad V(S) = \frac{n(N+1)(N-n)}{12}$$

αντίστοιχα.

Τα ποσοστιαία σημεία της κατανομής της στατιστικής συνάρτησης T δίνονται στον πίνακα 5 του παραρτήματος. Παράλληλα, τα ποσοστιαία σημεία της κατανομής της στατιστικής συνάρτησης T_1 προσεγγίζονται από τα αντίστοιχα ποσοστιαία σημεία της τυποποιημένης κανονικής κατανομής, τα οποία υπάρχουν στον πίνακα 6 του παραρτήματος.

Είναι προφανές ότι οι πολύ μεγάλες ή οι πολύ μικρές τιμές της στατιστικής συνάρτησης είναι ένδειξη απόρριψης της μηδενικής υπόθεσης. Δηλαδή, η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας ίσο με α, αν η τιμή της στατιστικής συνάρτησης είναι μικρότερη από το $\alpha/2$-ποσοστιαίο σημείο της κατανομής της ή αν είναι μεγαλύτερη από το $(1-\alpha/2)$-ποσοστιαίο σημείο της κατανομής της. Επομένως, η μηδενική υπόθεση θα απορρίπτεται αν $T < w_{a/2}$ ή αν $T > w_{1-a/2}$. Αν χρησιμοποιείται η T_1, η μηδενική υπόθεση θα απορρίπτεται αν $T_1 < w_{a/2}$ ή αν $T_1 > w_{1-a/2}$.

Συμπερασματικά, ο έλεγχος Mann-Whitney χρησιμοποιείται ως έλεγχος ύπαρξης διαφοράς στην θέση (μέση τιμή ή διάμεσο) δύο πληθυσμών.

<u>Παράδειγμα</u>

Σε ένα λύκειο, ο καθηγητής των μαθηματικών έβαλε όλους τους μαθητές του, κορίτσια και αγόρια, να γράψουν ένα διαγώνισμα με άριστα το 20. Κράτησε τους βαθμούς των γραπτών από 12 κορίτσια και από 36 αγόρια και ήθελε να εξετάσει αν οι μέσες τιμές των 2 ομάδων ισούνται-ταυτίζονται. Τα αποτελέσματα αυτά λαμβάνουν θέση στον ακριβώς παρακάτω πίνακα:

ΚΟΡΙΤΣΙΑ		ΑΓΟΡΙΑ					
14.8	10.6	12.7	16.9	7.6	2.4	6.2	9.9
7.3	12.5	14.2	7.9	11.3	6.4	6.1	10.6
5.6	12.9	12.6	16.0	8.3	9.1	15.3	14.8
6.3	16.1	2.1	10.6	6.7	6.7	10.6	5.0
9.0	11.4	17.7	5.6	3.6	18.6	1.8	2.6
4.2	2.7	11.8	5.6	1.0	3.2	5.9	4.0

Πίνακας 9: Οι βαθμοί των αγοριών και των κοριτσιών

Εφαρμόζοντας τον έλεγχο Mann-Whitney θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν υπάρχει ισότητα των μέσων τιμών των δύο πληθυσμών, με βάση τα δύο ανεξάρτητα δείγματα που επέλεξε ο καθηγητής.

Οι υποθέσεις που θα ελεγχθούν είναι οι εξής:

- $H_0: E(X) = E(Y)$
- $H_1: E(X) \neq E(Y)$

Δηλαδή:

- H_0: Τα αγόρια και τα κορίτσια έχουν ίσες μέσες τιμές βαθμών στα μαθηματικά
- H_1: Τα αγόρια και τα κορίτσια δεν έχουν ίσες μέσες τιμές βαθμών στα μαθηματικά

Συνδυάζουμε τα δύο δείγματα σε ένα ενιαίο δείγμα, τα διατάσσουμε κατά αύξουσα σειρά και βρίσκουμε και τον βαθμό. Έτσι προκύπτει ο παρακάτω πίνακας:

X	Y	Βαθμός	X	Y	Βαθμός	X	Y	Βαθμός
	1.0	1		6.2	17		11.3	33

	1.8	2	6.3		18	11.4		34
	2.1	3		6.4	19		11.8	35
	2.4	4		6.7	20.5	12.5		36
	2.6	5		6.7	20.5		12.6	37
2.7		6	7.3		22		12.7	38
	3.2	7		7.6	23	12.9		39
	3.6	8		7.9	24		14.2	40
	4.0	9		8.3	25		14.8	41.5
4.2		10	9.0		26	14.8		41.5
	5.0	11		9.1	27		15.3	43
	5.6	13		9.9	28		16.0	43
	5.6	13		10.6	30.5	16.1		45
5.6		13		10.6	30.5		16.9	46
	5.9	15	10.6		30.5		17.7	47
	6.1	16		10.6	30.5		18.6	48

Πίνακας 10: Τα δύο δείγματα και οι βαθμοί των παρατηρήσεων

Στο δείγμα υπάρχουν συνολικά 4 περιπτώσεις ταύτισης τιμών, επομένως η ελεγχοσυνάρτηση που θα χρησιμοποιηθεί είναι η T_1.

Υπολογίζουμε n =12 και $m = 36$. Άρα, $N = n + m = 48$.

Επίσης,

$$T = \sum_{i=1}^{n} R(X_i)$$

$$= 6+10+13+18+22+26+30.5+34+36+39+41.5+45$$

$$= 321$$

και $\sum_{i=1}^{n} {R_i}^2 = 38016$

Επομένως, $$T_1 = \frac{T - \frac{n(N+1)}{2}}{\sqrt{\frac{nm}{N(N-1)}\sum_{i=1}^{N} {R_i}^2 - \frac{nm(N+1)^2}{4(N-1)}}}$$

$$= \frac{321 - 13(\frac{49}{2})}{\sqrt{\frac{(12)(36)}{(48)(47)}(38016) - \frac{(12)(36)(49)^2}{4(47)}}}$$

$$= 0.6431$$

Σύμφωνα με τον πίνακα 2 του παραρτήματος, η κρίσιμη περιοχή μεγέθους $\alpha = 0.05$ εμπεριέχει τιμές της στατιστικής συνάρτησης T_1 μεγαλύτερες από $z_{0.95}$ =1.645. Επομένως, η T_1 δεν βρίσκεται μέσα στη κρίσιμη περιοχή και για αυτό η μηδενική υπόθεση δεν απορρίπτεται σε επίπεδο σημαντικότητας 0.05. Δηλαδή δεν απορρίπτεται η υπόθεση ότι οι δύο πληθυσμοί έχουν ίσες μέσες τιμές.

Λόγω του ότι οι περιπτώσεις ταύτισης τιμών ήταν μόνο 4, θα μπορούσαμε να χρησιμοποιήσουμε και την στατιστική συνάρτηση T. Σε αυτή την περίπτωση, η κρίσιμη περιοχή αντιστοιχεί στις τιμές της στατιστικής συνάρτησης T που υπερβαίνουν την τιμή:

$$w_{0.95} = n(N+1)/2 + z_{0.95}\sqrt{nm(N+1)/12}$$

$$= 294 + (1.645)\ (42)$$

$$= 363.1$$

Η T όμως έχει τιμή 321, επομένως η μηδενική υπόθεση δεν απορρίπτεται σε επίπεδο σημαντικότητας 0.05. Είναι λογικό να προκύπτει το ίδιο αποτέλεσμα με την περίπτωση που χρησιμοποιήσαμε την στατιστική συνάρτηση T_1.

Μέθοδος προσήμων

Ο προσημικός έλεγχος ή έλεγχος προσήμων (sign test) χρησιμοποιήθηκε για πρώτη φορά το 1710 και είναι ο αρχαιότερος μη παραμετρικός έλεγχος. Στην πραγματικότητα, αποτελεί μια παραλλαγή του διωνυμικού ελέγχου, η οποία απαιτεί ξεχωριστή μελέτη λόγω του ενδιαφέροντος που παρουσιάζει η ευέλικτη μορφή της που οφείλεται στο γεγονός ότι p_0 =1/2. Χρησιμοποιείται κυρίως για τον έλεγχο της υπόθεσης ότι οι τιμές μιας από τις τυχαίες μεταβλητές του ζεύγους (X, Y) τείνουν να είναι μεγαλύτερες από τις τιμές της άλλης. Εμείς όμως θα τον χρησιμοποιήσουμε για να αποδείξουμε την ισότητα των μέσων δύο ανεξάρτητων πληθυσμών. Σε πολλές από τις περιπτώσεις όπου μπορεί να χρησιμοποιηθεί ο προσημικός έλεγχος, υπάρχουν διαθέσιμοι περισσότερο ισχυροί μη παραμετρικοί έλεγχοι για το ίδιο μοντέλο. Στην περίπτωση μας, ένας πιο ισχυρός έλεγχος είναι ο Mann Whitney. Παρ' όλα αυτά, ο προσημικός έλεγχος προτιμάται λόγω της απλότητάς του.

Έστω το τυχαίο δείγμα $X_1, X_2, \dots, X_n$ μεγέθους n από τον πληθυσμό 1 και έστω ένα άλλο τυχαίο δείγμα $Y_1, Y_2, \dots, Y_m$ μεγέθους m από τον πληθυσμό 2, τα οποία είναι ανεξάρτητα μεταξύ τους. Σε κάθε ζεύγος (X_i , Y_i) γίνεται μια εσωτερική σύγκριση και το ζεύγος αυτό ταξινομείται ως "$+$" ή "συν" ζεύγος, εάν $X_i < Y_i$, ως "$-$" ή "πλην" ζεύγος, αν $X_i > Y_i$, ή ως "0" ή "ισοβαθμούν" ζεύγος, αν $X_i = Y_i$, $i = 1, 2, \dots, n$. Επομένως, η κλίμακα μέτρησης των δεδομένων αρκεί να είναι κλίμακα διάταξης.

Αν οι τιμές της μεταβλητής X είναι μεγαλύτερες από τις τιμές της μεταβλητής Y, τότε τα ζεύγη (X_i , Y_i) είναι εσωτερικά συνεπή, με την έννοια ότι αν $P(+) > P(-)$ για κάποιο ζεύγος (X_i , Y_i), τότε $P(+) > P(-)$, για όλα τα ζεύγη. Η εσωτερική συνέπεια ορίζεται με ανάλογο τρόπο και στην περίπτωση κατά την οποία οι τιμές της μεταβλητής Y τείνουν να είναι μεγαλύτερες από τις τιμές της μεταβλητής X (οπότε οι αντίστοιχες πιθανότητες των "$+$" και "$-$" ζευγών ικανοποιούν την ανισότητα $P(+) < P(-)$), όπως επίσης και στην περίπτωση που δεν υπάρχει τάση στις τιμές της μιας μεταβλητής να υπερβαίνουν τις τιμές της άλλης (οπότε οι πιθανότητες των "$+$" και "$-$" ζευγών είναι ίσες, δηλαδή, $P(+) = P(-)$).

Έστω $E(X)$ και $E(Y)$ οι μέσες τιμές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα. Οι υποθέσεις που θα ελεγχτούν για την ισότητα των μέσων τιμών των δύο πληθυσμών, είναι οι εξής:

- H_0: $E(X) = E(Y)$
- H_1: $E(X) \neq E(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες μέσες τιμές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες μέσες τιμές.

Δηλαδή:

- H_0: Ίσες μέσες τιμές
- H_1: Δεν έχουν ίσες μέσες τιμές

Είναι προφανές ότι ο προσημικός έλεγχος μπορεί να χρησιμοποιηθεί ως έλεγχος υποθέσεων και για τις διαμέσους των μεταβλητών X και Y.

