



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

Έλεγχοι Τυχαιοποίησης, Monte Carlo και Bootstrap

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΛΑΜΠΕΑ ΘΕΟΔΩΡΑΣ

Επιβλέπων: Φουσκάκης Δημήτριος
Καθηγητής ΕΜΠ

Αθήνα, Σεπτέμβριος 2024



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

Έλεγχοι Τυχαιοποίησης, Monte Carlo και Bootstrap

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΛΑΜΠΕΑ ΘΕΟΔΩΡΑΣ

Επιβλέπων: Φουσκάκης Δημήτριος
Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 24 Σεπτεμβρίου 2024.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Φουσκάκης Δημήτριος
Καθηγητής ΕΜΠ

.....
Βόντα Φιλία
Καθηγήτρια ΕΜΠ

.....
Λουλάκης Μιχαήλ
Καθηγητής ΕΜΠ

Αθήνα, Σεπτέμβριος 2024



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

Copyright ©-All rights reserved Λαμπέα Θεοδώρα, 2024.

Με την επιφύλαξη παντός δικαιώματος.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της επιτροπής που την ενέκρινε.

Υπεύθυνη Δήλωση

Βεβαιώνω ότι είμαι συγγραφέας αυτής της πτυχιακής εργασίας, και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην πτυχιακή εργασία. Επίσης έχω αναφέρει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επίσης, βεβαιώνω ότι αυτή η πτυχιακή εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για τις απαιτήσεις του προγράμματος σπουδών του τμήματος Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής του Πανεπιστημίου Πατρών.

(Υπογραφή)

.....
Λαμπέα Θεοδώρα

στον Στέφανο

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο της ολοκλήρωσης του προπτυχιακού κύκλου σπουδών μου, στη Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβιου Πολυτεχνείου.

Η ολοκλήρωση της διπλωματικής αυτής εργασίας θα ήταν αδύνατη χωρίς την πολύτιμη υποστήριξη και καθοδήγηση του καθηγητή μου, Καθηγητή ΕΜΠ, κύριου Δημητρίου Φουσκάκη. Του εκφράζω ένα ειλικρινές ευχαριστώ για όλη την βοήθεια που μου προσέφερε.

Επιπλέον θέλω να ευχαριστήσω από τα βάθη της καρδιάς μου την οικογένεια μου, για την ανεκτίμητη στήριξη τους, τόσο στη διάρκεια των σπουδών μου, όσο και γενικότερα.

Ακόμα οφείλω ένα τεράστιο ευχαριστώ στη δεύτερη μου οικογένεια, που ήταν σταθερά δίπλα μου, αναδεικνύοντας διαρκώς κίνητρα και υπενθυμίζοντας μου ότι κάτι αλλάζει.

Τέλος οφείλω ένα ευχαριστώ σε όλους τους κοντινούς μου ανθρώπους για την υποστήριξη τους.

Περίληψη

Η παρούσα διπλωματική εργασία εκπονήθηκε στα πλαίσια του προπτυχιακού κύκλου σπουδών της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβιου Πολυτεχνείου.

Σκοπός της εργασίας είναι η παρουσίαση, τόσο θεωρητικά όσο και προγραμματιστικά, υπολογιστικών μεθόδων στη στατιστική, οι οποίες είναι χρήσιμες σε ελέγχους υποθέσεων. Αρχικά υπενθυμίζονται βασικά στοιχεία των ελέγχων υποθέσεων, αλλά και μέθοδοι πραγματοποίησης τους, βασισμένες στην κλασσική στατιστική. Έπειτα παρουσιάζονται οι Έλεγχοι Τυχαιοποίησης, οι Έλεγχοι Monte Carlo και οι Έλεγχοι Bootstrap. Στο θεωρητικό μέρος αναγράφονται η μαθηματική βάση των μεθόδων, τα αλγοριθμικά βήματα τους, προβλήματα στην εφαρμογή τους και ενδιαφέρουσες περιπτώσεις που εφαρμόζονται.

Στο προγραμματιστικό, και τελευταίο μέρος, υλοποιούνται οι μέθοδοι, με τη χρήση της γλώσσας προγραμματισμού R. Παρουσιάζονται συγκεκριμένα είδη των ελέγχων υποθέσεων, βασισμένα σε πλασματικά δεδομένα, εφαρμόζονται σε αυτά όποιες μέθοδοι είναι κατάλληλες και προκύπτουν συμπεράσματα σχετικά με τις αποφάσεις που οδηγούν οι μέθοδοι αλλά και με τις διαφοροποιήσεις των αποφάσεων.

Λέξεις Κλειδιά

Έλεγχοι Τυχαιοποίησης, Monte Carlo, Bootstrap, R, Έλεγχοι Υποθέσεων, Υπολογιστική Στατιστική

Abstract

This thesis was conducted as part of the undergraduate program of the School of Applied Mathematical and Physical Sciences at the National Technical University of Athens.

The aim of this thesis is to present, both theoretically and programmatically, computational methods in statistics that are useful for hypothesis testing. Initially, basic elements of hypothesis testing are recalled, along with methods for conducting them based on classical statistics. Then, Randomization Tests, Monte Carlo Tests, and Bootstrap Tests are presented. In the theoretical part, the mathematical basis of the methods, their algorithmic steps, problems in their application, and interesting cases where they are applied are described.

In the programming and final part, the methods are implemented using the R programming language. Specific types of hypothesis tests are presented, based on simulated data, appropriate methods are applied to them, and conclusions are drawn regarding the decisions that the methods lead to and the variations in these decisions.

Keywords

Randomization tests, Monte Carlo, Bootstrap, R, Hypothesis testing, Computational Statistics

Περιεχόμενα

Ευχαριστίες	iii
Περίληψη	v
Abstract	vii
Περιεχόμενα	xi
Κατάλογος Σχημάτων	xiii
Κατάλογος Πινάκων	xv
Κατάλογος Κωδίκων	xviii
1 Εισαγωγή	1
1.1 Σκοπός	1
1.2 Δομή	2
2 Παραμετρικοί Έλεγχοι Υποθέσεων	5
2.1 Εισαγωγή	5
2.2 Είδη Υποθέσεων και Ελέγχων	5
2.3 Συνάρτηση Στατιστικής Ελέγχου, Κρίσιμη Περιοχή και Επίπεδο Σημαντικότητας	6
2.4 Είδη Σφαλμάτων και Ισχύς	8
2.5 Παρατηρούμενο Επίπεδο Σημαντικότητας	9
2.6 Ενδιαφέρουσες Περιπτώσεις	9
2.6.1 Έλεγχοι Υποθέσεων για την Μέση Τιμή	9
2.6.2 Έλεγχος Υποθέσεων για τον Συντελεστή Συσχέτισης - Ο Συντελεστής Γραμμικής Συσχέτισης Pearson	15
2.7 Συμπεράσματα	16

3 Μη Παραμετρικοί Έλεγχοι Υποθέσεων	17
3.1 Εισαγωγή	17
3.2 Έλεγχοι για Δύο Πληθυσμούς	17
3.2.1 Ο Έλεγχος Wilcoxon Rank-Sum	17
3.3 Έλεγχος για Τρεις ή Περισσότερους Πληθυσμούς	19
3.3.1 Έλεγχος Kruskal-Wallis	19
3.3.2 Έλεγχος Friedman-Rank	20
3.4 Έλεγχος για Ύπαρξη Συσχέτισης	20
3.4.1 Ο Συντελεστής Συσχέτισης Spearman	20
3.4.2 Ο Συντελεστής Συσχέτισης Pearson χ^2	21
3.5 Συμπεράσματα	23
4 Έλεγχοι Τυχαιοποίησης	25
4.1 Εισαγωγή	25
4.2 Ακριβείς Έλεγχοι Τυχαιοποίησης	25
4.3 Προσεγγιστικοί Έλεγχοι Τυχαιοποίησης	27
4.3.1 Εκτίμηση p-value	27
4.4 Ενδιαφέρουσες Περιπτώσεις	28
4.4.1 Έλεγχος για τη Μέση Τιμή Δύο Πληθυσμών	28
4.4.2 Συσχέτιση Δύο Μεταβλητών	29
4.4.3 Έλεγχος Διακύμανσης ANOVA	30
4.4.4 Ακριβής Έλεγχος του Fisher για Πίνακες Συνάφειας	31
4.5 Συμπεράσματα	33
5 Έλεγχοι Monte Carlo	35
5.1 Εισαγωγή	35
5.2 Περιγραφή Ελέγχου	35
5.3 Ενδιαφέροντα Παραδείγματα	36
5.3.1 Έλεγχος για τη Μέση Τιμή Ενός Πληθυσμού	36
5.3.2 Monte Carlo Καλής Προσαρμογής	37
5.4 Συμπεράσματα	38
6 Bootstrap	41
6.1 Εισαγωγή	41
6.2 Περιγραφή Ελέγχου	41
6.3 Προβλήματα της Μεθόδου	42
6.4 Ενδιαφέρουσες Περιπτώσεις	43
6.4.1 Έλεγχος της Μέσης Τιμής Ενός Πληθυσμού	43

6.4.2 Έλεγχος για Γραμμική Συσχέτιση Δύο Μεταβλητών	44
6.5 Συμπεράσματα	45
7 Εφαρμογές στην R	47
7.1 Εισαγωγή	47
7.2 Έλεγχος για τη Μέση Τιμή Ενός Πληθυσμού	47
7.3 Έλεγχος της Διαφοράς των Μέσων Δύο Πληθυσμών	54
7.4 Έλεγχος της Διαφοράς των Μέσων Τριών Πληθυσμών	60
7.5 Γραμμική Συσχέτιση	63
7.6 Συσχέτιση σε Πίνακες Συνάφειας	67
7.7 Έλεγχος Καλής Προσαρμογής	71
7.8 Συμπεράσματα	75
8 Συμπεράσματα	77

Κατάλογος Σχημάτων

2.1	Κρίσιμη περιοχή για μονόπλευρο και αμφίπλευρο έλεγχο υποθέσεων .	7
7.1	Απεικόνιση των επειγόντων περιστατικών και επισήμανση της μέσης τιμής	49
7.2	Ιστόγραμμα των τιμών της συνάρτησης ελέγχου από την προσομοίωση δειγμάτων με Monte Carlo για 1 πληθυσμό	51
7.3	Ιστόγραμμα των τιμών της συνάρτησης ελέγχου από την προσομοίωση δειγμάτων με τη μέθοδο Bootstrap για 1 πληθυσμό	53
7.4	Απεικόνιση των δεδομένων των δύο φαρμάκων	55
7.5	Ιστόγραμμα των τιμών της συνάρτησης ελέγχου με τη μέθοδο της τυ- χαιοποίησης για 2 πληθυσμούς	58
7.6	Ιστόγραμμα των τιμών της συνάρτησης ελέγχου με τη μέθοδο Bootstrap για 2 πληθυσμούς	59
7.7	Απεικόνιση των δεδομένων των τριών τύπων κυττάρων	61
7.8	Διάγραμμα Συστολικής - Διαστολικής Πίεσης	64
7.9	Διάγραμμα συσχέτισης Καπνίσματος - Καρκίνου του Πνεύμονα	68
7.10	Ιστόγραμμα των τιμών της συνάρτησης ελέγχου για τους διαφορετικούς δυνατούς πίνακες συνάφειας με τη μέθοδο Fisher	71
7.11	Παρατηρούμενες και Αναμενόμενες Τιμές	73
7.12	Ιστόγραμμα των τιμών της συνάρτησης ελέγχου της μεθόδου Monte Carlo καλής προσαρμογής	75