Η στατιστική συνάρτηση T για τον έλεγχο των συγκεκριμένων υποθέσεων ισούται με τον αριθμό των "+" ζευγών, δηλαδή:

$$T = \text{αριθμός των "+" ζευγών.}$$

Με άλλα λόγια, η ελεγχοσυνάρτηση συμβολίζει τον αριθμό των ζευγών (X_i, Y_i) στα οποία η μεταβλητή X_i είναι μικρότερη από τη μεταβλητή Y_i. Αγνοώντας τα "0" ενδεχόμενα, και περιοριζόμενοι σε n ζεύγη,

$$\text{όπου } n = (\text{αριθμός των "+" ζευγών}) + (\text{αριθμός των "−" ζευγών}),$$

μπορούμε εύκολα να παρατηρήσουμε ότι ο έλεγχος των παραπάνω υποθέσεων βρίσκεται στο πλαίσιο του διωνυμικού μοντέλου.

Κάθε ζεύγος (X_i, Y_i) από τα απομένοντα n ζεύγη παρατηρήσεων θεωρείται ως μια δοκιμή, η οποία μπορεί να οδηγήσει σε ένα από δύο δυνατά αποτελέσματα:

$$\text{επιτυχία, "+" } (\{X_i < Y_i\}) \text{ με πιθανότητα } P(+)$$

$$\text{αποτυχία, "−" } (\{X_i > Y_i\}) \text{ με πιθανότητα } P(-)$$

Κατά συνέπεια, η ελεγχοσυνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους n και $p = P(+)$.

$$T \sim διωνυμική\ (n, p = P(+)).$$

Για τον έλεγχο των παραπάνω υποθέσεων μπορεί να χρησιμοποιηθεί ο διωνυμικός έλεγχος, αφού η στατιστική συνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους n και $p = P(+) = 1/2$.

$$T \sim διωνυμική\ (n, p = P(+) = 1/2).$$

Επομένως, σε επίπεδο σημαντικότητας περίπου ίσο με α, η H_0 απορρίπτεται αν $T \leq t_1$ ή $T > t_2$, όπου t_1, t_2 ορίζονται έτσι ώστε $P(T \leq t_1 \mid n, p = 1/2) \cong \alpha/2$ και $P(T > t_2 \mid n, p = 1/2) \cong \alpha/2$.

Η κρίσιμη περιοχή έχει την μορφή $T \leq t$ και $T \geq n - t$, όπου η τιμή t προσδιορίζεται από τον πίνακα 7 του παραρτήματος, έτσι ώστε $P(T \leq t \mid n, p = 1/2) \cong \alpha/2$.

<u>Παράδειγμα</u>

Μια ομάδα διαιτολόγων–διατροφολόγων ανέθεσαν σε έξι γυναίκες να παρακολουθήσουν μια συγκεκριμένη διατροφή και άσκηση για ένα ολόκληρο μήνα. Τα αποτελέσματα, όσον αφορά τα κιλά τους πριν(X_i) και μετά(Y_i), βρίσκονται στον παρακάτω πίνακα:

X_i	174	191	188	182	201	188
Y_i	165	186	183	178	203	181

Πίνακας 11: Τα κιλά των γυναικών πριν και μετά

Θέλουμε να ελέγξουμε αν οι μέσες τιμές των δύο πληθυσμών ταυτίζονται. Εφαρμόζοντας τον προσημικό έλεγχο, που διακρίνεται από το πόσο απλός είναι, θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν υπάρχει ισότητα των μέσων των δύο πληθυσμών, με βάση τα δύο ανεξάρτητα δείγματα που επιλέχθηκαν.

Έστω $E(X)$ και $E(Y)$ οι μέσες τιμές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα. Οι υποθέσεις που θα ελεγχτούν για την ισότητα των μέσων τιμών των δύο πληθυσμών, είναι οι εξής:

- $H_0: E(X) = E(Y)$
- $H_1: E(X) \neq E(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες μέσες τιμές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες μέσες τιμές.

Δηλαδή:

- H_0: Ίσες μέσες τιμές
- H_1: Δεν έχουν ίσες μέσες τιμές

Η στατιστική συνάρτηση T για τον έλεγχο των συγκεκριμένων υποθέσεων ισούται με τον αριθμό των "+" ζευγών, δηλαδή:

$$T = \text{αριθμός των "+" ζευγών.}$$

Επιπλέον, ακολουθεί την διωνυμική κατανομή με παραμέτρους $n = 6$ και $p = 1/2$, δηλαδή:

$$T \sim \text{διωνυμική } (6, 1/2).$$

Παρακάτω βρίσκουμε το πρόσημο του κάθε ζεύγους (X_i, Y_i):

X_i	174	191	188	182	201	188
Y_i	165	186	183	178	203	181
Πρόσημο	−	−	−	−	+	−

Πίνακας 12: Τα κιλά των γυναικών, πριν και μετά, μαζί με το πρόσημο

Επομένως, η τιμή της στατιστικής συνάρτησης ή αλλιώς το πλήθος των "+" ζευγών είναι ίσο με 1, $\tau = 1$.

Σύμφωνα με τον πίνακα 7 του παραρτήματος, προκύπτει ότι το κρίσιμο επίπεδο του ελέγχου είναι 10.94%.

Δηλαδή: $P(T \leq 1 \mid n = 6, p = 1/2) = 0.1094$

Επομένως, η μηδενική υπόθεση θα απορρίπτεται σε κάθε επίπεδο σημαντικότητας α, αν $\alpha \geq 0.1094$.

Πιο συγκεκριμένα, λόγω του ότι $\alpha = 0.05 < 0.1094$, η μηδενική υπόθεση δεν απορρίπτεται στα συνήθη επίπεδα σημαντικότητας.

Μέθοδος Προσημικής διάταξης του Wilcoxon

Ο έλεγχος των προσημικών βαθμών μεγέθους του Wilcoxon σε προβλήματα σύγκρισης των μέσων δύο πληθυσμών χρησιμοποιείται όταν η μελέτη βασίζεται σε ένα δείγμα ζευγών παρατηρήσεων. Πιο λεπτομερώς, τα δεδομένα αποτελούνται από n΄ ζεύγη τιμών $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, που αποτελούν πραγματοποίηση της ακολουθίας των διδιάστατων τυχαίων μεταβλητών $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, η οποία αποτελεί το τυχαίο δείγμα μεγέθους n' από τον διδιάστατο πληθυσμό που περιγράφεται από την μεταβλητή (X, Y).

Τα βήματα που θα ακολουθήσουμε είναι τα εξής:

1. Θεωρούμε τις διαφορές $D_i = Y_i - X_i$, $i = 1, 2, \ldots, n'$
2. Θεωρούμε τις απόλυτες διαφορές $|D_i| = |Y_i - X_i|$, $i = 1, 2, \ldots, n'$
3. Παραλείπουμε (αγνοούμε) τα ζεύγη (X_i, Y_i) για τα οποία $X_i = Y_i$, δηλαδή $D_i = 0$, και διατάσσουμε κατά αύξουσα σειρά μεγέθους τις n $(n \leq n')$ διαφορές $|D_i|$ που απομένουν
4. Αντιστοιχίζουμε στα n ζεύγη παρατηρήσεων που απομένουν βαθμούς από το 1 μέχρι το n, οι οποίοι δεν είναι τίποτε άλλο από τους βαθμούς των απολύτων διαφορών $|D_i| = |Y_i - X_i|$, $i = 1, 2, \ldots, n$. Έτσι, ο βαθμός 1 αντιστοιχεί στο ζεύγος (X_i, Y_i) με την μικρότερη απόλυτη διαφορά $|D_i|$, ο βαθμός 2 αντιστοιχεί στο ζεύγος (X_i, Y_i) με την αμέσως μεγαλύτερη απόλυτη διαφορά, ... , ο βαθμός n αντιστοιχεί στο ζεύγος (X_i, Y_i) με την μεγαλύτερη απόλυτη διαφορά
5. Αν αρκετά ζεύγη έχουν απόλυτες διαφορές οι οποίες ταυτίζονται, τότε σε κάθε ένα από αυτά τα ζεύγη παρατηρήσεων αντιστοιχίζεται ως βαθμός ο μέσος των βαθμών που τα ζεύγη θα είχαν αν οι διαφορές τους δεν ήταν ίσες.
6. Υποθέτουμε ότι η κατανομή της μεταβλητής $D = Y - X$ είναι συμμετρική και ότι η κλίμακα μέτρησης των δεδομένων είναι τουλάχιστον κλίμακα διαστήματος.
7. Συμβολίζουμε με $d_{0.5}$ το μέσο του πληθυσμού των διαφορών, ο οποίος περιγράφεται από την μεταβλητή $D = Y - X$. Προφανώς, η τιμή του μέσου $d_{0.5}$ παρέχει πληροφορίες για τις σχετικές θέσεις των δύο πληθυσμών.
   - Όταν οι τιμές της τυχαίας μεταβλητής Y είναι μεγαλύτερες από τις αντίστοιχες τιμές της τυχαίας μεταβλητής X και αυτό συμβαίνει τις

περισσότερες φορές στον πληθυσμό των διαφορών $D = Y - X$, τότε η διαφορά των μέσων είναι θετική, δηλαδή $d_{0.5} > 0$.

- Με όμοιο τρόπο, όταν οι τιμές της Y είναι μικρότερες από τις αντίστοιχες τιμές της X και αυτό συμβαίνει τις περισσότερες φορές στον πληθυσμό των διαφορών $D = Y - X$, τότε η διαφορά των μέσων είναι αρνητική, δηλαδή $d_{0.5} < 0$.
- Τέλος, αν οι τιμές της μεταβλητής $D = Y - X$ στον πληθυσμό των διαφορών είναι θετικές και αρνητικές με ίσες συχνότητες, τότε η διαφορά των μέσων είναι ίση με το μηδέν, δηλαδή $d_{0.5} = 0$. Είναι, δηλαδή, η $d_{0.5}$ μία παράμετρος η οποία μετρά την διαφορά στις θέσεις των δύο πληθυσμών.

Έστω $E(X)$ και $E(Y)$ οι μέσες τιμές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα. Ορίζουμε $d_{0.5} = E(X) - E(Y)$

Οι υποθέσεις που θα ελεγχτούν για την ισότητα των μέσων τιμών των δύο πληθυσμών, είναι οι εξής:

- H_0: $d_{0.5} = 0$, ή αλλιώς $E(X) = E(Y)$
- H_1: $d_{0.5} \neq 0$, ή αλλιώς $E(X) \neq E(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες μέσες τιμές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες μέσες τιμές.

Δηλαδή:

- H_0: Ίσες μέσες τιμές
- H_1: Δεν έχουν ίσες μέσες τιμές

Η στατιστική συνάρτηση που θα χρησιμοποιήσουμε βασίζεται και πάλι στους προσημικούς βαθμούς μεγέθους $R_i, i = 1, 2, \dots, n$ των απολύτων διαφορών $|D_i|, i = 1, 2, \dots, n$, οι οποίες ορίζονται όπως φαίνεται παρακάτω.

Δηλαδή,

$$R_i = \begin{cases} +R(|D_i|), \; αν \; D_i = Y_i - X_i > 0 \\ -R(|D_i|), \; αν \; D_i = Y_i - X_i < 0 \end{cases}, \quad i = 1,2, \dots, n.$$

Επομένως, έχουμε την:

$$T = \frac{\sum_{i=1}^{n} R_i}{\sum_{i=1}^{n} R_i{}^2},$$

Όταν όλες οι διαφορές $D_i = Y_i - X_i$ είναι διακεκριμένες, έχουμε την:

$$T = \frac{\sum_{i=1}^{n} R_i}{\sqrt{n(n+1)(2n+1)/6}}.$$

Η κρίσιμη περιοχή καθορίζεται με τη βοήθεια του πίνακα 6. Συγκεκριμένα, η κρίσιμη περιοχή μεγέθους α των του ελέγχου ορίζεται από τις ανισότητες:

$$T < z_{a/2} \text{ και } T > z_{1-a/2}.$$

Παράδειγμα

Εκτελέστηκε ένα πείραμα διάρκειας ενός μήνα σχετικά με το κάπνισμα. Συγκεκριμένα, 40 εθελοντές χωρίστηκαν σε 20 ζεύγη. Στον έναν από τους δύο στο κάθε ζεύγος, και μόνο σε αυτόν, γινόταν μια παρουσίαση σχετικά με το πόσο βλαβερό είναι το κάπνισμα, τις συνέπειες του αλλά και μαρτυρίες θυμάτων που έχουν καταστραφεί από αυτό. Στην συνέχεια, ανατέθηκε και στους δύο από κάθε ζεύγος να μετρήσουν για ένα ολόκληρο μήνα τον αριθμό των τσιγάρων που κάπνιζαν. Τα αποτελέσματα συνοψίζονται στον παρακάτω πίνακα:

Ζεύγος(i)	Περίπτωση $1(X_i)$	Περίπτωση $2(Y_i)$	Ζεύγος(i)	Περίπτωση $1(X_i)$	Περίπτωση $2(Y_i)$
1	85	90	11	70	50
2	50	60	12	85	90

3	85	113	13	70	93
4	70	100	14	77	115
5	30	45	15	70	62
6	75	99	16	50	85
7	67	86	17	55	72
8	60	85	18	75	125
9	60	85	19	81	103
10	85	87	20	20	60

Πίνακας 13: Αριθμός τσιγάρων σε κάθε ζεύγος

Εφαρμόζοντας τον έλεγχο των προσημικών βαθμών μεγέθους του Wilcoxon θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν είναι δυνατόν η διαφορά των μέσων να είναι 0.

Έστω $E(X)$ και $E(Y)$ οι μέσες τιμές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα. Ορίζουμε $d_{0.5} = E(X) - E(Y)$

Οι υποθέσεις που θα ελεγχτούν για την ισότητα των μέσων τιμών των δύο πληθυσμών, είναι οι εξής:

- H_0: $d_{0.5} = 0$, ή αλλιώς $E(X) = E(Y)$
- H_1: $d_{0.5} \neq 0$, ή αλλιώς $E(X) \neq E(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες μέσες τιμές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες μέσες τιμές.

Δηλαδή:

- H_0: Ίσες μέσες τιμές
- H_1: Δεν έχουν ίσες μέσες τιμές

Πραγματοποιώντας τα βήματα του συγκεκριμένου ελέγχου που προαναφέρθηκαν προκύπτει ο παρακάτω πίνακας:

Ζεύγος(i)	Περίπτωση $1(X_i)$	Περίπτωση $2(Y_i)$	$D_i = Y_i - X_i$	$\lvert D_i \rvert$	$R(\lvert D_i \rvert)$	R_i
1	85	90	+5	5	2.5	+2.5
2	50	60	+10	10	5	+5
3	85	113	+28	28	15	+15
4	70	100	+30	30	16	+16
5	30	45	+15	15	6	+6
6	75	99	+24	24	12	+12
7	67	86	+19	19	8	+8
8	60	85	+25	25	13.5	+13.5
9	60	85	+25	25	13.5	+13.5
10	85	87	+2	2	1	+1
11	70	50	-20	20	9	-9
12	85	90	+5	5	2.5	+2.5
13	70	93	+23	23	11	+11
14	77	115	+38	38	18	+18
15	70	62	-8	8	4	-4
16	50	85	+35	35	17	+17
17	55	72	+17	17	7	+7
18	75	125	+50	50	20	+20
19	81	103	+22	22	10	+10
20	20	60	+40	40	19	+19

Πίνακας 14: Υλοποίηση βημάτων του ελέγχου

Οι τιμές των R_i δεν είναι διακεκριμένες, επομένως η στατιστική συνάρτηση είναι η:

$$T = \frac{\sum_{i=1}^{20} R_i}{\sqrt{\sum_{i=1}^{20} {R_i}^2}} = \frac{184}{2869} = 3.435.$$

Η κρίσιμη περιοχή μεγέθους 0.05 ορίζεται από τις ανισότητες $T < z_{0.025}$ και $T > z_{0.975}$, όπου $z_{0.025} = -196$ και $z_{0.975} = 1.96$. Αφού η $T = 3.435$, βρίσκεται μέσα στην κρίσιμη περιοχή, με αποτέλεσμα η μηδενική υπόθεση να απορρίπτεται σε επίπεδο σημαντικότητας 0.05.

ΚΕΦΑΛΑΙΟ 4 - Έλεγχοι για την ισότητα των παραμέτρων κλίμακας δύο ανεξάρτητων πληθυσμών

Στη θεωρία πιθανοτήτων και τη στατιστική, η διακύμανση είναι η αναμενόμενη τιμή της τετραγωνικής απόκλισης της τυχαίας μεταβλητής από τη μέση τιμή, και άτυπα μετρά πόσο μακριά ένα σύνολο (τυχαίων) αριθμών απλώνεται από τη μέση τιμή του. Η μικρή διακύμανση δείχνει ότι η τυχαία μεταβλητή κατανέμεται κοντά στη μέση τιμή, ενώ η μεγάλη διακύμανση δείχνει ότι η τυχαία μεταβλητή κατανέμεται πολύ μακριά από τη μέση τιμή. Η διακύμανση έχει κεντρικό ρόλο στη στατιστική. Η διακύμανση είναι το τετράγωνο της τυπικής απόκλισης και συχνά συμβολίζεται σ^2 ή $Var(X)$. Υπολογίζεται:

$$Var(X) = E[(X - \mu)^2].$$

Παράλληλα, υπάρχει στην βιβλιογραφία ως μέσος αριθμητικός των τετραγώνων των διαφορών των τιμών μιας μεταβλητής από το μέσο αριθμητικό της. Υπολογίζεται:

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N} = \frac{\sum x_i^2}{N} - \mu^2$$

Η τυπική απόκλιση είναι απλά η τετραγωνική ρίζα της διακύμανσης.. Η διακύμανση και η τυπική απόκλιση είναι πολύ σημαντικές όταν συγκρίνουμε σύνολα δεδομένων ή όταν συγκρίνουμε ένα σύνολο δεδομένων με μία θεωρητική τιμή.

Στην έρευνα, στην περίπτωση που έχουμε δύο ανεξάρτητα δείγματα παρατηρήσεων, πολλές φορές τίθεται ένα ερώτημα για τις διασπορές τους. Συγκεκριμένα, ο ερευνητής ενδέχεται να ενδιαφέρεται να προσδιορίσει αν υπάρχει διαφορά στην μορφή των πληθυσμών μέσω της σύγκρισης των διασπορών τους. Για αυτό, έχουν αναπτυχθεί διάφορες στατιστικές μεθοδολογίες για τον έλεγχο αν δύο σύνολα δεδομένων μπορούν να θεωρηθούν ότι έχουν ίδιες διασπορές. Οι έλεγχοι που θα μελετηθούν παρακάτω είναι κατάλληλοι στις περιπτώσεις αυτές.

Μέθοδος Siegel - Tukey

Ο συγκεκριμένος έλεγχος, που πήρε το όνομά του από τους Sidney Siegel και John Tukey, μπορεί να είναι απλός στον τρόπο που διεξάγεται, αλλά δεν είναι ιδιαίτερα ισχυρός. Η βασική ιδέα πίσω από την τεχνική που χρησιμοποιείται είναι ότι, αν δύο δείγματα προέρχονται από πληθυσμούς που διαφέρουν μόνο κατά την διασπορά τους, το δείγμα που προέρχεται από τον πληθυσμό με την μεγαλύτερη διασπορά θα είναι περισσότερο απλωμένο και με μεγαλύτερες ακραίες τιμές.

Εάν διατάξουμε κατά αύξουσα σειρά μεγέθους τις παρατηρήσεις του δείγματος που προκύπτει αν συνενώσουμε τα δύο επιμέρους δείγματα και αντιστοιχίσουμε τον βαθμό 1 στην μικρότερη παρατήρηση, 2 στην μεγαλύτερη, 3 στην επόμενη μεγαλύτερη, 4 και 5 στις επόμενες δύο πιο μικρές, 6 και 7 στις επόμενες πιο μεγάλες και ούτω καθεξής, το άθροισμα των βαθμών που αντιστοιχούν στις παρατηρήσεις του δείγματος που προέρχεται από τον πληθυσμό με την μεγαλύτερη διασπορά αναμένεται να είναι μικρότερο από ό,τι εάν δεν διαφέρουν οι διασπορές των πληθυσμών.

Όμως, αν οι δύο πληθυσμοί διαφέρουν και ως προς την θέση, δηλαδή εάν υπάρχουν και διαφορές στις τιμές των μέσων τους ή άλλων μέτρων θέσης, η εφαρμογή του ελέγχου δεν θα είναι το ίδιο ικανοποιητική. Ένας τρόπος αντιμετώπισης αυτού του "προβλήματος", όταν υπάρχουν ενδείξεις ότι οι πληθυσμοί διαφέρουν και ως προς την θέση, είναι να "μετασχηματίσουμε" τις παρατηρήσεις του δείγματος από τον πληθυσμό με την μεγαλύτερη μέση τιμή ή διάμεσο αφαιρώντας από κάθε μία από αυτές μία εκτίμηση της διαφοράς των μέσων ή των διαμέσων των πληθυσμών, αντίστοιχα. Εναλλακτικά, μπορούμε να μετασχηματίσουμε τις παρατηρήσεις του δείγματος που προέρχεται από τον πληθυσμό με την χαμηλότερη μέση τιμή ή διάμεσο προσθέτοντας σε κάθε μία από τις παρατηρήσεις του την εκτίμηση της διαφοράς στις μέσες τιμές ή στις διαμέσους των πληθυσμών, αντίστοιχα. Είναι δεδομένο ότι η διασπορά δεν επηρεάζεται από μεταβολές στις θέσεις των πληθυσμών. Το αποτέλεσμα αυτού του μετασχηματισμού είναι η αύξηση της ισχύος του ελέγχου των Siegel-Tukey.

Επιπλέον, η εφαρμογή του συγκεκριμένου ελέγχου είναι παρόμοια με την εφαρμογή του ελέγχου Wilcoxon – Mann Whitney(έλεγχος που χρησιμοποιείται για την ισότητα των μέσων δύο πληθυσμών).

Έστω $V(X)$ και $V(Y)$ οι διασπορές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα . Οι υποθέσεις που θα ελεγχτούν για την ισότητα των διασπορών των δύο πληθυσμών, είναι οι εξής:

- $H_0: V(X) = V(Y)$
- $H_1: V(X) \neq V(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες διασπορές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες διασπορές.

Δηλαδή:

- H_0: Ίσες διασπορές
- H_1: Δεν έχουν ίσες διασπορές

Έστω το τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ μεγέθους n από τον πληθυσμό 1 και έστω ένα άλλο τυχαίο δείγμα $Y_1, Y_2, \ldots, Y_m$ μεγέθους m από τον πληθυσμό 2, τα οποία είναι ανεξάρτητα μεταξύ τους. Το δείγμα που προκύπτει από την συνένωση των δυο δειγμάτων έχει μέγεθος $N = n + m$.