Κατάλογος Πινάκων

2.1	Είδη Σφαλμάτων και Αποφάσεων	8
3.1	Πίνακας συνάφειας $r \times c$	22
4.1	Πίνακας συνάφειας 2×2	31
7.1	Αριθμός επειγόντων περιστατικών ενός νοσοκομείου σε διάστημα 100 ημερών	48
7.2	Σύνοψη αποτελεσμάτων για τον έλεγχο της μέσης τιμής ενός πληθυσμού	53
7.3	Ποσοστό εξάλειψης καρκινικών κυττάρων για δύο τύπους φαρμάκων .	54
7.4	Σύνοψη αποτελεσμάτων για τον έλεγχο της διαφοράς των μέσων τιμών δύο πληθυσμών	59
7.5	Ποσοστό απορρόφησης ακτινοβολίας για τρεις τύπους κυττάρων	60
7.6	Σύνοψη αποτελεσμάτων για τον έλεγχο της διαφοράς μεταξύ τριών πληθυσμών	63
7.7	Μετρήσεις Συστολικής και Διαστολικής Πίεσης σε διάστημα 7 ημερών .	63
7.8	Σύνοψη αποτελεσμάτων για τον έλεγχο της γραμμικής συσχέτισης . . .	67
7.9	Δεδομένα συσχέτισης καπνίσματος - καρκίνου του πνεύμονα	67
7.10	Διαφορετική μορφή δεδομένων συσχέτισης καπνίσματος - καρκίνου του πνεύμονα	69
7.11	Συχνότητα εμφάνισης κάθε δυνατού πίνακα συνάφειας και τιμή Pearson	70
7.12	Αριθμός επειγόντων περιστατικών ενός νοσοκομείου σε διάστημα 100 ημερών	72

Κατάλογος Κωδίκων

7.1	Εγκατάσταση βιβλιοθήκης BSDA	48
7.2	Ενσωμάτωση της BSDA στον κώδικα	48
7.3	Εισαγωγή δεδομένων επειγόντων περιστατικών και υπολογισμός χαρακτηριστικών μεγεθών του δείγματος	49
7.4	Παραμετρικός έλεγχος για έναν πληθυσμό	50
7.5	Δημιουργία 99 δειγμάτων από την Poisson	50
7.6	Υπολογισμός T και p-value για τα δείγματα της Poisson	51
7.7	Bootstrap για τη μέση τιμή ενός πληθυσμού	52
7.8	Υπολογισμός p-value για τα bootstrap δείγματα για τη μέση τιμή ενός πληθυσμού	52
7.9	Εισαγωγή δεδομένων φαρμάκων	55
7.10	Παραμετρικός έλεγχος δύο πληθυσμών	56
7.11	Μη παραμετρικός έλεγχος δύο πληθυσμών	56
7.12	Έλεγχος τυχαιοποίησης δύο πληθυσμών και υπολογισμός p-value	57
7.13	Έλεγχος Bootstrap δύο πληθυσμών και υπολογισμός p-value	58
7.14	Εισαγωγή δεδομένων κυττάρων και υπολογισμός χαρακτηριστικών μεγεθών των δειγμάτων	61
7.15	Παραμετρικός έλεγχος τριών πληθυσμών	62
7.16	Μη παραμετρικός έλεγχος τριών πληθυσμών	62
7.17	Έλεγχος τυχαιοποίησης για τρεις πληθυσμούς	62
7.18	Εισαγωγή δεδομένων συστολικής και διαστολικής πίεσης	64
7.19	Υπολογισμός συντελεστή συσχέτισης Pearson	64
7.20	Υπολογισμός συντελεστή συσχέτισης Spearman	65
7.21	Έλεγχος τυχαιοποίησης για γραμμική συσχέτιση	65
7.22	Έλεγχος bootstrap για γραμμική συσχέτιση	66
7.23	Υπολογισμός p-value των δειγμάτων bootstrap για γραμμική συσχέτιση	66
7.24	Εισαγωγή δεδομένων για τη συσχέτιση καρκίνου-καπνίσματος με τη μορφή πίνακα συνάφειας	68
7.25	Υπολογισμός συντελεστή συσχέτισης Pearson χ^2	69

7.26 Δημιουργία των δυνατών συνδυασμών του πίνακα συνάφειας	69
7.27 Εισαγωγή δεδομένων για έλεγχο κατανομής και υπολογισμός χαρακτηριστικών μεγεθών του δείγματος	72
7.28 Δημιουργία συνάρτησης για προσομοίωση απο την Poisson	74
7.29 Δημιουργία προσομοιωμένων δειγμάτων βάσει του δοσμένου, υπολογισμός της T και του p -value	74

Κεφάλαιο 1

Εισαγωγή

Η εξέλιξη της τεχνολογίας, η αναπτυξη των υπολογιστών και η ευρεία χρήση τους έχουν δημιουργήσει τις προϋποθέσεις για την ανάπτυξη υπολογιστικά πολυπλοκότερων στατιστικών μεθόδων καθώς και την ικανότητα οπτικοποίησης των δεδομένων. Τα δεδομένα είναι ολοένα περισσότερα και πιο σύνθετα και οι κατανομές των πληθυσμών από τους οποίους προέρχονται είναι συνήθως άγνωστες.

Η ανάλυση των δεδομένων σήμερα είναι απαραίτητη για τη διεξαγωγή συμπερασμάτων, χρήσιμα σε ολόκληρη την κοινωνία. Για παράδειγμα, μέσω ιατρικών και κλινικών δεδομένων είναι εφικτό να προκύψουν συμπεράσματα σχετικά με τα χαρακτηριστικά ιών, φαρμάκων και θεραπειών και κατ' επέκταση να εφαρμοστούν αυτά στην πρόληψη και τη βελτίωση του συνόλου των ανθρώπων.

Η Ιατρική είναι μόνο ένας από τους τομείς, που οι συγκεκριμένες μέθοδοι έχουν εφαρμογή. Άλλοι είναι το Περιβάλλον, η Ψυχολογία, η Βιομηχανική Παραγωγή.

Παρόλ' αυτά, αν και οι συγκεκριμένες μέθοδοι έχουν ευρεία και κυρίως αναγκαία χρήση, σήμερα η χρήση τους είναι περιορισμένη, αφού προτεραιότητα είναι η αξιολόγηση του κόστους και όχι οι ανάγκες. Αυτό όμως δεν αναιρεί το γεγονός ότι η περαιτέρω έρευνα και ανάπτυξη των μεθόδων αυτών σε περισσότερα πεδία εφαρμογών είναι απαραίτητη.

1.1 Σκοπός

Σκοπός της παρούσας διπλωματικής είναι η παράθεση τριών μεθόδων στατιστικής που εφαρμόζονται σε ελέγχους υποθέσεων, των ελέγχων τυχαιοποίησης, της μεθόδου Monte Carlo και της μεθόδου Bootstrap, κοινό στοιχείο των οποίων είναι η ανάγκη

χρήσης υπολογιστών.

Η παρουσίαση των μεθόδων γίνεται τόσο θεωρητικά, όσο και με εφαρμογές. Σκοπός του θεωρητικού τμήματος είναι η παράθεση του μαθηματικού υποβάθρου και των βημάτων των μεθόδων, των πλεονεκτημάτων και των προβλημάτων κάθε μιας. Το κομμάτι των εφαρμογών εξυπηρετεί την καλύτερη εποπτεία σχετικά με την υλοποίηση των μεθόδων προγραμματιστικά.

1.2 Δομή

Η παρούσα διπλωματική εργασία αποτελείται από 7 Κεφάλαια.

Το παρών **Κεφάλαιο 1** αποτελεί μία εισαγωγή της εργασίας, και συγκεκριμένα περιγράφει το σκοπό και τη δομή της.

Στο **Κεφάλαιο 2** ορίζονται, αρχικά, βασικά στοιχεία των ελέγχων υποθέσεων, δηλαδή τα είδη των υποθέσεων, η συνάρτηση ελέγχου, η κρίσιμη περιοχή, το επίπεδο σημαντικότητας, τα είδη σφαλμάτων, η ισχύς και το παρατηρούμενο επίπεδο σημαντικότητας. Στη συνέχεια παρουσιάζονται κάποιοι παραμετρικοί έλεγχοι, τόσο για τη μέση τιμή, όσο και για τον συντελεστή συσχέτισης, έλεγχοι δηλαδή που χρήζουν προϋποθέσεων, κυρίως περί κανονικότητας του πληθυσμού ή των πληθυσμών που μελετούνται, προκειμένου να υλοποιηθούν.

Στο **Κεφάλαιο 3** παρουσιάζονται μη παραμετρικοί έλεγχοι υποθέσεων για τη μέση τιμή και τον συντελεστή συσχέτισης, έλεγχοι δηλαδή που δεν χρήζουν προϋποθέσεων, κυρίως περί κανονικότητας, αλλά βασίζονται στην τάξη μεγέθους των παρατηρήσεων του δείγματος.

Στο **Κεφάλαιο 4** παρουσιάζονται οι έλεγχοι τυχαιοποίησης, τόσο τα βήματα με τα οποία πραγματοποιούνται γενικά, όσο και η εφαρμογή τους σε συγκεκριμένα είδη ελέγχων. Οι έλεγχοι αυτοί δεν χρήζουν καμίας υπόθεσης και βασίζονται πλήρως στο δείγμα.

Στο **Κεφάλαιο 5** παρουσιάζεται η μέθοδος Monte Carlo, η οποία χρησιμοποιείται για την προσομοίωση δεδομένων και την εξαγωγή στατιστικών συμπερασμάτων βασισμένη στη γνώση της κατανομής που ακολουθεί ο πληθυσμός ή οι πληθυσμοί που μελετώνται.

Στο **Κεφάλαιο 6** παρουσιάζεται η μέθοδος Bootstrap, η οποία για την εξαγωγή στατιστικών συμπερασμάτων βασίζεται αποκλειστικά και μόνο στο δείγμα, και στη τυχαία δειγματοληψία με επανάθεση, από αυτό, για τη δημιουργία νέων δειγμάτων.

Τέλος, στο **Κεφάλαιο 7**, υλοποιούνται όλες οι παραπάνω μέθοδοι με τη χρήση της

γλώσσας προγραμματισμού R, η οποία είναι κατάλληλη για τη διεξαγωγή στατιστικής ανάλυσης, σε ειδικό περιβάλλον, το RStudio. Στο συγκεκριμένο κεφάλαιο παρατίθενται 7 παραδείγματα με συγκεκριμένα σύνολα δεδομένων, ορίζονται οι υποθέσεις βάσει του ζητούμενου του προβλήματος, πραγματοποιούνται οι απαραίτητοι έλεγχοι για τη διεξαγωγή συμπερασμάτων και παρουσιάζονται τα κατάλληλα διαγράμματα.

Κεφάλαιο 2

Παραμετρικοί Έλεγχοι Υποθέσεων

2.1 Εισαγωγή

Οι έλεγχοι υποθέσεων χρησιμοποιούνται για την επιβεβαίωση μιας στατιστικής υπόθεσης H , δηλαδή κάθε υπόθεσης που αφορά την κατανομή $p(x|\theta)$ μιας τυχαίας μεταβλητής (τ.μ.) X . Αυτή μπορεί να αφορά είτε την άγνωστη παράμετρο θ της συνάρτησης πυκνότητας πιθανότητας (σ.π.π.) $p(x|\theta)$ (όταν η τ.μ. είναι συνεχής) ή την συνάρτηση μάζας πιθανότητας (σ.μ.π.) $p(x|\theta)$ (όταν η τ.μ. είναι διακριτή), όταν η p έχει γνωστή συναρτησιακή μορφή, είτε την συναρτησιακή μορφή της p (Φουσκάκης, 2009). Οι παραμετρικοί έλεγχοι εξετάζουν την τιμή μιας παραμέτρου βασιζόμενοι στην κατανομή της, ενώ οι μη παραμετρικοί έλεγχοι, που θα αναφερθούν στο επόμενο κεφάλαιο, βασίζονται σε άλλες ιδιότητες της παραμέτρου. Η επιλογή του είδους του ελέγχου σχετίζεται με την παράμετρο που χρειάζεται να ελεγχθεί, το μέγεθος του δείγματος και στην περίπτωση περισσότερων δειγμάτων την ανεξαρτησία των παρατηρήσεων.