Έστω, επίσης, ότι $R(X_i)$ και $R(Y_i)$ οι βαθμοί των μεταβλητών X_i και Y_i για κάθε i και j. Σε περίπτωση που αρκετές τιμές του δείγματος ταυτίζονται, αντιστοιχίζουμε σε κάθε μία από αυτές τον μέσο των βαθμών που θα είχαν αν δεν ταυτίζονταν.

Η στατιστική συνάρτηση T για τον έλεγχο των συγκεκριμένων υποθέσεων ισούται με το άθροισμα των βαθμών που οι παρατηρήσεις $X_1, X_2, \ldots, X_n$ έχουν στο τελικό δείγμα των $n + m$ παρατηρήσεων, δηλαδή

$$T = \sum_{i=1}^{n} R(X_i)$$

Τα ποσοστιαία σημεία της κατανομής της στατιστικής συνάρτησης T δίνονται στον πίνακα 5 του παραρτήματος. Είναι προφανές ότι οι πολύ μεγάλες ή οι πολύ μικρές τιμές της στατιστικής συνάρτησης είναι ένδειξη απόρριψης της μηδενικής υπόθεσης. Δηλαδή, η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας ίσο με α, αν η τιμή της στατιστικής συνάρτησης είναι μικρότερη από το $\alpha/2$-ποσοστιαίο σημείο της κατανομής της ή αν είναι μεγαλύτερη από το $(1-\alpha/2)$-ποσοστιαίο σημείο της κατανομής της. Δηλαδή, η μηδενική υπόθεση θα απορρίπτεται αν $T<w_{a/2}$ή αν $T>w_{1-a/2}$.

<u>Παράδειγμα</u>

Ένας φοιτητής του Πανεπιστημίου Θεσσαλίας έχει στη βιβλιοθήκη του 257 βιβλία, από τα οποία τα 114 είναι βιβλία στατιστικής και τα 143 είναι βιβλία πληροφορικής. Ο φοιτητής επιλέγει δύο ανεξάρτητα τυχαία δείγματα βιβλίων, ένα από κάθε κατηγορία, μεγέθους 12 και 16 αντίστοιχα και καταγράφει τους αριθμούς σελίδων τους.

	Αριθμός σελίδων
Βιβλία στατιστικής	126, 142, 156, 228, 245, 246, 370, 419, 433, 454, 478, 503
Βιβλία πληροφορικής	29, 39, 60, 78, 82, 112, 125, 170, 192, 224, 263, 275, 276, 286, 369, 756

Πίνακας 15: Τα δύο δείγματα βιβλίων και οι αριθμοί σελίδων τους

Εφαρμόζοντας τον έλεγχο Siegel-Tukey θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν υπάρχει ισότητα των διασπορών των δύο πληθυσμών, με βάση τα δύο ανεξάρτητα δείγματα που επέλεξε ο φοιτητής.

Προκειμένου να έχουμε μία εκτίμηση της διαφοράς των θέσεων των δύο πληθυσμών, σχηματίζουμε όλες τις δυνατές διαφορές τιμών της μορφής $X_i - Y_j,\ i = 1, 2, \dots, 12$, $j = 1, 2, \dots, 16$ και χρησιμοποιούμε την διάμεσο των 192 διαφορών που προκύπτουν. Έτσι, εκτιμούμε ότι οι δύο πληθυσμοί παρουσιάζουν διαφορά ως προς την θέση ίση με 133.5. Αυτό είναι ορατό σύμφωνα με τον παρακάτω πίνακα:

Διαφορές αριθμών σελίδων βιβλίων												
y_i/x_i	126	142	156	228	245	246	370	419	433	454	478	503

29	97	113	127	199	216	217	341	390	404	425	449	474
39	87	103	117	189	206	207	331	380	394	415	439	464
60	66	82	96	168	185	186	310	359	373	394	418	443
78	48	64	78	150	167	168	292	341	355	376	400	425
82	44	60	74	146	163	164	288	337	351	372	396	421
112	14	30	44	116	133	134	258	307	321	342	366	391
125	1	17	31	103	120	121	245	294	308	329	353	378
170	-44	-28	-14	58	75	76	200	249	263	284	308	333
192	-66	-50	-36	36	53	54	178	227	241	262	286	311
224	-98	-82	-68	4	21	22	146	195	209	230	254	279
263	-137	-121	-107	-35	-18	-17	107	156	170	191	215	240
275	-149	-133	-119	-47	-30	-29	95	144	158	179	203	228
276	-150	-134	-120	-48	-31	-30	94	143	157	178	202	227
286	-160	-144	-130	-58	-41	-40	84	133	147	168	192	217
369	-243	-227	-213	-141	-124	-123	1	50	64	85	109	134
736	-630	-614	-600	-528	-511	-510	-386	-337	-323	-302	-278	-253

Πίνακας 16: Διαφορές αριθμών σελίδων βιβλίων

Στη συνέχεια προσθέτουμε τη διάμεσο στις τιμές των βιβλίων πληροφορικής, συνενώνουμε τα δύο δείγματα σελίδων και τα διατάσσουμε κατά αύξουσα σειρά μεγέθους. Σε κάθε παρατήρηση αντιστοιχίζουμε τον βαθμό που της αντιστοιχεί, δηλαδή τον βαθμό 1 στην μικρότερη παρατήρηση, 2 στην μεγαλύτερη, 3 στην επόμενη μεγαλύτερη, 4 και 5 στις επόμενες δύο πιο μικρές, 6 και 7 στις επόμενες πιο μεγάλες και ούτω καθεξής. Το αποτέλεσμα είναι:

Πλήθος σελίδων	126	142	156	162.5	172.5	193.5	211.5	215.5
Βαθμός	1	4	5	8	9	12	13	16
Πλήθος σελίδων	228	245	245.5	246	258.5	303.5	325.5	357.5
Βαθμός	17	20	21	24	25	28	27	26
Πλήθος σελίδων	370	396.5	408.5	409.5	419	419.5	433	454
Βαθμός	23	22	19	18	15	14	11	10
Πλήθος σελίδων	478	502.5	503	889.5				
Βαθμός	7	6	3	2				

Πίνακας 17: Πλήθος σελίδων με τον βαθμό τους

Οι υπογραμμισμένες σελίδες ανήκουν στα βιβλία στατιστικής.

Χρησιμοποιούμε την στατιστική συνάρτηση:

$$T = \sum_{i=1}^{n} R(X_i) = 140.$$

Τα ποσοστιαία σημεία της κατανομής της στατιστικής συνάρτησης T δίνονται στον πίνακα 5 του παραρτήματος. Δηλαδή, η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας ίσο με α, αν η τιμή της στατιστικής συνάρτησης είναι μικρότερη από το $\alpha/2$-ποσοστιαίο σημείο της κατανομής της ή αν είναι μεγαλύτερη από το $(1-\alpha/2)$-ποσοστιαίο σημείο της κατανομής της. Δηλαδή, η μηδενική υπόθεση θα απορρίπτεται αν $T<w_{a/2}$ ή αν $T>w_{1-a/2}$.

Στην περίπτωση μας: για $n = 12$ και $\mu = 16$ έχουμε $w_{0.025} = 132$ και $w_{0.975} = 12(12 + 16 + 1) - w_{0.025} = 216$. Επομένως, σύμφωνα με τα προηγούμενα δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας 5%.

Μέθοδος των Τετραγωνικών Βαθμών Μεγέθους

Ο έλεγχος Mood αποτελεί τον πρώτο μη παραμετρικό έλεγχο για τις διασπορές και παρουσιάστηκε από τον Mood το 1954, από όπου πήρε και το όνομά του. Ο έλεγχος αυτός, στην εφαρμογή του, είναι ανάλογος με τους ελέγχους που εξετάσθηκαν για την ισότητα μέσων τιμών. Για παράδειγμα, για τον έλεγχο της υπόθεσης $H_0: E(X) = E(Y)$, τα δύο ανεξάρτητα δείγματα που χρησιμοποιούνται, ένα από κάθε πληθυσμό, συνενώνονται σε ένα δείγμα, οι παρατηρήσεις διατάσσονται κατά αύξουσα σειρά μεγέθους και το άθροισμα των βαθμών που αντιστοιχούν στις παρατηρήσεις που προέρχονται από το ένα δείγμα, π.χ. το δείγμα των X, χρησιμοποιείται ως ελεγχοσυνάρτηση.

Στην στατιστική, η διασπορά ενός πληθυσμού ορίζεται ως η αναμενόμενη τιμή του τετραγώνου της διαφοράς της από την μέση τιμή της. Για αυτό για τον έλεγχο της υπόθεσης $H_0: E[(X - \mu_X)^2] = E[(Y - \mu_Y)^2]$, φαίνεται εύλογο να μεταχειρισθούμε τις τιμές των $(X_i - \mu_X)^2$ και $(Y_j - \mu_Y)^2$ από δύο ανεξάρτητα δείγματα $\{X_1, X_2, \dots, X_n\}$ και $\{Y_1, Y_2, \dots, Y_m\}$ όπως μεταχειριζόμαστε τις τιμές των X_i και Y_j για την περίπτωση των μέσων και να χρησιμοποιήσουμε το άθροισμα των βαθμών των τιμών $(X - \mu_X)^2$ ως ελεγχοσυνάρτηση.

Έστω το τυχαίο δείγμα $X_1, X_2, \dots, X_n$ μεγέθους n από τον πληθυσμό 1 και έστω ένα άλλο τυχαίο δείγμα $Y_1, Y_2, \dots, Y_m$ μεγέθους m από τον πληθυσμό 2, τα οποία είναι ανεξάρτητα μεταξύ τους. Το δείγμα που προκύπτει από την συνένωση των δυο δειγμάτων έχει μέγεθος $N = n + m$.

Έστω $V(X)$ και $V(Y)$ οι διασπορές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα . Οι υποθέσεις που θα ελεγχτούν για την ισότητα των διασπορών των δύο πληθυσμών, είναι οι εξής:

- $H_0: V(X) = V(Y)$
- $H_1: V(X) \neq V(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες διασπορές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες διασπορές.

Δηλαδή:

- H_0: Ίσες διασπορές

- H_1: Δεν έχουν ίσες διασπορές

Η μεθοδολογία που θα ακολουθήσουμε είναι η εξής:

1. Θεωρούμε τις μεταβλητές $U_i = |X_i - \mu_X|$ και $V_j = |Y_j - \mu_Y|$, με $i = 1,2,\ldots,n$ και $j = 1,2,\ldots,m$, όπου $\mu_X = E(X)$ και $\mu_Y = E(Y)$. Σε περίπτωση που δεν γνωρίζουμε τις τιμές των μ_X και μ_Y τις αντικαθιστούμε με τους δειγματικούς μέσους $\overline{X_n}$ και $\overline{Y_m}$
2. Παρατήρηση: Η διάταξη κατά αύξουσα σειρά μεγέθους των μεταβλητών $U_i = |X_i - \mu_X|$ και $V_j = |Y_j - \mu_Y|$, ισοδυναμεί με την διάταξη κατά αύξουσα σειρά μεγέθους των τιμών $(X_i - \mu_X)^2$ και $(Y_j - \mu_Y)^2$
3. Πραγματοποιείται συνένωση των δειγμάτων των U_i και V_j
4. Διατάσσουμε κατά αύξουσα σειρά μεγέθους το δείγμα που προέκυψε και θεωρούμε το δείγμα των βαθμών που αντιστοιχίζονται στις τιμές του δείγματος. Αν δύο ή περισσότερες τιμές ταυτίζονται, σε κάθε μία από αυτές αντιστοιχίζεται ο μέσος των βαθμών που θα είχαν, αν δεν ταυτίζονταν

Έστω $R(U_i), i = 1,2,\ldots,n$ και $R(V_j), j = 1,2,\ldots,m$ τους βαθμούς που αντιστοιχούν στις τιμές που προέρχονται από το δείγμα των $X_i, i = 1,2,\ldots,n$ και στις τιμές που προέρχονται από το δείγμα των $Y_j, j = 1,2,\ldots,m$ αντίστοιχα.

Η στατιστική συνάρτηση T που θα χρησιμοποιηθεί είναι το άθροισμα των τετραγώνων των βαθμών που έχουν οι παρατηρήσεις $U_1, U_2, \ldots, U_n$ στο ενιαίο δείγμα των $n + m$ παρατηρήσεων. Δηλαδή:

$$T = \sum_{i=1}^{n}[R(U_i)]^2$$

Σε περίπτωση που υπάρχουν πολλές περιπτώσεις ταύτισης τιμών στο δείγμα χρησιμοποιείται η T_1.

$$T_1 = \frac{T - n\bar{R}^2}{\left[\frac{nm}{(N(N-1)}\sum_{k=1}^{N} {R_k}^4 - \frac{nm}{N-1}\left(\overline{R^2}\right)^2\right]^{1/2}}$$

Όπου $N = n + m$, $\bar{R}^2$ συμβολίζει το μέσο των τετραγώνων των βαθμών των παρατηρήσεων στο συνενωμένο δείγμα, δηλαδή:

$$\bar{R}^2 = \frac{1}{N}\left\{ \sum_{i=1}^{n}[R(U_i)]^2 + \sum_{j=1}^{m}\left[R\left(V_j\right)\right]^2 \right\}$$

και $\sum \mathrm{R}_\mathrm{k}^4$ συμβολίζει το άθροισμα της τέταρτης δύναμης των βαθμών των παρατηρήσεων στο συνολικό δείγμα. Δηλαδή:

$$\sum_{k=1}^{N} R_k^4 = \sum_{i=1}^{n}[R(U_i)]^4 + \sum_{j=1}^{m}\left[R\left(V_j\right)\right]^4.$$

Τα ποσοστιαία σημεία της ακριβούς μορφής της κατανομής της στατιστικής συνάρτησης T δίνονται στον πίνακα 5 του παραρτήματος όταν δεν υπάρχουν περιπτώσεις ταύτισης τιμών. Στην περίπτωση μεγάλων δειγμάτων, τα ποσοστιαία σημεία της κατανομής της στατιστικής συνάρτησης T προσεγγίζονται από τα αντίστοιχα ποσοστιαία σημεία της τυποποιημένης κανονικής κατανομής σύμφωνα με την σχέση:

$$w_p = \frac{n(N+1)(2N+1)}{6} + z_p\sqrt{\frac{mn(N+1)(2N+1)(8N+1)}{180}},$$

Η κατανομή της στατιστικής συνάρτησης T_1 προσεγγίζεται ικανοποιητικά από την τυποποιημένη κανονική κατανομή. Είναι προφανές ότι η μηδενική υπόθεση θα απορρίπτεται σε επίπεδο σημαντικότητας α, αν η στατιστική συνάρτηση T έχει τιμή μικρότερη από το $\alpha/2$-ποσοστιαίο σημείο της ή μεγαλύτερη από το $(1-\alpha/2)$-ποσοστιαίο σημείο της. Δηλαδή, η μηδενική υπόθεση απορρίπτεται αν $T < w_{a/2}$ ή $T > w_{1-a/2}$. Με αντίστοιχο τρόπο ορίζεται η κρίσιμη περιοχή στην περίπτωση που χρησιμοποιείται η στατιστική συνάρτηση T_1.

<u>Παράδειγμα</u>

Σε ένα λύκειο, ο καθηγητής των μαθηματικών έβαλε όλους τους μαθητές του, κορίτσια και αγόρια, να γράψουν ένα διαγώνισμα με άριστα το 20. Κράτησε τους βαθμούς των γραπτών από 12 κορίτσια και από 16 αγόρια και ήθελε να εξετάσει αν οι διασπορές των δύο πληθυσμών ισούνται-ταυτίζονται. Τα αποτελέσματα αυτά λαμβάνουν θέση στον ακριβώς παρακάτω πίνακα:

	Βαθμοί
Κορίτσια	126, 142, 156, 228, 245, 246, 370, 419, 433, 454, 478, 503
Αγόρια	29, 39, 60, 78, 82, 112, 125, 170, 192, 224, 263, 275, 276, 286, 369, 756

Πίνακας 18: Τα δύο δείγματα βιβλίων και ο αριθμός σελίδων τους

Εφαρμόζοντας τον έλεγχο των Τετραγωνικών Βαθμών Μεγέθους θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν υπάρχει ισότητα των διασπορών των δύο πληθυσμών, με βάση τα δύο ανεξάρτητα δείγματα που επέλεξε ο καθηγητής.

Έστω $V(X)$ και $V(Y)$ οι διασπορές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα . Οι υποθέσεις που θα ελεγχτούν για την ισότητα των διασπορών των δύο πληθυσμών, είναι οι εξής:

- $H_0: V(X) = V(Y)$
- $H_1: V(X) \neq V(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες διασπορές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες διασπορές.

Δηλαδή:

- H_0: Ίσες διασπορές
- H_1: Δεν έχουν ίσες διασπορές

Έστω $X_i, i = 1,2, \dots, 12$, οι βαθμοί στο δείγμα των 12 κοριτσιών και $Y_j, j = 1, 2, \dots, 16$, οι βαθμοί στο δείγμα των 16 αγοριών. Έστω $U_i = |X_i - \overline{X_n}|$ και $V_j = |Y_j - \overline{Y_m}|$, με $i = 1,2, \dots, n$ και $j = 1,2, \dots, m$, τα δείγματα των απολύτων διαφορών των βαθμών των δύο κατηγοριών από τους αντίστοιχους μέσους όρους τους. Οι τιμές των μέσων είναι $\bar{x}$ = 316.67 και $\bar{y}$ = 208.5.

Συνενώνουμε τις τιμές αυτών των απολύτων διαφορών σε ένα ενιαίο δείγμα και τις διατάσσουμε κατά αύξουσα σειρά μεγέθους(1^{η} γραμμή) μαζί με τις τιμές που αντιστοιχούν στα κορίτσια(υπογραμμισμένες τιμές). Επιπλέον, σημειώνονται οι βαθμοί των παρατηρήσεων και τα τετράγωνά τους στον πίνακα που ακολουθεί:

Απόκλιση	15.5	16.5	38.5	53.3	54.5	66.5	67.5	70.7	71.7
Βαθμός	1	2	3	4	5	6	7	8	9
Τετράγωνο	1	4	9	16	25	36	49	64	81
Απόκλιση	77.5	83.5	88.7	96.5	102.3	116.3	126.5	130.5	137.3
Βαθμός	10	11	12	13	14	15	16	17	18
Τετράγωνο	100	121	144	169	196	225	256	289	324
Απόκλιση	148.5	160.5	160.7	161.3	169.5	174.7	179.5	186.3	190.7
Βαθμός	19	20	21	22	23	24	25	26	27
Τετράγωνο	361	400	441	484	529	576	625	676	729
Απόκλιση	547.5								
Βαθμός	28								
Τετράγωνο	784								

Πίνακας 19: Αποκλίσεις, βαθμοί και τετράγωνα βαθμών

Έστω $R(U_i), i = 1,2, \dots, n$ και $R(V_j), j = 1,2, \dots, m$ τους βαθμούς που αντιστοιχούν στις τιμές που προέρχονται από το δείγμα των $X_i, i = 1,2, \dots, n$ και στις τιμές που προέρχονται από το δείγμα των $Y_j, j = 1,2, \dots, m$ αντίστοιχα.

Η στατιστική συνάρτηση T που θα χρησιμοποιηθεί είναι το άθροισμα των τετραγώνων των βαθμών που έχουν οι παρατηρήσεις $U_1, U_2, \dots, U_n$ στο ενιαίο δείγμα των $n+m$ παρατηρήσεων. Δηλαδή:

$$T = \sum_{i=1}^{n} [R(U_i)]^2$$

Επομένως, το άθροισμα των υπογραμμισμένων τετραγώνων των βαθμών ισούται με 3956, επομένως η στατιστική συνάρτηση T έχει την τιμή 3956. Τα ποσοστιαία σημεία της ακριβούς μορφής της κατανομής της στατιστικής συνάρτησης T δίνονται στον πίνακα 5 του παραρτήματος όταν δεν υπάρχουν περιπτώσεις ταύτισης τιμών. Σύμφωνα με τον πίνακα, έχουμε $w_{0.025}$ = 132.

Στην περίπτωση μεγάλων δειγμάτων, τα ποσοστιαία σημεία της κατανομής της στατιστικής συνάρτησης T προσεγγίζονται από τα αντίστοιχα ποσοστιαία σημεία της τυποποιημένης κανονικής κατανομής σύμφωνα με την σχέση:

$$w_p = \frac{n(N+1)(2N+1)}{6} + z_p \sqrt{\frac{mn(N+1)(2N+1)(8N+1)}{180}},$$

Επομένως, $w_{0.975} = 3306 + z_{0.975}$ 629.96 = 3306 + (1.96)(629.96) = 4540.53

Ξέρουμε ότι η μηδενική υπόθεση θα απορρίπτεται σε επίπεδο σημαντικότητας α, αν η στατιστική συνάρτηση T έχει τιμή μικρότερη από το $\alpha/2$-ποσοστιαίο σημείο της ή μεγαλύτερη από το (1$-\alpha/2$)-ποσοστιαίο σημείο της. Δηλαδή, η μηδενική υπόθεση απορρίπτεται αν $T < w_{a/2}$ ή $T > w_{1-a/2}$. Με βάση τα προηγούμενα, η τιμή της στατιστικής συνάρτησης T βρίσκεται εκτός της κρίσιμης περιοχής. Επομένως, δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας 0.05.

Μέθοδος των Freund, Ansari και Bradley

Παρουσιάστηκε από τους Freund και Ansari το 1957 και αναπτύχθηκε περεταίρω από τους Ansari και Bradley το 1960. Τον συμβολίζουμε με F-A-B και στην ουσία αποτελεί μια παραλλαγή του ελέγχου Mood.

Πιο συγκεκριμένα, έστω το τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ μεγέθους n από τον πληθυσμό 1 και έστω ένα άλλο τυχαίο δείγμα $Y_1, Y_2, \ldots, Y_m$ μεγέθους m από τον πληθυσμό 2, τα οποία είναι ανεξάρτητα μεταξύ τους. Υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες διαμέσους. Το δείγμα που προκύπτει από την συνένωση των δυο δειγμάτων έχει μέγεθος $N = n + m$. Έστω $V(X)$ και $V(Y)$ οι διασπορές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα . Οι υποθέσεις που θα ελεγχτούν για την ισότητα των διασπορών των δύο πληθυσμών, είναι οι εξής:

- $H_0: V(X) = V(Y)$
- $H_1: V(X) \neq V(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες διασπορές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες διασπορές.