2.2 Είδη Υποθέσεων και Ελέγχων

Προκειμένου να πραγματοποιηθεί ένας έλεγχος υποθέσεων θα πρέπει να οριστούν οι υποθέσεις του ελέγχου. Αρχικά ορίζεται η **μηδενική υπόθεση**, η οποία συμβολίζεται με H_0 , και είναι η υπόθεση η οποία εξετάζεται ως προς την αλήθεια της. Η υπόθεση H_0 θεωρείται αληθής μέχρι να υπάρξουν αρκετές ενδείξεις ώστε να απορριφθεί. Όταν είναι της μορφής $H_0: \theta = \theta_0$ η υπόθεση λέγεται **απλή**, ενώ όταν είναι της μορφής $H_0: \theta \leq \theta_0$ ή $H_0: \theta \geq \theta_0$ λέγεται **σύνθετη**.

Ορίζεται επίσης η **εναλλακτική υπόθεση** H_1 , η οποία ισχύει όταν απορρίπτεται η μηδενική υπόθεση. Εάν είναι της μορφής $H_1: \theta < \theta_0$ ή $H_1: \theta > \theta_0$ ο έλεγχος λέγεται **μονόπλευρος** (αριστερά και δεξιά αντίστοιχα), ενώ στην περίπτωση που η εναλλακτική υπόθεση είναι της μορφής $H_1: \theta \neq \theta_0$ τότε ο έλεγχος λέγεται **αμφίπλευρος**.

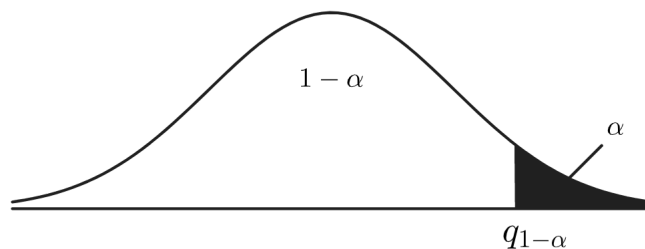
2.3 Συνάρτηση Στατιστικής Ελέγχου, Κρίσιμη Περιοχή και Επίπεδο Σημαντικότητας

Αφού οριστούν οι δύο υποθέσεις γίνεται διαχωρισμός του παραμετρικού χώρου (του χώρου των δυνατών τιμών της παραμέτρου θ) στην περιοχή που δεν απορρίπτεται η H_0 και ονομάζεται **περιοχή αποδοχής** και στην περιοχή που απορρίπτεται η H_0 και ονομάζεται **κρίσιμη περιοχή** (περιοχή απόρριψης), η οποία συμβολίζεται με R .

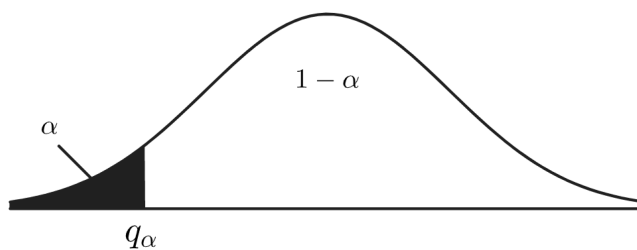
Για να πραγματοποιηθεί αυτός ο διαχωρισμός θα πρέπει να υπολογιστούν τα κρίσιμα σημεία. Αυτό μπορεί να είναι υπολογιστικά δύσκολο από την κατανομή της παραμέτρου θ . Είναι λοιπόν αναγκαίο να γίνει χρήση μιας άλλης συνάρτησης, η οποία εξαρτάται από την τιμή θ_0 της θ για την οποία η H_0 είναι αληθής. Η συνάρτηση αυτή ονομάζεται **στατιστική συνάρτηση ελέγχου** (ελεγχοσυνάρτηση) και συμβολίζεται με q . Η κατανομή της ελεγχοσυνάρτησης δείχνει την πυκνότητα πιθανότητας να πάρει αυτή κάποια τιμή αν η μηδενική υπόθεση είναι αληθής. Επομένως μεγάλες τιμές της ελεγχοσυνάρτησης ενισχύουν την ορθότητα της H_0 . Για να οριστεί η συνάρτηση αυτή και να βρεθεί η κατανομή της χρησιμοποιείται σαν δεδομένο στην αρχή του ελέγχου ότι ισχύει η H_0 .

Εκτός από την στατιστική συνάρτηση ελέγχου, η κρίσιμη περιοχή εξαρτάται και από μία πιθανότητα, η οποία ονομάζεται **επίπεδο σημαντικότητας** (ε.σ.) και συμβολίζεται με α . Το ε.σ. α δείχνει την πιθανότητα να απορριφθεί η H_0 ενώ αυτή είναι αληθής, όταν ο έλεγχος είναι απλός. Επομένως το ε.σ. καθορίζει το εύρος των δύο περιοχών. Με δεδομένο το ε.σ. μπορεί να οριστεί η κρίσιμη περιοχή.

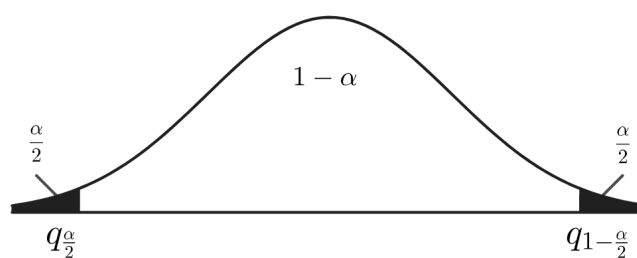
Υποθέτοντας για την ελεγχοσυνάρτηση γνωστή συμμετρική κατανομή, στην περίπτωση του μονόπλευρου ελέγχου ορίζεται ως $R = \{q|q < q_\alpha\}$ ή $R = \{q|q > q_{1-\alpha}\}$, ενώ στην περίπτωση του αμφίπλευρου ελέγχου ως $R = \{q|q < q_{\frac{\alpha}{2}} \vee q > q_{1-\frac{\alpha}{2}}\}$. Οι τιμές που διαχωρίζουν τις δύο περιοχές ονομάζονται **κρίσιμες τιμές**. Μπορεί λοιπόν να παρατηρηθεί ότι τόσο στον μονόπλευρο όσο και στον αμφίπλευρο έλεγχο η πιθανότητα απόρριψης της H_0 είναι συνολικά ίση με α . Στο Σχήμα 2.1 φαίνεται σκιαγραφημένη η περιοχή απόρριψης στον μονόπλευρο και στον αμφίπλευρο απλό έλεγχο.



(α) Κρίσιμη περιοχή (σκιαγραφημένη) για δεξιά μονόπλευρο έλεγχο υποθέσεων



(β) Κρίσιμη περιοχή (σκιαγραφημένη) για αριστερά μονόπλευρο έλεγχο υποθέσεων



(γ) Κρίσιμη περιοχή (σκιαγραφημένη) για αμφίπλευρο έλεγχο υποθέσεων

Σχήμα 2.1: Κρίσιμη περιοχή για μονόπλευρο και αμφίπλευρο έλεγχο υποθέσεων

Πραγματικότητα	Αποφάσεις	
	Μη απόρ. της H_0	Απόρ. της H_0
H_0 αληθής	Σωστή Απόφαση	Σφάλμα Τύπου I (πιθανότητα α)
H_1 αληθής	Σφάλμα Τύπου II (πιθανότητα β)	Σωστή Απόφαση

Πίνακας 2.1: Είδη Σφαλμάτων και Αποφάσεων

2.4 Είδη Σφαλμάτων και Ισχύς

Ανάλογα με την υπόθεση που είναι αληθής και την απόφαση που λαμβάνεται τελικά, ορίζονται δύο είδη σφαλμάτων. Αν η υπόθεση H_0 είναι αληθής αλλά απορριφθεί, το σφάλμα ονομάζεται **σφάλμα τύπου I**. Αν είναι αληθής η H_1 αλλά δεν απορριφθεί η H_0 , το σφάλμα ονομάζεται **σφάλμα τύπου II**.

Επομένως,

- $P(\text{σφάλμα τύπου I}) = P(\text{απόρ. } H_0 \mid H_0 \text{ αληθής})$
- $P(\text{σφάλμα τύπου II}) = P(\text{μη απόρ. } H_0 \mid H_1 \text{ αληθής})$

Στον Πίνακα 2.1 συνοψίζονται τα σφάλματα του ελέγχου ανάλογα με την απόφαση που λαμβάνεται και την υπόθεση που ισχύει πραγματικά.

Όταν η μηδενική υπόθεση είναι απλή τότε η πιθανότητα του σφάλματος τύπου I είναι ίση με το επίπεδο σημαντικότητας. Το επίπεδο σημαντικότητας υποδηλώνει δηλαδή το μέγεθος του ελέγχου. Η πιθανότητα του σφάλματος τύπου II συμβολίζεται με β . Ιδανικά, στην επιλογή της κρίσιμης περιοχής ελαχιστοποιούνται οι δύο αυτές πιθανότητες. Επειδή όμως η ταυτόχρονη ελαχιστοποίηση μπορεί να είναι δύσκολη, συνήθως επιλέγεται μια σταθερή τιμή για το α και ελαχιστοποιείται το β .

Ορίζεται επίσης η ποσότητα **1- β** , η οποία ονομάζεται **ισχύς** του ελέγχου. Πιο συγκεκριμένα, η ισχύς δείχνει την πιθανότητα να απορριφθεί η μηδενική υπόθεση H_0 ενώ ισχύει η εναλλακτική H_1 . Δηλαδή, $1-\beta = P(\text{απόρ. } H_0 \mid H_1 \text{ αληθής})$. Είναι προφανές ότι σε ελέγχους που ελαχιστοποιείται το β , μεγιστοποιείται η ισχύς του ελέγχου.

2.5 Παρατηρούμενο Επίπεδο Σημαντικότητας

Το **παρατηρούμενο επίπεδο σημαντικότητας** ή αλλιώς **p-value** δείχνει την πιθανότητα να παρατηρηθεί μια τιμή πιο ακραία από την παρατηρούμενη τιμή, δεδομένου ότι η μηδενική υπόθεση ισχύει. Επομένως, σε περιπτώσεις που δεν μπορεί να γίνει χρήση του επιπέδου σημαντικότητας α (π.χ. διακριτές κατανομές), υπολογίζεται το p-value και συγκρίνεται με το α . Εάν το p-value είναι μεγαλύτερο ή ίσο του α , τότε η μηδενική υπόθεση δεν απορρίπτεται. Είναι σημαντικό να σημειωθεί ότι το παρατηρούμενο επίπεδο σημαντικότητας δεν δείχνει την πιθανότητα η μηδενική υπόθεση να είναι αληθής.

2.6 Ενδιαφέρουσες Περιπτώσεις

2.6.1 Έλεγχοι Υποθέσεων για την Μέση Τιμή

Ο έλεγχος υποθέσεων για την μέση τιμή μπορεί να αφορά είτε τον έλεγχο της σε έναν πληθυσμό είτε την σύγκριση μέσων τιμών για δύο ή περισσότερους πληθυσμούς. Σε κάθε περίπτωση επιλέγεται η κατάλληλη ελεγχοσυνάρτηση ανάλογα με τις γνώσεις που υπάρχουν για τις διασπορές των πληθυσμών και το μέγεθος του τυχαίου δείγματος (τ.δ.) \mathbf{X} , γίνεται επιλογή των κρίσιμων σημείων ανάλογα με το είδος του ελέγχου (μονόπλευρος ή αμφίπλευρος) και λαμβάνεται η απόφαση απόρριψης ή μη της μηδενικής υπόθεσης ανάλογα με την τιμή της ελεγχοσυνάρτησης από τα δεδομένα του τυχαίου δείγματος.

Μέση Τιμή μιας Ποσοτικής Μεταβλητής ενός Πληθυσμού

Στους ελέγχους τη μέσης τιμής μιας ποσοτικής μεταβλητής ενός πληθυσμού η επιλογή της ελεγχοσυνάρτησης γίνεται ανάλογα με το αν η διασπορά σ^2 του πληθυσμού είναι γνωστή ή άγνωστη και αν το μέγεθος n του τ.δ. \mathbf{X} είναι μικρό ($n < 30$) ή μεγάλο ($n \geq 30$). Έστω ότι η μηδενική υπόθεση έχει την μορφή $H_0: \mu = \mu_0$, όπου μ η τιμή της μέσης τιμής για την οποία αυτή εξετάζεται. Τότε και η εναλλακτική υπόθεση θα είναι είτε της μορφής $H_1: \mu < \mu_0$ ή $H_1: \mu > \mu_0$ (μονόπλευρος έλεγχος) είτε $H_1: \mu \neq \mu_0$ (αμφίπλευρος έλεγχος). Θα εξεταστούν παρακάτω τρεις περιπτώσεις του ελέγχου. Σε κάθε περίπτωση θα πρέπει ο πληθυσμός από τον οποίο προέρχεται το τ.δ. \mathbf{X} να είναι κανονικά κατανομημένος ή το μέγεθος του δείγματος να είναι μεγάλο.

- **n μεγάλο, σ^2 γνωστή:**

Στην περίπτωση αυτή χρησιμοποιείται ως ελεγχοσυνάρτηση του ελέγχου η συ-

νάρτηση

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}, \quad (2.1)$$

όπου σ είναι η τυπική απόκλιση του πληθυσμού. Μια εκτίμηση της μέσης τιμής του πληθυσμού είναι η \bar{X} . Αφού ο πληθυσμός είναι κανονικά κατανομημένος με μέση τιμή μ και διασπορά σ^2 και το δείγμα μεγάλο, τότε από το Κεντρικό Οριακό Θεώρημα (Κ.Ο.Θ.) για την τ.μ. \bar{X} θα ισχύει $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$. Τυποποιώντας την μεταβλητή αυτή προκύπτει η τ.μ. Z για την οποία ισχύει ότι $Z \sim N(0, 1)$. Με βάση όσα αναφέρθηκαν παραπάνω για τις κρίσιμες περιοχές διακρίνουμε τις εξής περιπτώσεις:

- Αν $H_1: \mu > \mu_0 \rightarrow R: \{Z > Z_\alpha\}$
- Αν $H_1: \mu < \mu_0 \rightarrow R: \{Z < -Z_\alpha\}$
- Αν $H_1: \mu \neq \mu_0 \rightarrow R: \{|Z| > Z_{\frac{\alpha}{2}}\}$.

• **n μεγάλο, σ^2 άγνωστη:**

Στην περίπτωση που η διασπορά του πληθυσμού είναι άγνωστη ακολουθείται η ίδια διαδικασία με πριν με τη διαφορά ότι στην ελεγχουσυνάρτηση αντικαθίσταται η τυπική απόκλιση σ του πληθυσμού από την τυπική απόκλιση S του δείγματος. Η S υπολογίζεται από το τ.δ. \mathbf{X} και χρησιμοποιείται σαν εκτιμήτρια της σ , όταν αυτή είναι άγνωστη. Επομένως η ελεγχουσυνάρτηση έχει τη μορφή

$$Z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \quad (2.2)$$

και ακολουθεί την $N(0, 1)$. Οι κρίσιμες περιοχές είναι και πάλι οι:

- Αν $H_1: \mu > \mu_0 \rightarrow R: \{Z > Z_\alpha\}$
- Αν $H_1: \mu < \mu_0 \rightarrow R: \{Z < -Z_\alpha\}$
- Αν $H_1: \mu \neq \mu_0 \rightarrow R: \{|Z| > Z_{\frac{\alpha}{2}}\}$.

Η υλοποίηση του ελέγχου γίνεται στο υποκεφάλαιο 7.2.

• **n μικρό, σ^2 άγνωστη, κανονικός πληθυσμός από τον οποίο προέρχεται η \mathbf{X} :**

Στην περίπτωση που το τ.δ. είναι μικρό, δεν μπορεί να χρησιμοποιηθεί η προσέγγιση του Κ.Ο.Θ. Τότε χρησιμοποιείται σαν ελεγχουσυνάρτηση η

$$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}, \quad (2.3)$$

η οποία ακολουθεί την κατανομή t_{n-1} , όπου $n-1$ οι βαθμοί ελευθερίας της κατανομής. Συνεπώς οι κρίσιμες περιοχές διαμορφώνονται ως εξής:

- Αν $H_1: \mu > \mu_0 \rightarrow R: \{t > t_{n-1;a}\}$
- Αν $H_1: \mu < \mu_0 \rightarrow R: \{t < t_{n-1;a}\}$
- Αν $H_1: \mu \neq \mu_0 \rightarrow R: \{|t| > t_{n-1;\frac{\alpha}{2}}\}$.

Διαφορά των Μέσων Τιμών δύο Πληθυσμών

Έστω δύο πληθυσμοί κανονικά κατανομημένοι με μέσες τιμές μ_1, μ_2 και διασπορές σ_1^2, σ_2^2 , δύο τ.δ. $\mathbf{X}_1, \mathbf{X}_2$ και έστω ότι η H_0 είναι της μορφής $H_0: \mu_1 - \mu_2 = \delta$, όπου $\delta \in \mathbb{R}$. Ο έλεγχος διαφοροποιείται ανάλογα με την εξάρτηση των δύο δειγμάτων μεταξύ τους.

• Ανεξάρτητα Δείγματα

Όταν τα δύο τ.δ. είναι ανεξάρτητα και έχουν μεγάλο μέγεθος ή προέρχονται από κανονικούς πληθυσμούς τότε η διαφορά $\bar{X}_1 - \bar{X}_2$ είναι μια εκτίμηση της διαφοράς των μέσων τιμών και ακολουθεί κανονική κατανομή με μ.τ. $\mu_1 - \mu_2$ και διασπορά $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$. Όπως στους ελέγχους υποθέσεων για την μέση τιμή ενός πληθυσμού, έτσι και στους ελέγχους υποθέσεων για την διαφορά των μέσων τιμών δύο πληθυσμών, η επιλογή της ελεγχουσυνάρτησης διαφοροποιείται ανάλογα με τα μεγέθη n_1, n_2 των τ.δ. $\mathbf{X}_1, \mathbf{X}_2$, την γνώση των διασπορών σ_1^2, σ_2^2 και τη σχέση αυτών των δύο μεταξύ τους. Διακρίνονται οι εξής τρεις περιπτώσεις:

• n_1, n_2 μεγάλα, σ_1^2, σ_2^2 γνωστές ή άγνωστες:

Σε αυτή την περίπτωση δεν απαιτείται κανονικότητα των δύο πληθυσμών.

Η ελεγχουσυνάρτηση που επιλέγεται όταν είναι γνωστές οι διασπορές είναι η

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \quad (2.4)$$

ενώ όταν είναι άγνωστες οι διασπορές επιλέγεται η

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}. \quad (2.5)$$

Σε κάθε περίπτωση λόγω του Κ.Ο.Θ και της τυποποίησης της τ.μ. ισχύει $Z \sim N(0, 1)$. Επομένως οι περιοχές απόρριψης είναι οι:

- Αν $H_1: \mu_1 - \mu_2 > \delta \rightarrow R: \{Z > Z_\alpha\}$
- Αν $H_1: \mu_1 - \mu_2 < \delta \rightarrow R: \{Z < -Z_\alpha\}$
- Αν $H_1: \mu_1 - \mu_2 \neq \delta \rightarrow R: \{|Z| > Z_{\frac{\alpha}{2}}\}$.

- n_1, n_2 **μικρά**, σ_1^2, σ_2^2 **άγνωστες αλλά** $\sigma_1^2 = \sigma_2^2$:

Σε αυτή την περίπτωση δεν μπορεί να γίνει χρήση των παραπάνω προσεγγίσεων και απαιτείται κανονικότητα των δύο πληθυσμών. Συνεπώς χρησιμοποιείται σαν συνάρτηση ελέγχου η t . Επειδή οι διασπορές των πληθυσμών είναι άγνωστες, χρησιμοποιούνται στην ελεγχοσυνάρτηση οι διασπορές των τ.δ. S_1^2, S_2^2 σαν εκτιμήτριες των σ_1^2, σ_2^2 αντίστοιχα. Όμως ισχύει ότι $\sigma_1^2 = \sigma_2^2$, οπότε θα χρησιμοποιηθεί η σταθμισμένη εκτιμήτρια S_p^2 της διασποράς, όπου

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}. \quad (2.6)$$

Οπότε η συνάρτηση ελέγχου είναι η

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - \delta}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}. \quad (2.7)$$

Για την t ισχύει ότι $t \sim t_{n_1+n_2-2}$. Επομένως για τις περιοχές απόρριψης θα ισχύει:

- Αν $H_1: \mu_1 - \mu_2 > \delta \rightarrow R: \{t > t_{n_1+n_2-2; \alpha}\}$
- Αν $H_1: \mu_1 - \mu_2 < \delta \rightarrow R: \{t < -t_{n_1+n_2-2; \alpha}\}$
- Αν $H_1: \mu_1 - \mu_2 \neq \delta \rightarrow R: \{|t| > t_{n_1+n_2-2; \frac{\alpha}{2}}\}$.

- n_1, n_2 **μικρά**, σ_1^2, σ_2^2 **άγνωστες και** $\sigma_1^2 \neq \sigma_2^2$:

Στην περίπτωση που οι πληθυσμοί είναι κανονικοί, αλλά οι διασπορές τους δεν είναι ίσες, δεν χρησιμοποιείται η S_p^2 , αλλά οι επιμέρους εκτιμήτριες S_1^2, S_2^2 των διασπορών. Επομένως η συνάρτηση ελέγχου έχει τη μορφή

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - \delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}. \quad (2.8)$$

Όμως τώρα η t ακολουθεί προσεγγιστικά την t_m , όπου m είναι οι βαθμοί ελευθερίας της συνάρτησης και υπολογίζονται από τον τύπο

$$m = \frac{(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1-1} + \frac{(\frac{S_2^2}{n_2})^2}{n_2-1}}. \quad (2.9)$$

Οι περιοχές απόρριψης είναι οι:

- Αν $H_1: \mu_1 - \mu_2 > \delta \rightarrow R: \{t > t_{n; \alpha}\}$
- Αν $H_1: \mu_1 - \mu_2 < \delta \rightarrow R: \{t < -t_{n; \alpha}\}$
- Αν $H_1: \mu_1 - \mu_2 \neq \delta \rightarrow R: \{|t| > t_{n; \frac{\alpha}{2}}\}$.

Η υλοποίηση του ελέγχου γίνεται στο υποκεφάλαιο 7.3.

Διαφορά των Μέσων Τιμών άνω των δύο Ανεξάρτητων Πληθυσμών - Η μέθοδος ANOVA

Η μέθοδος ANOVA είναι μία παραμετρική μέθοδος που χρησιμοποιείται στους ελέγχους υποθέσεων για τη σύγκριση των μέσων τιμών 3 ή περισσότερων ανεξάρτητων πληθυσμών. Έστω ότι χρειάζεται να εξετασθεί αν οι μέσες τιμές k ανεξάρτητων πληθυσμών διαφέρουν ως προς μία κατηγορική μεταβλητή, η οποία ονομάζεται **παράγοντας**. Οι πληθυσμοί αποτελούν τις δυνατές τιμές που μπορεί να λάβει ο παράγοντας και ονομάζονται **επίπεδα**. Έστω ότι έχουν ληφθεί k τυχαία δείγματα, μεγέθους n_i , $i = 1, \dots, k$ από κάθε πληθυσμό, με αντίστοιχες μέσες τιμές μ_i , $i = 1, \dots, k$.