Δηλαδή:

- H_0: Ίσες διασπορές
- H_1: Δεν έχουν ίσες διασπορές

Η ελεγχοσυνάρτηση που θα χρησιμοποιήσουμε έχει τη μορφή:

$$AB = \sum_{i=1}^{m} R_i(X),$$

όπου AB είναι το άθροισμα των βαθμών στις τιμές της X.

Η μεθοδολογία που θα ακολουθήσουμε είναι η εξής:

1. Τοποθετούμε το σύνολο των δεδομένων $N = n + m$, από την μικρότερη τιμή στην μεγαλύτερη τιμή
2. Η μικρότερη και η μεγαλύτερη τιμή στο ενιαίο δείγμα λαμβάνουν το βαθμό 1
3. Η αμέσως επόμενη μικρότερη και η αμέσως επόμενη μεγαλύτερη τιμή στο ενιαίο δείγμα λαμβάνουν το βαθμό 2 και ούτω καθεξής

4. Αν το N είναι άρτιος αριθμός, οι βαθμοί εισάγονται ως εξής: $1,2,3, \dots, N/2, N/2, \dots, 3,2,1$
5. Αν το N είναι περιττός αριθμός, οι βαθμοί εισάγονται ως εξής: $1,2,3, \dots, (N-1)/2, (N+1)/2, (N-1)/2, \dots, 3,2,1$

Η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας ίσο με α, αν η τιμή της στατιστικής συνάρτησης είναι μικρότερη από το $(1-\alpha/2)$-ποσοστιαίο σημείο της κατανομής της ή αν είναι μεγαλύτερη από το $\alpha/2$-ποσοστιαίο σημείο της κατανομής της. Δηλαδή, η μηδενική υπόθεση θα απορρίπτεται αν $AB < w_{1-\alpha/2}$ ή αν $AB > w_{\alpha/2}$.

<u>Παράδειγμα</u>

Έχουμε δεδομένα από έναν καρδιακό δείκτη σε δύο ομάδες ασθενών. Στην πρώτη ομάδα έχουμε τους ασθενείς με κανονική προσθετική λειτουργία βαλβίδας, ενώ στην δεύτερη έχουμε τους ασθενείς με μη κανονική προσθετική λειτουργία βαλβίδας.

Group1(X)	3.84	2.6	1.19	2.00	
Group2(Y)	3.97	2.5	2.7	3.36	2.30

Πίνακας 20: Καρδιακός δείκτης των δύο ομάδων

Εφαρμόζοντας τον έλεγχο των Freund, Ansari και Bradley θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν οι διασπορές των 2 πληθυσμών ισούνται, με βάση τα δύο ανεξάρτητα δείγματα που επιλέχθηκαν.

Έστω $V(X)$ και $V(Y)$ οι διασπορές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα . Οι υποθέσεις που θα ελεγχτούν για την ισότητα των διασπορών των δύο πληθυσμών, είναι οι εξής:

- $H_0: V(X) = V(Y)$
- $H_1: V(X) \neq V(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες διασπορές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες διασπορές.

Δηλαδή:

- H_0: Ίσες διασπορές
- H_1: Δεν έχουν ίσες διασπορές

Συνδυάζοντας τα δύο δείγματα($N = 9$) έχουμε:

	1.19	2	2.3	2.5	2.6	2.7	3.36	3.84	3.97
Group	x	x	y	y	x	y	y	x	y
Βαθμός	1	2	3	4	5	4	3	2	1

Πίνακας 21: Συνενωμένο δείγμα και βαθμοί

Η ελεγχοσυνάρτηση που θα χρησιμοποιήσουμε έχει τη μορφή:

$$AB = \sum_{i=1}^{m} R_i(X),$$

$$= 1 + 2 + 5 + 2$$

$$= 10$$

Η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας ίσο με α, αν η τιμή της στατιστικής συνάρτησης είναι μικρότερη από το $(1-\alpha/2)$-ποσοστιαίο σημείο της κατανομής της ή αν είναι μεγαλύτερη από το $\alpha/2$-ποσοστιαίο σημείο της κατανομής της. Δηλαδή, η μηδενική υπόθεση θα απορρίπτεται αν $AB < w_{1-\alpha/2}$ ή αν $AB > w_{\alpha/2}$.

Στην περίπτωση μας, έχουμε $w_{1-\alpha/2} = 8$ και $w_{\alpha/2} = 16$. Επομένως, δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση.

Μέθοδος των Barton-David

Ο B-D έλεγχος, όπως τον συμβολίζουμε, παρουσιάστηκε από τους Barton και David το 1958 λίγο μετά τον έλεγχο F-A-B και είναι καταρχήν παρόμοιος με αυτόν. Ενώ οι τιμές του F-A-B είναι σε τριγωνικό σχήμα, οι τιμές του B-D ακολουθούν ένα σχήμα V με μεγάλες τιμές στα άκρα και μικρές τιμές στη μεγάλη διάμεσο. Το αποτέλεσμα είναι ένας έλεγχος ίδιας ευρωστίας και ισχύος με το F-A-B.

Πιο συγκεκριμένα, έστω το τυχαίο δείγμα $X_1, X_2, \dots, X_n$ μεγέθους n από τον πληθυσμό 1 και έστω ένα άλλο τυχαίο δείγμα $Y_1, Y_2, \dots, Y_m$ μεγέθους m από τον πληθυσμό 2, τα οποία είναι ανεξάρτητα μεταξύ τους. Το δείγμα που προκύπτει από την συνένωση των δυο δειγμάτων έχει μέγεθος $N = n + m$ και το διατάσσουμε σε αύξουσα σειρά μεγέθους.

Έστω $V(X)$ και $V(Y)$ οι διασπορές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα . Οι υποθέσεις που θα ελεγχτούν για την ισότητα των διασπορών των δύο πληθυσμών, είναι οι εξής:

- $H_0: V(X) \ = \ V(Y)$
- $H_1: V(X) \ \neq \ V(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες διασπορές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες διασπορές.

Δηλαδή:

- H_0: Ίσες διασπορές
- H_1: Δεν έχουν ίσες διασπορές

Η ελεγχοσυνάρτηση των David και Barton είναι πιο πολύπλοκη και έχει τη μορφή:

$$B_N = \sum_{i=1}^{\left[\frac{N+1}{2}\right]} \left(\left[\frac{N+2}{2}\right] - i\right) z_i + \sum_{i=\left[\frac{N+1}{2}\right]+1}^{N} \left(i - \left[\frac{N+1}{2}\right]\right) z_i.$$

Όπου $z_i \ = \begin{cases} 0, \alpha\nu\ \eta\ i - o\sigma\tau\acute{\eta}\ \pi\alpha\rho\alpha\tau\acute{\eta}\rho\eta\sigma\eta\ \alpha\nu\acute{\eta}\kappa\varepsilon\iota\ \sigma\tau o\ \delta\varepsilon\acute{\iota}\gamma\mu\alpha\ \tau\omega\nu\ Y \\ 1,\ \alpha\nu\ \eta\ i - o\sigma\tau\acute{\eta}\ \pi\alpha\rho\alpha\tau\acute{\eta}\rho\eta\sigma\eta\ \alpha\nu\acute{\eta}\kappa\varepsilon\iota\ \sigma\tau o\ \delta\varepsilon\acute{\iota}\gamma\mu\alpha\ \tau\omega\nu\ X \end{cases}$

Η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας ίσο με α, αν η τιμή της στατιστικής συνάρτησης είναι μικρότερη από το $(1-\alpha/2)$-ποσοστιαίο σημείο της κατανομής της ή αν είναι μεγαλύτερη από το $\alpha/2$-ποσοστιαίο σημείο της κατανομής της. Δηλαδή, η μηδενική υπόθεση θα απορρίπτεται αν $B_N < w_{1-\alpha/2}$ ή αν $B_N > w_{\alpha/2}$.

Παράδειγμα

Σε ένα εργοστάσιο τα ελατήρια παράγονται από δύο μηχανές. Τυχαία δείγματα 5 ελατηρίων από την πρώτη μηχανή και 4 ελατηρίων από τη δεύτερη, έδωσαν τα παρακάτω μήκη (σε εκατοστά):

Πρώτη μηχανή(X)	19.5, 20.1, 20.0, 19.7, 19.4
Δεύτερη μηχανή(Y)	20.2, 19.6, 20.5, 20.3

Πίνακας 22: Μήκη ελατηρίων για κάθε μηχανή

Εφαρμόζοντας τον έλεγχο των Barton-David θα εξετάσουμε σε επίπεδο σημαντικότητας 5% αν οι διασπορές των 2 πληθυσμών ισούνται, με βάση τα δύο ανεξάρτητα δείγματα που επιλέχθηκαν.

Έστω $V(X)$ και $V(Y)$ οι διασπορές των τυχαίων μεταβλητών X και Y των πληθυσμών 1 και 2 αντίστοιχα . Οι υποθέσεις που θα ελεγχτούν για την ισότητα των διασπορών των δύο πληθυσμών, είναι οι εξής:

- $H_0: V(X) = V(Y)$
- $H_1: V(X) \neq V(Y)$

Πρακτικά, στην H_0 υποθέτουμε ότι οι δύο πληθυσμοί έχουν ίσες διασπορές, ενώ στην εναλλακτική H_1 υποθέτουμε ότι οι δύο πληθυσμοί δεν έχουν ίσες διασπορές.

Δηλαδή:

- H_0: Ίσες διασπορές
- H_1: Δεν έχουν ίσες διασπορές

Συνδυάζοντας τα δύο δείγματα$(N = 9)$ έχουμε:

	19.4	19.5	19.6	19.7	20.0	20.1	20.2	20.3	20.5
Group	X	X	Y	X	X	X	Y	Y	Y

Πίνακας 23: Συνενωμένο δείγμα και η ομάδα στην οποία ανήκει

Η ελεγχοσυνάρτηση των David και Barton είναι πιο πολύπλοκη και έχει τη μορφή:

$$B_N = \sum_{i=1}^{[\frac{N+1}{2}]}(\left[\frac{N+2}{2}\right] - i)z_i + \sum_{i=\left[\frac{N+1}{2}\right]+1}^{N}(i - \left[\frac{N+1}{2}\right])z_i.$$

$$= 10 + 1$$

$$= 11$$

Η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας ίσο με α, αν η τιμή της στατιστικής συνάρτησης είναι μικρότερη από το $(1-\alpha/2)$-ποσοστιαίο σημείο της κατανομής της ή αν είναι μεγαλύτερη από το $\alpha/2$-ποσοστιαίο σημείο της κατανομής της. Δηλαδή, η μηδενική υπόθεση θα απορρίπτεται αν $B_N < w_{1-\alpha/2}$ ή αν $B_N > w_{\alpha/2}$.

Στην περίπτωση μας, έχουμε $w_{1-\alpha/2} = 8$ και $w_{\alpha/2} = 16$. Επομένως, δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση.

Μέθοδος του Capon και Klotz

Αντί για τη χρήση των scores που είναι μια τετραγωνική συνάρτηση των βαθμών όπως είχε κάνει ο Mood, ο Capon (1961) πρότεινε την επιλογή των scores που δίνουν τη βέλτιστη ισχύ υπό κάποια έννοια. Το αποτέλεσμα είναι αυτός ο κανονικός έλεγχος scores, ο οποίος είναι τοπικά ο ισχυρότερος μεταξύ των βαθμωτών ελέγχων έναντι στις κανονικού τύπου εναλλακτικές, και τοπικά ασυμπτωτικά ο ισχυρότερος μεταξύ όλων των ελέγχων για αυτήν την εναλλακτική. Λίγο αργότερα, ο Klotz (1962) εισήγαγε έναν άλλο κανονικό έλεγχο scores όπου χρησιμοποίησε τα καταλληλότερα κανονικά ποσοστημόρια (quantiles). Το αποτέλεσμα έχει πιθανότατα μικρότερη ισχύ τοπικά για δείγματα μικρού μεγέθους, αλλά έχει τις ίδιες ασυμπτωτικές ιδιότητες όπως το Capon. Εξαιτίας αυτής της ευκολίας, προτιμάται ο έλεγχος του Klotz, αλλά όχι ο αντίστοιχος του Capon.

Μέθοδος των Flinger και Killeen

Οι Figner και Killeen το 1976 πρότειναν ταξινόμηση του $|X_{ij}|$ και καθορισμένη αύξηση των τιμών $\alpha_{N,i} = i$, $\alpha_{N,i} = i^2$, και $\alpha_{N,i} = \Phi^{-1}(1/2 + (i/2(N+1)))$ βάσει αυτών των ταξινομήσεων.

Στη συνέχεια, ακολουθούν οι παραλλαγές αυτού του ελέγχου. Προτείνουμε τη χρήση των βαθμών(ranks) $|X_{ij} - \overline{X}_i|$ και ονομάζουμε τον πρώτο έλεγχο T-G από τους Talwar και Gentle (1977) οι οποίοι χρησιμοποίησαν έναν αποκομμένο μέσο αντί του X_i. Ο δεύτερος έλεγχος, που ονομάζεται squared ranks έλεγχος, S-R, είχε μελετηθεί από τους Conover και Iman (1978), αλλά έχει ρίζες σε προγενέστερες δημοσιεύσεις των Shorack (1965), Duran και Mielke (1968), και άλλων.

ΚΕΦΑΛΑΙΟ 5 - Σύνοψη

Όπως έχει προαναφερθεί, στην διαδικασία της έρευνας τίθονται πολλά ερωτήματα. Αυτά προέρχονται από τον ίδιο τον ερευνητή, ο οποίος ενδέχεται να ενδιαφέρεται να προσδιορίσει αν οι δύο συναρτήσεις κατανομής που σχετίζονται με τους δύο πληθυσμούς ταυτίζονται ή όχι, αλλά και αν ισούνται οι παράμετροι θέσης και κλίμακας δύο ανεξάρτητων πληθυσμών. Είναι λογικό ότι με το πέρασμα του χρόνου έχουν αναπτυχθεί διάφορες στατιστικές μεθοδολογίες, οι οποίες είναι κατάλληλες για τους ανωτέρω ελέγχους. Όμως, είναι πολύ συνηθισμένο η κατανομή των δεδομένων να μην είναι γνωστή, επομένως δεν πληρούνται οι προϋποθέσεις για παραμετρικό έλεγχο. Για αυτό επινοήθηκαν εξίσου ισχυρές τεχνικές που μπορούν να χρησιμοποιηθούν τόσο για κανονικούς πληθυσμούς όσο και για μη κανονικούς πληθυσμούς και ονομάζονται μη παραμετρικές τεχνικές. Στην εργασία αυτή παρουσιάστηκαν αρκετοί μη παραμετρικοί έλεγχοι, οι οποίοι βοηθούν τον ερευνητή στα ερωτήματά του. Καθένας από αυτούς έχει τη δική του ξεχωριστή ελεγχοσυνάρτηση, η οποία καθορίζει πρακτικά αν θα αποδεχτούμε ή όχι τη μηδενική υπόθεση. Ως επίπεδο σημαντικότητας χρησιμοποιήσαμε το 5%, το πιο διαδεδομένο ποσοστό που έχει επικρατήσει στη Στατιστική.

ΠΗΓΕΣ – ΑΝΑΦΟΡΕΣ

Ελληνική Βιβλιογραφία

Αντζουλάκος, Δ. (2013). *Ανάλυση Δεδομένων με τη Χρήση Στατιστικών Πακέτων*, Πανεπιστημιακές Σημειώσεις, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς.

Εμβαλωτής, Α., Κατσής, Α., & Σιδερίδης, Γ. (2006). *Στατιστική Μεθοδολογία Εκπαιδευτικής Έρευνας*, Πιλοτική Έκδοση.

Ζιντζαράς, Η.(2020). *Περιγραφική Στατιστική*, Πανεπιστημιακές Σημειώσεις, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας.

Κολυβά Μαχαίρα, Φ., & Χατζόπουλος, Σ. (2015). *Μαθηματική στατιστική*, Εκδόσεις Κάλλιπος.

Μίγγος, Δ. (2010). *Παραμετρικοί και μη παραμετρικοί έλεγχοι υποθέσεων*, Διατριβή, ΤΕΙ Δυτικής Μακεδονίας.

Ξεκαλάκη, Ε. (2001). *Μη Παραμετρική Στατιστική*, Εκδόσεις Μπένου.

Παπαδόπουλος, Γ. (2013). *Περιγραφική Στατιστική*, Πανεπιστημιακές Σημειώσεις, Γεωπονικό Πανεπιστημίο Αθηνών.

Τσιτσιρίγκος, Κ. (2011). *Μελέτη αλγορίθμων εύρεσης ελαχίστων ζειγνυόντων δέντρων*, Μεταπτυχιακή Διατριβή, Τμήμα Πληροφορικής Πανεπιστήμιο Πειραιώς.

Ξενόγλωσση Βιβλιογραφία

Anderson, T. W. (1962). On the Annals of Mathematical Statistics, *The Distribution of the Two-Sample Cramér-von Mises Criterion*, *33*(3), 1148–1159.

Friedman, J. H., & Rafsky, L. C. (1979). On the Annals of Statistics, *Multivariate Generalizations of the Wald-Wolfowitz and Smirnov Two-Sample Tests, 7*(4), 697–717.

Hart, A. (2001). On the BMJ, *Mann-Whitney test is not just a test of medians: differences in spread can be important*, *323*(7309), 391–393.

Kieu-Xuan, T., & Koo, I. (2011). *Cramer-von Mises test based spectrum sensing for cognitive radio systems*. IEEE Xplore.

Li, S. (2017). *Nonparametric variance control charts based on siegel-tukey test*. IEEE Xplore.

Nornadiah Mohd Razali, & Bee Wah Yap. (2011). *Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests*. ResearchGate.

Sprent, P., & Smeeton, N. (2016). *Applied Nonparametric Statistical Methods*, 4th edition, CRC Press, New York.

Tucker, T. (2020). *Έλεγχος Υποθέσεων*, Πανεπιστημιακές Σημειώσεις, Wilfrid Laurier Univeristy.

Vickers, A. J. (2005). On the BMC Medical Research Methodology, *Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data*, *5*(1).

ΕΥΡΕΤΗΡΙΟ ΟΡΩΝ

ΠΑΡΑΡΤΗΜΑ

Πίνακες

Πίνακας 1

p- ποσοστιαία σημεία της Ελεγχοσυνάρτησης Smirnov για την περίπτωση δύο δειγμάτων μεγέθους n:

	$p=.80$	.90	.95	.98	.99
$n=3$	2/3	2/3			
4	3/4	3/4	3/4		
5	3/5	3/5	4/5	4/5	4/5
6	3/6	4/6	4/6	5/6	5/6
7	4/7	4/7	5/7	5/7	5/7
8	4/8	4/8	5/8	5/8	6/8
9	4/9	5/9	5/9	6/9	6/9
10	4/10	5/10	6/10	6/10	7/10
11	5/11	5/11	6/11	7/11	7/11
12	5/12	5/12	6/12	7/12	7/12
13	5/13	6/13	6/13	7/13	8/13
14	5/14	6/14	7/14	7/14	8/14
15	5/15	6/15	7/15	8/15	8/15
16	6/16	6/16	7/16	8/16	9/16
17	6/17	7/17	7/17	8/17	9/17
18	6/18	7/18	8/18	9/18	9/18
19	6/19	7/19	8/19	9/19	9/19
20	6/20	7/20	8/20	9/20	10/20
21	6/21	7/21	8/21	9/21	10/21
22	7/22	8/22	8/22	10/22	10/22
23	7/23	8/23	9/23	10/23	10/23
24	7/24	8/24	9/24	10/24	11/24
25	7/25	8/25	9/25	10/25	11/25
26	7/26	8/26	9/26	10/26	11/26
27	7/27	8/27	9/27	11/27	11/27
28	8/28	9/28	10/28	11/28	12/28
29	8/29	9/29	10/29	11/29	12/29
30	8/30	9/30	10/30	11/30	12/30
31	8/31	9/31	10/31	11/31	12/31
32	8/32	9/32	10/32	12/32	12/32
34	8/34	10/34	11/34	12/34	13/34
36	9/36	10/36	11/36	12/36	13/36
38	9/38	10/38	11/38	13/38	14/38
40	9/40	10/40	12/40	13/40	14/40

$n > 40$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.73}{\sqrt{n}}$	$\frac{1.92}{\sqrt{n}}$	$\frac{2.15}{\sqrt{n}}$	$\frac{2.30}{\sqrt{n}}$

<u>Πίνακας 2</u>

p-ποσοστιαία σημεία της Ελεγχοσυνάρτησης Smirnov για την περίπτωση δύο δειγμάτων διαφορετικού μεγέθους n και m:

		$p = .80$	.90	.95	.98	.99
N_1 = 1	N_2 = 9	17/18				
	10	9/10				
N_1 = 2	N_2 = 3	5/6				
	4	3/4				
	5	4/5	4/5			
	6	5/6	5/6			
	7	5/7	6/7			
	8	3/4	7/8	7/8		
	9	7/9	8/9	8/9		
	10	7/10	4/5	9/10		
N_1 = 3	N_2 = 4	3/4	3/4			
	5	2/3	4/5	4/5		
	6	2/3	2/3	5/6		
	7	2/3	5/7	6/7	6/7	
	8	5/8	3/4	3/4	7/8	
	9	2/3	2/3	7/9	8/9	8/9
	10	3/5	7/10	4/5	9/10	9/10
	12	7/12	2/3	3/4	5/6	11/12
N_1 = 4	N_2 = 5	3/5	3/4	4/5	4/5	
	6	7/12	2/3	3/4	5/6	5/6
	7	17/28	5/7	3/4	6/7	6/7
	8	5/8	5/8	3/4	7/8	7/8
	9	5/9	2/3	3/4	7/9	8/9
	10	11/20	13/20	7/10	4/5	4/5
	12	7/12	2/3	2/3	3/4	5/6
	16	9/16	5/8	11/16	3/4	13/16
N_1 = 5	N_2 = 6	23/42	2/3	2/3	5/6	5/6
	7	1/2	23/35	5/7	29/35	6/7
	8	1/2	5/8	27/40	4/5	4/5
	9	1/2	3/5	31/45	7/9	4/5
	10	1/2	3/5	7/10	7/10	4/5
	15	8/15	3/5	2/3	11/15	11/15
	20	1/2	11/20	3/5	7/10	3/4
N_1 = 6	N_2 = 7	23/42	4/7	29/42	5/7	5/6
	8	1/2	7/12	2/3	3/4	3/4