Ο έλεγχος υποθέσεων θα είναι:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_1: \mu_i \neq \mu_j, i \neq j, i = 1, \dots, k, j = 1, \dots, k$$

Η μηδενική υπόθεση υποδηλώνει την απουσία διαφοράς των μέσων τιμών μεταξύ των επιπέδων. Υποδηλώνει δηλαδή ότι ο παράγοντας δεν επηρεάζει τις μέσες τιμές των πληθυσμών.

Επειδή ο συγκεκριμένος έλεγχος είναι παραμετρικός θα πρέπει να ισχύουν οι παρακάτω **υποθέσεις**:

- Τα δείγματα πρέπει να είναι ανεξάρτητα μεταξύ τους. Δεν μπορεί κάποιο υποκείμενο να ανήκει σε 2 ή περισσότερους πληθυσμούς.
- Οι πληθυσμοί πρέπει να είναι κανονικά κατανομημένοι. Αυτό μπορεί να ελεγχθεί για παράδειγμα με την κατασκευή ιστογράμματος των παρατηρήσεων κάθε πληθυσμού.
- Οι τυπικές αποκλίσεις των πληθυσμών θα πρέπει να είναι ίσες.

Σημειώνεται ότι ο έλεγχος ANOVA είναι ανθεκτικός σε μέτρια παραβίαση των υποθέσεων περί κανονικότητας και ίσων αποκλίσεων.

Σημαντικά μεγέθη για την εκτέλεση του ελέγχου είναι τα εξής:

- **Διακύμανση εντός του κάθε πληθυσμού:**

$$SS(W) = \sum_{i=1}^k df_i S_i^2, \quad (2.10)$$

όπου $df_i = n_i - 1$ και S_i^2 η διακύμανση του κάθε δείγματος. Αυτή η διακύμανση δεν επηρεάζεται από την επίδραση του παράγοντα ενδιαφέροντος.

- Διακύμανση μεταξύ των πληθυσμών:

$$SS(B) = \sum_{i=1}^k n_i(\bar{x}_i - \bar{x})^2, \quad (2.11)$$

όπου \bar{x}_i : η μέση τιμή κάθε i δείγματος και \bar{x} η συνολική μέση τιμή όλων των παρατηρήσεων, όπως θα υπολογιζόταν αν δεν υπήρχαν οι παράγοντες. Η διακύμανση αυτή επηρεάζεται από την επίδραση του παράγοντα.

- Βαθμοί ελευθερίας εντός κάθε πληθυσμού:

$$df(W) = \sum_{i=1}^k (n_i - 1) \quad (2.12)$$

- Βαθμοί ελευθερίας μεταξύ των πληθυσμών:

$$df(B) = k - 1 \quad (2.13)$$

- Σταθμισμένο άθροισμα τετραγώνων:

$$MS = \frac{SS}{df}. \quad (2.14)$$

Επομένως

$$MS(W) = \frac{SS(W)}{df(W)} \quad (2.15)$$

και

$$MS(B) = \frac{SS(B)}{df(B)}. \quad (2.16)$$

Η συνάρτηση ελέγχου που χρησιμοποιείται είναι η

$$F = \frac{MS(B)}{MS(W)}, \quad (2.17)$$

η οποία ακολουθεί την $F_{(k-1, n-k)}$. Από τους γνωστούς πίνακες της κατανομής υπολογίζεται το p-value, αφού $p\text{-value} = P(F_{(k-1, n-k)} > F_{obs})$, όπου F_{obs} η τιμή της ελεγχουσυνάρτησης από το δείγμα.

Η μηδενική υπόθεση απορρίπτεται όταν $p\text{-value} < \alpha$.

Είναι σημαντικό να αναφερθεί ότι στην περίπτωση που απορριφθεί η μηδενική υπόθεση, ο έλεγχος υποδηλώνει ότι υπάρχει διαφορά μεταξύ των πληθυσμών, χωρίς όμως να είναι σε θέση να υποδείξει μεταξύ ποιων πληθυσμών.

Στην περίπτωση που δεν πληρούνται οι προϋποθέσεις για τη χρήση του παραμετρικού ελέγχου, μπορεί να γίνει χρήση του μη παραμετρικού, όπως θα αναφερθεί στο Κεφάλαιο 3.

Η υλοποίηση της μεθόδου ANOVA γίνεται στο υποκεφάλαιο 7.4.

2.6.2 Έλεγχος Υποθέσεων για τον Συντελεστή Συσχέτισης - Ο Συντελεστής Γραμμικής Συσχέτισης Pearson

Πολλές φορές είναι χρήσιμη η γνώση για την ύπαρξη συσχέτισης δύο ή περισσότερων ποσοτικών χαρακτηριστικών ενός τυχαίου δείγματος, και κατ' επέκταση των πληθυσμών από τους οποίους προέρχονται. Με τον όρο **συσχέτιση** περιγράφεται ο τρόπος με τον οποίο η μεταβολή ενός χαρακτηριστικού θα επηρεάσει γραμμικά τη μεταβολή κάποιου άλλου χαρακτηριστικού. Η συσχέτιση αυτή μπορεί να ποσοτικοποιηθεί και να ερμηνευτεί με τον υπολογισμό διάφορων μέτρων.

Ο **συντελεστής γραμμικής συσχέτισης Pearson**, συμβολίζεται με r και χρησιμοποιείται για τη διερεύνηση γραμμικής σχέσης μεταξύ δύο ποσοτικών μεταβλητών. Γραμμική σχέση σημαίνει πως η μεταβολή της μιας μεταβλητής θα επηρεάσει αναλογικά την άλλη. Ο έλεγχος υποθέσεων για τον συντελεστή Pearson είναι ένας παραμετρικός έλεγχος. Έστω ένα τυχαίο δείγμα μεγέθους n και x_i, y_i οι παρατηρήσεις για τις μεταβλητές X, Y αντίστοιχα. Αρχικά υπολογίζεται ο δειγματικός συντελεστής

$$r = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{\sqrt{[n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2][n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2]}}. \quad (2.18)$$

Σαν μηδενική υπόθεση ορίζεται η απουσία συσχέτισης. Η μηδενική υπόθεση, λοιπόν, θα είναι της μορφής $H_0: r = 0$ και η εναλλακτική θα είναι $H_1: r \neq 0$. Αυτό προκύπτει από το γεγονός ότι ο συντελεστής Pearson όταν είναι 0 υποδηλώνει την απουσία συσχέτισης, ενώ οι τιμές κοντά στο 1 και το -1 υποδηλώνουν θετική και αρνητική συσχέτιση αντίστοιχα.

Ως συνάρτηση ελέγχου χρησιμοποιείται η

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}, \quad (2.19)$$

όπου r είναι ο δειγματικός συντελεστής, ο οποίος κάτω από την H_0 ακολουθεί την κατανομή $t_{n-2, \alpha}$. Η κρίσιμη περιοχή είναι $R = t > t_{n-2, \alpha}$. Και τελικά αν $t >$ κρίσιμη τιμή, τότε απορρίπτεται η H_0 .

Επειδή ο παραπάνω έλεγχος είναι παραμετρικός, για να γίνει η χρήση του πρέπει να ικανοποιούνται οι παρακάτω προϋποθέσεις:

- Όλες οι μεταβλητές πρέπει να είναι ποσοτικές.
- Οι μεταβλητές πρέπει να ακολουθούν κανονική κατανομή.

Η εφαρμογή του ελέγχου του συντελεστή Pearson γίνεται στο υποκεφάλαιο 7.5.

2.7 Συμπεράσματα

Συνοψίζοντας, οι παραμετρικοί έλεγχοι υποθέσεων χρησιμοποιούνται όταν χρειάζεται να ελεγχθεί η τιμή μιας παραμέτρου ενός πληθυσμού ή να συγκριθούν δύο τιμές αντίστοιχων παραμέτρων δύο πληθυσμών. Συνοπτικά τα βήματα του ελέγχου υποθέσεων είναι τα εξής:

- Ορισμός της μηδενικής υπόθεσης H_0
- Ορισμός της εναλλακτικής υπόθεσης H_1
- Επιλογή της κατάλληλης ελεγκοσυνάρτησης
- Υπολογισμός της τιμής της ελεγκοσυνάρτησης από τα δεδομένα του τυχαίου δείγματος
- Υπολογισμός της κρίσιμης τιμής χρησιμοποιώντας το επίπεδο σημαντικότητας α
- Απόφαση για απόρριψη ή μη της H_0 .

Σε αυτό το κεφάλαιο αναφέρθηκαν έλεγχοι για συγκεκριμένες παραμέτρους βασισμένοι σε υποθέσεις για την κατανομή της. Στο επόμενο κεφάλαιο θα αναφερθούν οι μη παραμετρικοί έλεγχοι υποθέσεων, οι έλεγχοι δηλαδή που χρησιμοποιούνται για τη λήψη μιας απόφασης βασισμένοι σε άλλες ιδιότητες της παραμέτρου και όχι στην κατανομή της.

Κεφάλαιο 3

Μη Παραμετρικοί Έλεγχοι Υποθέσεων

3.1 Εισαγωγή

Οι μη παραμετρικοί έλεγχοι υποθέσεων χρησιμοποιούνται σε περιπτώσεις όπου δεν υπάρχουν οι συνθήκες για να χρησιμοποιηθούν οι παραμετρικοί έλεγχοι (πχ. μη κανονικότητα των πληθυσμών, μικρά δείγματα). Γενικά αν μπορεί να γίνει χρήση των παραμετρικών ελέγχων αυτή προτιμάται. Παρόλ' αυτά, ακόμα και σε περιπτώσεις με πολύ μικρά δείγματα μπορεί να προκύψει μία αρκετά σημαντική στατιστική απόφαση. Ως επί τω πλείστον στους μη παραμετρικούς ελέγχους εξετάζεται αν δύο ή περισσότεροι πληθυσμοί ακολουθούν ίδια κατανομή. Αν είναι γνωστό ότι οι πληθυσμοί ακολουθούν ίδια κατανομή, μπορεί η στατιστική υπόθεση να εξειδικευτεί στον έλεγχο ισότητας των κεντρικών θέσεων των πληθυσμών. Ως κεντρικές θέσεις χρησιμοποιούνται σε ποσοτικά δεδομένα η διάμεσος ή η μέση τιμή, ενώ σε διατακτικά δεδομένα μόνο η διάμεσος. Αξίζει να σημειωθεί ότι οι μη παραμετρικοί έλεγχοι έχουν μικρότερη ισχύ από τους παραμετρικούς.

3.2 Έλεγχοι για Δύο Πληθυσμούς

3.2.1 Ο Έλεγχος Wilcoxon Rank-Sum

Ο έλεγχος Wilcoxon Rank-Sum βασίζεται στη διάταξη των παρατηρήσεων. Έστω δύο πληθυσμοί A και B και έστω δύο τυχαία δείγματα από τους πληθυσμούς A και B μεγέθους n_A και n_B αντίστοιχα. Εξετάζεται η μηδενική υπόθεση H_0 : «οι δύο πληθυσμοί έχουν ίδιες κατανομές», ή συμβολικά $H_0: A = B$. Η εναλλακτική

υπόθεση μπορεί να είναι $H_1: A > B$ (μετακίνηση της A προς τα δεξιά), ή $H_1: A < B$ (μετακίνηση της A προς τα αριστερά) ή $H_1: A \neq B$ (μετακίνηση της A είτε προς τα αριστερά είτε προς τα δεξιά).

Για την πραγματοποίηση του ελέγχου χρειάζεται να βαθμολογηθούν τα δεδομένα ανάλογα με την θέση τους, αν αυτά τοποθετηθούν σε αύξουσα σειρά. Δηλαδή στην μικρότερη τιμή ανατίθεται η τιμή 1, στην αμέσως επόμενη η τιμή 2 κ.ο.κ. Σε περίπτωση ισοβαθμίας η τιμή που παίρνουν τα δεδομένα είναι ο μέσος όρος των τιμών που θα έπαιρναν αυτά αν δεν ήταν ίσα. Σε αυτή την περίπτωση, αν έχει δοθεί για παράδειγμα η τιμή 3.5 σε δύο ισόπαλες τιμές τότε η αρίθμηση συνεχίζει από το 5 (θεωρείται δηλαδή ότι έχει δοθεί και το 3 και το 4).

Στον έλεγχο ορίζεται η μεταβλητή W_A , η οποία είναι μια τυχαία μεταβλητή, που προσδιορίζει το άθροισμα των βαθμολογιών, όταν ισχύει η μηδενική υπόθεση. Η τιμή της για το δοσμένο δείγμα συμβολίζεται w_A .

Ανάλογα με τον τύπο του ελέγχου (μονόπλευρος ή αμφίπλευρος) προκύπτουν οι παρακάτω σχέσεις για το p-value:

- Αν $A > B$: p-value = $P(W_A \geq w_A)$
- Αν $A < B$: p-value = $P(W_A \leq w_A)$
- Αν $A \neq B$: p-value = $2P(W_A \geq w_A)$ (αν το w_A είναι στην πάνω ουρά της W_A)
ή p-value = $2P(W_A \leq w_A)$ (αν το w_A είναι στην κάτω ουρά της W_A).

Στην πράξη όμως χρησιμοποιείται μία άλλη διαδικασία προκειμένου να απορριφθεί ή όχι η H_0 . Με δεδομένα τα μεγέθη n_A, n_B , το επίπεδο σημαντικότητας και την κατανομή που ακολουθεί η W_A κάτω από την μηδενική υπόθεση υπολογίζονται τα ακόλουθα όρια:

- κάτω όριο: $P(W_A \leq w_A) = \varepsilon.σ.$
- άνω όριο: $P(W_A \geq w_A) = \varepsilon.σ.$

Έτσι έπειτα από υπολογισμό της τιμής w_A του δείγματος, μπορεί να απορριφθεί ή όχι η μηδενική υπόθεση. Πρακτικά η μηδενική υπόθεση απορρίπτεται όταν η παρατηρούμενη τιμή w_A είναι εκτός των ορίων της W_A .

Η υλοποίηση του ελέγχου Wilcoxon Rank-Sum γίνεται στο υποκεφάλαιο 7.3.

3.3 Έλεγχος για Τρεις ή Περισσότερους Πληθυσμούς

3.3.1 Έλεγχος Kruskal-Wallis

Ο έλεγχος Kruskal-Wallis είναι ένας μη παραμετρικός έλεγχος που χρησιμοποιείται συνήθως για την σύγκριση 3 ή περισσότερων ανεξάρτητων πληθυσμών. Είναι ουσιαστικά ο έλεγχος που χρησιμοποιείται όταν δεν πληρούνται οι προϋποθέσεις για να γίνει έλεγχος ANOVA. Η τεχνική που χρησιμοποιείται σχετίζεται με την κατάταξη των τιμών σε αύξουσα σειρά, και έπειτα την χρήση ελεγχουσυνάρτησης γνωστής κατανομής για την λήψη της απόφασης.

Πιο αναλυτικά, έστω k , $i = 1, 2, \dots, k$ ανεξάρτητοι πληθυσμοί και έστω ότι οι δύο υποθέσεις είναι οι εξής:

- H_0 : Όλοι οι πληθυσμοί είναι ισοδύναμοι
- H_1 : τουλάχιστον 1 πληθυσμός διαφέρει.

Επιλέγονται k δείγματα ένα από κάθε πληθυσμό. Το σύνολο των στοιχείων όλων των δειγμάτων ονομάζεται **δείγμα-ομάδα** και έχει πλήθος στοιχείων ίσο με $n = n_1 + n_2 + \dots + n_k$. Κατατάσσονται αρχικά όλες οι μετρήσεις σε αύξουσα σειρά και βαθμολογούνται από το 1 μέχρι το n , βάζοντας την τιμή 1 στην μικρότερη μέτρηση κ.ο.κ. Σε περίπτωση ισοβαθμίας, η τιμή που παίρνουν τα δεδομένα είναι ο μέσος όρος των τιμών που θα έπαιρναν αυτά αν δεν ήταν ίσα. Σε αυτή την περίπτωση, αν έχει δοθεί για παράδειγμα η τιμή 3.5 σε δύο ισόπαλες τιμές, τότε η αρίθμηση συνεχίζει από το 5 (θεωρείται δηλαδή ότι έχει δοθεί και το 3 και το 4).

Έπειτα υπολογίζονται οι ποσότητες T_i , $i = 1, 2, \dots, k$, όπου T_i το άθροισμα των βαθμολογιών του δείγματος i . Για την λήψη της απόφασης χρησιμοποιείται η εξής συνάρτηση ελέγχου

$$H = \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1). \quad (3.1)$$

Η συνάρτηση H ακολουθεί την γνωστή κατανομή X^2 , με $k - 1$ βαθμούς ελευθερίας (β.ε.). Η μηδενική υπόθεση απορρίπτεται όταν $H > X_a^2$, όπου a το ε.σ.

Συνήθως η προσέγγιση στην απόρριψη ή μη της μηδενικής υπόθεσης είναι ικανοποιητική αν κάθε δείγμα έχει τουλάχιστον 5 παρατηρήσεις.

Η υλοποίηση του ελέγχου Kruskal-Wallis γίνεται στο υποκεφάλαιο 7.4.

3.3.2 Έλεγχος Friedman-Rank

Ο έλεγχος Friedman-Rank είναι ένας μη παραμετρικός έλεγχος που χρησιμοποιείται όταν οι παρατηρήσεις είναι ζευγαρωτές.

Έστω ότι πρέπει να ελεγχθεί αν $k, k \geq 3$ ομάδες ακολουθούν ίδια κατανομή. Τότε οι δύο υποθέσεις θα είναι:

- H_0 : Οι k ομάδες ακολουθούν ίδια κατανομή
- H_1 : Οι k ομάδες δεν ακολουθούν ίδια κατανομή.

Για την υλοποίηση του ελέγχου χωρίζεται το δείγμα σε υποομάδες (blocks), έτσι ώστε κάθε υποκείμενο της υποομάδας να ανήκει στη ομάδα, αλλά μεταξύ τους οι παρατηρήσεις να είναι ανεξάρτητες.

Έστω ότι το δείγμα χωρίζεται σε l blocks $l \geq 5$. Κάθε στοιχείο του κάθε block βαθμολογείται και υπολογίζονται τα αθροίσματα $T_i, i = 1, 2, \dots, l$ της κάθε ομάδας. Έπειτα χρησιμοποιείται η ελεγκοσυνάρτηση

$$F_r = \frac{12}{nl(l+1)} \sum_{i=1}^l T_i^2 - 3n(l+1) \quad (3.2)$$

Για την συνάρτηση αυτή ισχύει $F_r \sim X^2$, με $l-1$ βαθμούς ελευθερίας. Οπότε από γνωστό πίνακα της κατανομής υπολογίζεται η κρίσιμη τιμή για δεδομένο επίπεδο σημαντικότητας α και γίνεται απόρριψη της μηδενικής υπόθεσης αν $F_r > X_\alpha^2$.

Ο έλεγχος Friedman-Rank χρησιμοποιείται συχνά σε αναλύσεις δεδομένων στην Ιατρική, π.χ. ισοδυναμία θεραπειών, φαρμάκων κλπ.

3.4 Έλεγχος για Ύπαρξη Συσχέτισης

3.4.1 Ο Συντελεστής Συσχέτισης Spearman

Ο **συντελεστής συσχέτισης Spearman** (ρ) εκφράζει την σχέση που έχουν δύο μεταβλητές με την προϋπόθεση ότι αυτή είναι **μονότονη**, οι μεταβλητές συνεχείς ή τουλάχιστον κατηγορικές με διαβαθμίσεις.

Ο έλεγχος για τον συντελεστή είναι μη παραμετρικός, αφού βασίζεται στην τάξη μεγέθους των παρατηρήσεων όταν αυτές διατάσσονται σε αύξουσα σειρά και δεν χρειάζεται καμία υπόθεση για τον πληθυσμό από τον οποίο προέρχεται το τυχαίο δείγμα. Όπως και στον συντελεστή συσχέτισης Pearson έτσι και εδώ, τιμές κοντά στο 0 υποδηλώνουν την απουσία συσχέτισης, ενώ τιμές κοντά στο 1 και το -1 υποδηλώνουν την ύπαρξη θετικής ή αρνητικής συσχέτισης αντίστοιχα.

Έστω δύο μεταβλητές X, Y και ένα δείγμα ζευγαρωτών παρατηρήσεων x_i, y_i μεγέθους n . Οι τιμές της κάθε μεταβλητής βαθμολογούνται σε αύξουσα σειρά, 1 η μικρότερη, 2 η επόμενη, κ.ο.κ. Έπειτα υπολογίζεται ο συντελεστής συσχέτισης Spearman ως εξής:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \quad (3.3)$$

όπου d_i η διαφορά των βαθμολογιών του ζεύγους x_i, y_i .

Ο έλεγχος υποθέσεων είναι ίδιος με αυτόν για τον συντελεστή Pearson, όπως αναφέρθηκε στο υποκεφάλαιο 2.6.2, με τη μηδενική υπόθεση να υποδηλώνει την απουσία συσχέτισης. Ο συντελεστής συσχέτισης ρ ακολουθεί προσεγγιστικά κάτω από την μηδενική υπόθεση την κατανομή t_{n-2} .

Αξίζει να αναφερθούν κάποια συμπεράσματα που προκύπτουν από τη σύγκριση των δύο συντελεστών. Στην περίπτωση που δεν υπάρχει συσχέτιση οι δύο συντελεστές παίρνουν τιμές κοντά στο 0, ενώ όταν η σχέση είναι τέλεια γραμμική παίρνουν τιμές κοντά στο 1 ή το -1, ανάλογα με την κατεύθυνση της συσχέτισης. Παρόλ' αυτά, όταν η σχέση είναι γραμμική όμως ο ρυθμός μεταβολής δεν είναι σταθερός, ο συντελεστής συσχέτισης Spearman μπορεί να είναι 1, ενώ ο συντελεστής συσχέτισης Pearson κοντά στο 1. Τέλος, στην περίπτωση που οι συντελεστές βγαίνουν 0 πρέπει να ελεγχθεί η μορφή της συσχέτισης (π.χ. με διάγραμμα διασποράς) γιατί και στην περίπτωση της πολυωνυμικής οι συντελεστές αυτοί μπορεί να δώσουν τιμές κοντά στο 0.

Ο έλεγχος για τον συντελεστή Spearman γίνεται στο υποκεφάλαιο 7.5.

3.4.2 Ο Συντελεστής Συσχέτισης Pearson χ^2

Ο έλεγχος του συντελεστή συσχέτισης Pearson χ^2 χρησιμοποιείται για να διαπιστωθεί η ανεξαρτησία δύο ποιοτικών ή διακριτών ποσοτικών μεταβλητών.

Έστω δύο μεταβλητές A, B με r_1, r_2, \dots, r_r και c_1, c_2, \dots, c_c τιμές η κάθε μία αντίστοιχα. Έστω τυχαίο δείγμα μεγέθους N , στο οποίο εξετάζονται οι A και B . Τα δεδομένα αυτά μπορούν να αναπαρασταθούν με τη μορφή πίνακα συνάφειας, r γραμμών και c στηλών, με $x_{ij}, i = 1, 2, \dots, r, j = 1, 2, \dots, c$ οι συχνότητες του κάθε κελιού και m_i, n_j τα περιθώρια αθροίσματα, όπως φαίνεται στον Πίνακα 3.1.