	9	1/2	5/9	2/3	13/18	7/9
	10	1/2	17/30	19/30	7/10	11/15
	12	1/2	7/12	7/12	2/3	3/4
	18	4/9	5/9	11/18	2/3	13/18
	24	11/24	1/2	7/12	5/8	2/3
N_1 = 7	N_2 = 8	27/56	33/56	5/8	41/56	3/4
	9	31/63	5/9	40/63	5/7	47/63
	10	33/70	39/70	43/70	7/10	5/7
	14	3/7	1/2	4/7	9/14	5/7
	28	3/7	13/28	15/28	17/28	9/14
N_1 = 8	N_2 = 9	4/9	13/24	5/8	2/3	3/4
	10	19/40	21/40	23/40	27/40	7/10
	12	11/24	1/2	7/12	5/8	2/3
	16	7/16	1/2	9/16	5/8	5/8
	32	13/32	7/16	1/2	9/16	19/32
N_1 = 9	N_2 = 10	7/15	1/2	26/45	2/3	31/45
	12	4/9	1/2	5/9	11/18	2/3
	15	19/45	22/45	8/15	3/5	29/45
	18	7/18	4/9	1/2	5/9	11/18
	36	13/36	5/12	17/36	19/36	5/9
N_1 = 10	N_2 = 15	2/5	7/15	1/2	17/30	19/30
	20	2/5	9/20	1/2	11/20	3/5
	40	7/20	2/5	9/20	1/2	
N_1 = 12	N_2 = 15	23/60	9/20	1/2	11/20	7/12
	16	3/8	7/16	23/48	13/24	7/12
	18	13/36	5/12	17/36	19/36	5/9
	20	11/30	5/12	7/15	31/60	17/30
N_1 = 15	N_2 = 20	7/20	2/5	13/30	29/60	31/60
N_1 = 16	N_2 = 20	27/80	31/80	17/40	19/40	41/80
Για μεγάλα δείγματα		$1.07\sqrt{\frac{m+n}{mn}}$	$1.22\sqrt{\frac{m+n}{mn}}$	$1.36\sqrt{\frac{m+n}{mn}}$	$1.52\sqrt{\frac{m+n}{mn}}$	$1.63\sqrt{\frac{m+n}{mn}}$

Πίνακας 3

p-ποσοστιαία σημεία της Ελεγχοσυνάρτησης Cramer-Von Mises:

$W_{0.10} = 0.046$	$W_{0.50} = 0.119$	$W_{0.10} = 0.347$
$W_{0.20} = 0.062$	$W_{0.60} = 0.147$	$W_{0.10} = 0.461$
$W_{0.30} = 0.079$	$W_{0.70} = 0.184$	$W_{0.10} = 0.743$
$W_{0.40} = 0.097$	$W_{0.80} = 0.241$	$W_{0.10} = 1.168$

Πίνακας 4

Ποσοστιαία σημεία της ελεγχοσυνάρτησης T του ελέγχου των ροών:

	n_1	2	3	4	5	6	7	8	9	10
n_2										
2										
3						2 -	2 -	2 -	2 -	2 -
4					2 9	2 9	2 -	3 -	3 -	3 -
5				2 9	2 10	3 10	3 11	3 11	3 -	3 -
6			2 -	2 9	3 10	3 11	3 12	3 12	4 13	4 13
7			2 -	2 -	3 11	3 12	3 13	4 13	4 14	5 14
8			2 -	3 -	3 11	3 12	4 13	4 14	5 14	5 15
9			2 -	3 -	3 -	4 13	4 14	5 14	5 15	5 16
10			2 -	3 -	3 -	4 13	5 14	5 15	5 16	6 16

Όπου η πάνω γραμμή αποτελεί την τιμή $w_{0.025}$ και η κάτω γραμμή την τιμή $w_{0.975}$.

Πίνακας 5

p-ποσοστιαία σημεία της Ελεγχοσυνάρτησης Mann-Whitney:

n	p	m=2	3	4	5	6	7	8	9	10	16
2	.001	3	3	3	3	3	3	3	3	3	
	.005	3	3	3	3	3	3	3	3	3	
	.01	3	3	3	3	3	3	3	3	3	
	.025	3	3	3	3	3	3	4	4	4	
	.05	3	3	3	4	4	4	5	5	5	
	.10	3	4	4	5	5	5	6	6	7	
3	.001	6	6	6	6	6	6	6	6	6	
	.005	6	6	6	6	6	6	6	7	7	
	.01	6	6	6	6	6	7	7	8	8	
	.025	6	6	6	7	8	8	9	9	10	
	.05	6	7	7	8	9	9	10	11	11	
	.10	7	8	8	9	10	11	12	12	13	
4	.001	10	10	10	10	10	10	10	10	11	

	.005	10	10	10	10	11	11	12	12	13	
	.01	10	10	10	11	12	12	13	14	14	
	.025	10	10	11	12	13	14	15	15	16	
	.05	10	11	12	13	14	15	16	17	18	
	.10	11	12	14	15	16	17	18	20	21	
5	.001	15	15	15	15	15	15	16	17	17	
	.005	15	15	15	16	17	17	18	19	20	
	.01	15	15	16	17	18	19	20	21	21	
	.025	15	16	17	18	19	21	22	23	24	
	.05	16	17	18	20	21	22	24	25	27	
	.10	17	18	20	21	23	24	26	28	29	
6	.001	21	21	21	21	21	21	23	24	25	
	.005	21	21	22	23	24	25	26	27	28	
	.01	21	21	23	24	25	26	28	29	30	
	.025	21	23	24	25	27	28	30	32	33	
	.05	22	24	25	27	29	30	32	34	36	
	.10	23	25	27	29	31	33	35	37	39	
7	.001	28	28	28	28	29	30	31	32	34	
	.005	28	28	29	30	32	33	35	36	38	
	.01	28	29	30	32	33	35	36	38	40	
	.025	28	30	32	34	35	37	39	41	43	
	.05	29	31	33	35	37	40	42	44	46	
	.10	30	33	35	37	40	42	45	47	50	
8	.001	36	36	36	37	38	39	41	42	43	
	.005	36	36	38	39	41	43	44	46	48	
	.01	36	37	39	41	43	44	46	48	50	
	.025	37	39	41	43	45	47	50	52	54	
	.05	38	40	42	45	47	50	52	55	57	
	.10	39	42	44	47	50	53	56	59	61	
9	.001	45	45	45	47	48	49	51	53	54	
	.005	45	46	47	49	51	53	55	57	59	
	.01	45	47	49	51	53	55	57	60	62	
	.025	46	48	50	53	56	58	61	63	66	
	.05	47	50	52	55	58	61	64	67	70	
	.10	48	51	55	58	61	64	68	71	74	
10	.001	55	55	56	57	59	61	62	64	66	
	.005	55	56	58	60	62	65	67	69	72	
	.01	55	57	59	62	64	67	69	72	75	
	.025	56	59	61	64	67	70	73	76	79	
	.05	57	60	63	67	70	73	76	80	83	
	.10	59	62	66	69	73	77	80	84	88	
12	.001	78	78	79	81	83	86	88	91	93	110
	.005	78	80	82	85	88	91	94	97	100	120
	.01	78	81	84	87	90	93	96	100	103	125
	.025	80	83	86	90	93	97	101	105	108	132
	.05	81	84	88	92	96	100	105	109	111	139
	.10	83	87	91	96	100	105	109	114	118	146

Για τιμές του n και m μεγαλύτερες του 20, το p-ποσοστιαίο σημείο της ελεγχοσυνάρτησης Mann-Whitney μπορεί να προσεγγισθεί από την τιμή:

$$w_p = \left[{n(N+1)}/{2}\right] + Z_p\sqrt{{nm(N+1)}/{12}},$$

όπου $N = n + m$ και Z_p είναι το p-ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

<u>Πίνακας 6</u>

Ποσοστιαία σημεία της τυποποιημένης κανονικής κατανομής:

z	0.00	0.01	0.02	0.03	0.04
0.0	0.5000	0.5040	0.5080	0.5120	0.5160
0.1	0.5398	0.5438	0.5478	0.5517	0.5557
0.2	0.5793	0.5832	0.5871	0.5910	0.5948
0.3	0.6179	0.6217	0.6255	0.6293	0.6331
0.4	0.6554	0.6591	0.6628	0.6664	0.6700
0.5	0.6915	0.6950	0.6985	0.7019	0.7054
0.6	0.7257	0.7291	0.7324	0.7357	0.7389
0.7	0.7580	0.7611	0.7642	0.7673	0.7704
0.8	0.7881	0.7910	0.7939	0.7967	0.7995
0.9	0.8159	0.8186	0.8212	0.8238	0.8264
z	0.05	0.06	0.07	0.08	0.09
0.0	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6736	0.6722	0.6808	0.6844	0.6879
0.5	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8289	0.8315	0.8340	0.8365	0.8389

Πίνακας 7

Διωνυμική κατανομή για $p = 0.5$:

$\boldsymbol{n = 2}$								
$x = 0$	$x = 1$	$x = 2$						
0.2500	0.7500	1.0000						
$\boldsymbol{n = 3}$								
$x = 0$	$x = 1$	$x = 2$	$x = 3$					
0.1250	0.5000	0.8750	1.0000					
$\boldsymbol{n = 4}$								
$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$				
0.0625	0.3125	0.6875	0.9375	1.0000				
$\boldsymbol{n = 5}$								
$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$			
0.0312	0.1875	0.5000	0.8125	0.9688	1.000			
$\boldsymbol{n = 6}$								
$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$	$x = 6$		
0.0156	0.1094	0.3438	0.6562	0.8906	0.9844	1.0000		
$\boldsymbol{n = 7}$								
$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$	$x = 6$	$x = 7$	
0.0078	0.0625	0.2266	0.5000	0.7734	0.9375	0.9922	1.0000	
$\boldsymbol{n = 8}$								
$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$	$x = 6$	$x = 7$	$x = 8$
0.0039	0.0352	0.1445	0.3633	0.6367	0.8555	0.9648	0.9961	1.0000
$\boldsymbol{n = 9}$								
$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$	$x = 6$	$x = 7$	$x = 8$
0.0020	0.0195	0.0898	0.2539	0.5000	0.7461	0.9102	0.9805	0.9980
$x = 9$								
1.0000								
$\boldsymbol{n = 10}$								
$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$	$x = 6$	$x = 7$	$x = 8$
0.0010	0.0107	0.0547	0.1719	0.3770	0.6230	0.8281	0.9453	0.9893
$x = 9$	$x = 10$							
0.9990	1.0000							

Πίνακας 8

Κρίσιμες τιμές της μεθόδου Anderson-Darling για $a = 0.05$:

m	n	Alpha crit	m	n	Alpha crit
2	4	–	3	3	–
2	5	–	3	4	2.743
2	6	2.676	3	5	3.039
2	7	2.855	3	6	2.552

2	8	3.017	3	7	2.297
2	9	3.165	3	8	2.512
2	10	3.302	3	9	2.485
2	11	2.597	3	10	2.432
2	12	2.722	3	11	2.360
2	13	2.386	3	12	2.423
2	14	2.492	3	13	2.364
2	16	2.626	3	14	2.405
2	18	2.428	3	15	2.362
4	4	2.402	3	16	2.429
4	5	2.292	3	17	2.409
4	6	2.282	5	5	2.392
4	7	2.479	5	6	2.382
4	8	2.412	5	7	2.391
4	9	2.362	5	8	2.380
4	10	2.411	5	9	2.413
4	11	2.468	5	10	2.378
4	12	2.404	5	12	2.381
4	14	2.388	5	14	2.408
4	16	2.391	5	15	2.405
6	6	2.270	7	7	2.374
6	7	2.339	7	8	2.381
6	8	2.374	7	9	2.401
6	9	2.366	7	10	2.390
6	10	2.391	7	11	2.383
6	11	2.396	8	8	2.356
6	12	2.399			