Έστω οι δύο υποθέσεις:

- H_0 : Οι A και B είναι ανεξάρτητες
- H_0 : Οι A και B δεν είναι ανεξάρτητες,

A, B	c_1	c_2	c_c	
r_1	x_{11}	x_{12}	x_{1c}	m_1
r_2	x_{21}	x_{22}	x_{2c}	m_2
r_r	x_{r1}	x_{r2}	x_{rc}	m_r
	n_1	n_2	n_c	N

Πίνακας 3.1: Πίνακας συνάφειας $r \times c$

Είναι σημαντικό να αναφερθεί πως ακόμα και αν προκύψει ότι οι δύο μεταβλητές δεν είναι ανεξάρτητες, ο έλεγχος δεν είναι σε θέση να αποφανθεί ποια είναι η αιτία και ποιο το αποτέλεσμα.

Επιπλέον ο συγκεκριμένος έλεγχος εφαρμόζεται μόνο σε αμφίπλευρους ελέγχους αφού ο συντελεστής Pearson χ^2 μετράει αποκλίσεις και προς τις δύο κατευθύνσεις.

Η συνάρτηση ελέγχου του Pearson χ^2 είναι:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(x_{ij} - \hat{x}_{ij})^2}{\hat{x}_{ij}}, \quad (3.4)$$

όπου $\hat{x}_{ij} = \frac{m_i n_j}{N}$ οι αναμενόμενες συχνότητες των κελιών, αν οι μεταβλητές ήταν ανεξάρτητες.

Η περιοχή απόρριψης του ελέγχου είναι η

$$R = \chi^2 > \chi_{(r-1)(c-1)}^2.$$

Προφανώς ο έλεγχος χρησιμοποιεί ασυμπτωτικό αποτέλεσμα αφού παίρνεται σαν δεδομένο ότι η συνάρτηση ελέγχου ακολουθεί την $\chi_{(r-1)(c-1)}^2$.

Για να εφαρμοστεί ο έλεγχος χρειάζεται να πληρούνται κάποιες προϋποθέσεις:

- Το δείγμα πρέπει να είναι τυχαίο και οι παρατηρήσεις ανεξάρτητες.
- Οι αναμενόμενες τιμές πρέπει να είναι μεγαλύτερες ή ίσες του 5.

Ο έλεγχος για τον συντελεστή Pearson χ^2 γίνεται στο υποκεφάλαιο 7.6.

3.5 Συμπεράσματα

Συνοψίζοντας, οι μη παραμετρικοί έλεγχοι χρησιμοποιούνται όταν δεν ικανοποιούνται οι συνθήκες για να χρησιμοποιηθούν οι παραμετρικοί. Όμως στην περίπτωση που αυτές ικανοποιούνται, προτιμούνται οι παραμετρικοί. Με την χρήση των μη παραμετρικών ελέγχων μπορούν να συγκριθούν οι κατανομές πληθυσμών, και στην περίπτωση που οι πληθυσμοί ακολουθούν ίδια κατανομή μπορούν να συγκριθούν οι κεντρικές θέσεις. Γενικά οι μη παραμετρικοί έλεγχοι είναι αρκετά εύχρηστοι, γιατί δεν απαιτείται γνώση των παραμέτρων των πληθυσμών και μπορούν να εφαρμοστούν σε αρκετά μικρά δείγματα. Παρόλο που δεν χρησιμοποιούνται στοιχεία των πληθυσμών και τα δείγματα μπορεί να είναι μικρά, με τους μη παραμετρικούς ελέγχους μπορούν να ληφθούν αξιόπιστες αποφάσεις.

Κεφάλαιο 4

Έλεγχοι Τυχαιοποίησης

4.1 Εισαγωγή

Όπως αναφέρθηκε στα προηγούμενα κεφάλαια, τα βασικά στοιχεία των ελέγχων υποθέσεων είναι οι δυο υποθέσεις H_0 και H_1 , η συνάρτηση ελέγχου, το επίπεδο σημαντικότητας α και η κρίσιμη τιμή. Παρόλα αυτά, οι έλεγχοι τυχαιοποίησης δεν χρειάζονται καμία υπόθεση για τον πληθυσμό και μπορούν να απλοποιήσουν αρκετά τη διαδικασία εξαγωγής συμπερασμάτων, βασισμένοι μόνο στο δείγμα. Οι έλεγχοι τυχαιοποίησης επιλέγονται συχνότερα επειδή δεν υπάρχει καμία προϋπόθεση για να εφαρμοστούν. Στους κλασσικούς ελέγχους υποθέσεων χρειάζεται συνήθως η υπόθεση της κανονικότητας. Μπορεί το ΚΟΘ να αποτελεί μια προσέγγιση κανονικότητας σε μεγάλα δείγματα, στις περιπτώσεις όπου δεν υπάρχει πρόσβαση σε μεγάλο αριθμό δειγμάτων από τον πληθυσμό, όμως προτιμάται η υπόθεση της κανονικότητας για να μην υπάρχει ο περιορισμός του μεγάλου δείγματος. Επιπλέον δεν χρειάζεται το δείγμα να είναι τυχαίο, αφού έπειτα από την διαδικασία της τυχαιοποίησης δημιουργούνται τυχαία δείγματα. Η διαδικασία αυτή απλοποιεί αρκετά την στατιστική συμπερασματολογία, αφού σε πολλά πειράματα δεν είναι δυνατή η τυχαία επιλογή παρατηρήσεων.

Οι έλεγχοι τυχαιοποίησης είναι έλεγχοι σημαντικότητας. Αυτό σημαίνει ότι δεν καταλήγουν σε μια απόφαση όπως οι κλασσικοί έλεγχοι υποθέσεων, αλλά υπολογίζεται μια πιθανότητα για το πόσο ισχυρή είναι η μηδενική υπόθεση.

4.2 Ακριβείς Έλεγχοι Τυχαιοποίησης

Στους **ακριβείς ελέγχους τυχαιοποίησης** υπολογίζονται όλοι οι δυνατοί συνδυασμοί που μπορούν να προκύψουν από το δείγμα. Στις περιπτώσεις που ο αριθ-

μός αυτών των συνδυασμών είναι πολύ μεγάλος, και άρα δύσκολος υπολογιστικά, χρησιμοποιούνται οι προσεγγιστικοί έλεγχοι τυχαιοποίησης που θα αναφερθούν παρακάτω. Οι έλεγχοι αυτοί ονομάζονται έλεγχοι τυχαιοποίησης επειδή γίνεται τυχαία ο καταμερισμός των παρατηρήσεων στις διάφορες ομάδες.

Αρχικά ορίζονται οι δύο υποθέσεις. Σύμφωνα με την μηδενική υπόθεση H_0 δεν υπάρχει κάποια δομή στα δεδομένα, ενώ σύμφωνα με την εναλλακτική H_1 υπάρχει. Έπειτα πρέπει να επιλεγεί μία ελέγχουσυνάρτηση T η οποία να παίρνει διαφορετικές τιμές σε κάθε υπόθεση, να μπορεί δηλαδή να ξεχωρίζει όταν ισχύει η μηδενική υπόθεση και όταν ισχύει η εναλλακτική. Στη συνέχεια υπολογίζεται η τιμή της ελεγχουσυνάρτησης για κάθε έναν από τους συνδυασμούς. Το σύνολο των τιμών μπορεί να χρησιμοποιηθεί για την εκτίμηση της συνάρτησης κατανομής της ελεγχουσυνάρτησης, αφού είναι όλες οι δυνατές τιμές που μπορεί αυτή να πάρει για το συγκεκριμένο δείγμα. Κατόπιν, αφού έχει υπολογιστεί η τιμή t_{obs} της ελεγχουσυνάρτησης για το δείγμα, συγκρίνεται με τις τιμές της ελεγχουσυνάρτησης από τους συνδυασμούς. Έστω ότι ο συνολικός αριθμός των δυνατών συνδυασμών είναι N και ο αριθμός των συνδυασμών όπου $T \geq t_{obs}$ είναι m . Γενικά η μηδενική υπόθεση απορρίπτεται στις μικρές τιμές της T . Όμως με την κατάλληλη χρήση ελεγχουσυνάρτησης και την αλλαγή του προσήμου μπορεί η μηδενική υπόθεση να απορρίπτεται σε μεγάλες τιμές της ελεγχουσυνάρτησης. Και έτσι θα θεωρείται και από εδώ και πέρα. Το p-value υπολογίζεται σύμφωνα με τον παρακάτω τύπο:

$$\text{p-value} = \frac{m}{N}. \quad (4.1)$$

Υπολογίζεται δηλαδή πρακτικά πόσο πιο ακραία είναι η παρατηρούμενη τιμή από αυτές που θα έπαιρνε η συνάρτηση αν ίσχυε η μηδενική υπόθεση. Η μηδενική υπόθεση απορρίπτεται όταν $\text{p-value} < \alpha$.

Η επιλογή της ελεγχουσυνάρτησης είναι αρκετά σημαντική στους ελέγχους τυχαιοποίησης, όπως και στους κλασσικούς ελέγχους. Η συνάρτηση αυτή, όπως αναφέρθηκε παραπάνω, θα πρέπει να μπορεί να διακρίνει ανάμεσα στις δύο υποθέσεις. Πρέπει δηλαδή να παίρνει διαφορετικές τιμές όταν ισχύει η κάθε υπόθεση. Όμως τα πλεονεκτήματα στους ελέγχους τυχαιοποίησης σχετικά με την ελεγχουσυνάρτηση είναι αρκετά. Αρχικά δεν χρειάζεται να πινακοποιηθούν οι τιμές της, αφού όλες οι δυνατές τιμές που μπορεί να πάρει στο τυχαίο δείγμα και άρα η εκτίμηση της συνάρτησης πυκνότητας πιθανότητας της προκύπτουν από τον υπολογισμό της στα τυχαιοποιημένα δείγματα. Σημειώνεται πώς πολλές φορές οι ελεγχουσυναρτήσεις είναι αρκετά πολύπλοκες ώστε να πινακοποιηθούν. Επιπλέον, σε αρκετές περιπτώσεις η ελεγχουσυνάρτηση έχει γνωστή κατανομή κάτω από την μηδενική υπόθεση μόνο ασυμπτωτικά, κάτι που θα απαιτούσε πολύ μεγάλο δείγμα.

4.3 Προσεγγιστικοί Έλεγχοι Τυχαιοποίησης

Πολλές φορές ο αριθμός των δυνατών συνδυασμών είναι πολύ μεγάλος υπολογιστικά. Σε αυτές τις περιπτώσεις χρησιμοποιούνται οι **προσεγγιστικοί έλεγχοι τυχαιοποίησης**, όπου υπολογίζεται ένας αριθμός k συνδυασμών. Τα υπόλοιπα βήματα του ελέγχου είναι ίδια με αυτά στην περίπτωση των ακριβών ελέγχων τυχαιοποίησης, με την διαφορά ότι τώρα το p-value υπολογίζεται ως εξής:

$$\text{p-value} = \frac{m + 1}{k + 1}, \quad (4.2)$$

όπου m ο αριθμός των συνδυασμών με $T \geq t_{obs}$.

Όμως στην περίπτωση των προσεγγιστικών ελέγχων προκύπτει το εξής ερώτημα: Πόσοι πρέπει να είναι οι k συνδυασμοί που επιλέγονται; Η απάντηση είναι πως όσο μεγαλύτερο είναι το k , τόσο μεγαλύτερη είναι η βεβαιότητα της απόφασης. Η αύξηση όμως του k μπορεί να προκαλέσει υπολογιστικό κόστος.

Επιπλέον ανάλογα με την επιλογή των συνδυασμών δύο ερευνητές μπορεί να υπολογίσουν διαφορετικά p-values. Αν όμως οι διαφορετικές αυτές τιμές δείχνουν προς την ίδια κατεύθυνση, δηλαδή την απόρριψη ή μη της μηδενικής υπόθεσης, τότε δεν υπάρχει κάποιο πρόβλημα. Αυτό δείχνει πως το p-value διαφοροποιείται ανάλογα με τον αριθμό των συνδυασμών που θα επιλεγούν και επομένως είναι μια τυχαία μεταβλητή στην οποία μπορεί να γίνει στατιστική συμπερασματολογία.

4.3.1 Εκτίμηση p-value

Ενδιαφέρον αποτελεί το πως προκύπτει η παραπάνω προσέγγιση του p-value. Η επιλογή k συνδυασμών σημαίνει και επιλογή k τιμών της T . Επόμενως υπάρχουν $k + 1$ διαστήματα στη γραμμή των πραγματικών αριθμών στα οποία μπορεί να ανήκει η παρατηρούμενη τιμή της ελεγχουσυνάρτησης. Για να ερμηνευτεί ο αριθμητής πρέπει να επισημανθεί ότι το πραγματικό αλλά άγνωστο p-value είναι η πιθανότητα $P(T \geq t_{obs})$ και συμβολίζεται με p_{true} . Έστω ότι m είναι ο αριθμός των συνδυασμών για τους οποίους $T \geq t_{obs}$. Η μεταβλητή m είναι μία τυχαία μεταβλητή, η οποία ακολουθεί την διωνυμική κατανομή σε k επαναλήψεις με πιθανότητα επιτυχίας το p_{true} . Συνεπώς ισχύει

$$E(m) = kp_{true} \quad (4.3)$$

Υπολογίζεται επίσης η μέση τιμή του εκτιμώμενου p-value.

$$E(\text{p-value}) = E\left(\frac{m + 1}{k + 1}\right) = \frac{E(m) + 1}{k + 1}. \quad (4.4)$$

Αν γίνει αντικατάσταση της 4.3 στην 4.4 προκύπτει

$$E(\text{p-value}) = \frac{kp_{true} + 1}{k + 1}. \quad (4.5)$$

Είναι φανερό ότι $E(p\text{-value}) > p_{true}$. Επομένως η προσέγγιση αυτή υπερεκτιμά το πραγματικό p -value. Αν οριζόταν ως $p\text{-value} = \frac{m}{k+1}$ τότε θα το υποεκτιμούσε. Η υπερεκτίμηση είναι προτιμότερη όταν αφορά την λήψη αποφάσεων απόρριψης, γιατί έτσι απορρίπτεται η μηδενική υπόθεση δυσκολότερα, άρα με μεγαλύτερη βεβαιότητα.

Αν τώρα το p -value που υπολογίζεται από το δείγμα είναι κοντά στο επίπεδο σημαντικότητας πρέπει να αυξηθεί το k μέχρι το 95% Δ.Ε. να μην περιέχει το επίπεδο σημαντικότητας α .

Για πολύ μεγάλα k το p -value μπορεί να εκτιμηθεί ως

$$p\text{-value} = \frac{m}{k}. \quad (4.6)$$

4.4 Ενδιαφέρουσες Περιπτώσεις

4.4.1 Έλεγχος για τη Μέση Τιμή Δύο Πληθυσμών

Έστω δύο δείγματα X_A, X_B , μεγέθους n_1, n_2 αντίστοιχα, με μέσες τιμές \bar{x}_A, \bar{x}_B από δύο πληθυσμούς A, B . Ένας ερευνητής θέλει να ελέγξει αν οι δύο πληθυσμοί είναι ισοδύναμοι. Η υπόθεση μπορεί να απλοποιηθεί στον έλεγχο της ισότητας των μέσων τιμών της παρατηρούμενης ποσότητας.

Ορίζεται η μηδενική υπόθεση $H_0: \mu_A = \mu_B$.

Η εναλλακτική υπόθεση θα είναι $H_1: \mu_A > \mu_B$ είτε $H_1: \mu_A < \mu_B$ για μονόπλευρο έλεγχο, είτε $H_1: \mu_A \neq \mu_B$ για αμφίπλευρο.

Αν υπήρχαν οι υποθέσεις περί κανονικότητας θα μπορούσε να γίνει ένας παραμετρικός έλεγχος t -test, όπως αναφέρθηκε στο υποκεφάλαιο 2.6.1. Παρόλαυτα ένας έλεγχος τυχαιοποίησης προσπερνάει την άγνοια που υπάρχει για την συμπεριφορά των πληθυσμών.

Η λογική του ελέγχου τυχαιοποίησης βασίζεται στο γεγονός ότι οποιαδήποτε διαφορά εμφανίζεται στα δείγματα είναι τυχαία και οποιαδήποτε ανακατανομή των παρατηρήσεων στα δύο δείγματα είναι ισοπίθανη.

Έστω N ο συνολικός αριθμός των παρατηρήσεων. Ο συνολικός αριθμός των διαφορετικών συνδυασμών είναι ο συνολικός αριθμός των συνδυασμών N στοιχείων σε δύο ομάδες με n_1 και n_2 στοιχεία η κάθε μία, δηλαδή $\binom{N}{n_1}$.

Η περιοχή απόρριψης καθορίζεται από την H_1 . Μια καλή συνάρτηση ελέγχου είναι η

$$T = |\bar{x}_A - \bar{x}_B|$$

για αμφίπλευρο έλεγχο, και η

$$T = \bar{x}_A - \bar{x}_B$$

για μονόπλευρο.

Με κατάλληλο ορισμό της συνάρτησης ελέγχου, η T παίρνει τιμές κοντά στο 0 αν ισχύει η H_0 και μεγάλες αν δεν ισχύει.

Αφού δημιουργηθούν όλοι οι διαφορετικοί συνδυασμοί υπολογίζεται η T για κάθε ένα και συγκρίνεται με την παρατηρούμενη τιμή της, t_{obs} . Αυτές οι τιμές χρησιμοποιούνται για την εκτίμηση της συνάρτησης κατανομής της T , η οποία είναι άγνωστη και για αυτό χρειάζεται να γίνει ο έλεγχος τυχαιοποίησης. Έστω ότι το πλήθος των δειγμάτων για τα οποία ισχύει $T \geq t_{obs}$ είναι m . Στη συνέχεια υπολογίζεται το p-value

$$\text{p-value} = \frac{m}{N}.$$

Στην περίπτωση που το N είναι αρκετά μεγάλο, επιλέγονται k δείγματα και ακολουθηθεί η ίδια διαδικασία. Σε αυτή την περίπτωση το p-value υπολογίζεται ως εξής:

$$\text{p-value} = \frac{m + 1}{k + 1}.$$

Η υλοποίηση αυτού του ελέγχου παρουσιάζεται στο υποκεφάλαιο 7.3.

4.4.2 Συσχέτιση Δύο Μεταβλητών

Οι έλεγχοι τυχαιοποίησης μπορούν να χρησιμοποιηθούν στην περίπτωση που ο ερευνητής θέλει να ελέγξει αν υπάρχει σχέση ανάμεσα σε δύο μεταβλητές. Η μη-δενική υπόθεση ορίζει την απουσία κάποιας σχέσης, ενώ η εναλλακτική την ύπαρξη.

Συμπεπώς

H_0 : Δεν υπάρχει συσχέτιση

H_1 : Υπάρχει συσχέτιση

Η εναλλακτική υπόθεση θα πρέπει να είναι μονόπλευρη, δηλαδή η συσχέτιση να είναι θετική ή αρνητική, ώστε να γίνει κατάλληλη επιλογή ελεγχουσυνάρτησης. Για τον συγκεκριμένο έλεγχο μπορούν να χρησιμοποιηθούν διάφορα μέτρα συσχέτισης (συντελεστής Pearson, συντελεστής β της γραμμικής παλινδρόμησης κλπ), ανάλογα με το τι σχέση χρειάζεται να εξετασθεί. Στα πλαίσια της τυχαιοποίησης επιλέγονται k n -άδες ζευγών από το δείγμα, ή αν είναι υπολογιστικά δυνατό όλες. Η συγκεκριμένη διαδικασία μπορεί να υλοποιηθεί υπολογιστικά διατηρώντας τα στοιχεία της μιας μεταβλητής σταθερά και ανακατεύοντας τα στοιχεία της άλλης. Επομένως σε ένα δείγμα

(X_i, Y_i) μεγέθους n όλοι οι δυνατοί συνδυασμοί είναι $n!$. Στη συνέχεια ακολουθείται η διαδικασία που περιγράφηκε παραπάνω για τον υπολογισμό του p -value. Σημειώνεται ότι αν το k παραμείνει σταθερό και η διαδικασία επαναληφθεί πολλές φορές, τότε οι τιμές που θα προκύψουν μπορούν να χρησιμοποιηθούν σαν εκτίμηση της συνάρτησης κατανομής του p -value. Και άρα μπορεί να πραγματοποιηθεί και εκεί στατιστική συμπερασματολογία.

Στην περίπτωση που το δείγμα ακολουθεί την κανονική κατανομή μπορεί να γίνει χρήση των κλασικών ελέγχων για τον συντελεστή συσχέτισης. Παρόλ' αυτά χρησιμοποιώντας τον έλεγχο τυχαιοποίησης δεν χρειάζεται να γίνει καμία υπόθεση για κανονικότητα.

Ο έλεγχος για ύπαρξη γραμμικής συσχέτισης μεταξύ 2 μεταβλητών υλοποιείται στο υποκεφάλαιο 7.5.

4.4.3 Έλεγχος Διακύμανσης ANOVA

Έστω ότι χρειάζεται να συγκριθούν οι μέσες τιμές 3 ή περισσότερων ανεξάρτητων πληθυσμών. Αν πληρούνται οι υποθέσεις περί κανονικότητας μπορεί να γίνει χρήση του ελέγχου ANOVA. Παρόλ' αυτά η εξασφάλιση της κανονικότητας μπορεί να είναι δύσκολη σε πειραματικά δεδομένα, αφού μπορεί να μην υπάρχουν πληροφορίες για τον πληθυσμό ή τα δείγματα να μην είναι αντιπροσωπευτικά ώστε η κανονικότητα να γενικευτεί στους πληθυσμούς. Για αυτό το λόγο μπορεί να γίνει χρήση ANOVA με έλεγχο τυχειότητας.

Η λογική που ακολουθείται είναι η ίδια με τους προηγούμενους ελέγχους τυχαιοποίησης. Στο συγκεκριμένο παράδειγμα η κατασκευή των συνδυασμών μπορεί να γίνει κατασκευάζοντας μία στήλη με τις παρατηρήσεις, μία στήλη με τα επίπεδα που αυτές ανήκουν και έπειτα ανακατεύοντας τη μία στήλη και κρατώντας την άλλη σταθερή. Αν δεν είναι δυνατό να υπολογιστούν όλοι οι συνδυασμοί, μπορεί να γίνει επιλογή k συνδυασμών.

Το ερώτημα είναι ποια θα ήταν μία κατάλληλη ελεγχοσυνάρτηση. Μπορεί να γίνει χρήση της συνάρτησης F , που γίνεται χρήση στον απλό έλεγχο ANOVA. Παρόλ' αυτά για να απλοποιηθούν οι υπολογισμοί μπορεί να γίνει χρήση και άλλης συνάρτησης. Αν όλες οι μέσες τιμές ήταν ίσες, τότε οι τετραγωνικές αποκλίσεις των μέσων τιμών \bar{x}_i κάθε επιπέδου από την συνολική μέση τιμή \bar{x} αν δεν υπήρχε ο παράγοντας θα ήταν πολύ μικρές. Επιλέγεται, λοιπόν, η παρακάτω ελεγχοσυνάρτηση :

$$T = \sum_{i=1}^k (\bar{x}_i - \bar{x})^2. \quad (4.7)$$

Έπειτα το p-value υπολογίζεται όπως έχει αναφερθεί παραπάνω.

Ο έλεγχος ANOVA με τυχαιοποίηση γίνεται στο υποκεφάλαιο 7.4

4.4.4 Ακριβής Έλεγχος του Fisher για Πίνακες Συνάφειας

Ο ακριβής έλεγχος του Fisher για πίνακες συνάφειας χρησιμοποιείται για τον έλεγχο της ανεξαρτησίας δύο ποιοτικών ή διακριτών ποσοτικών μεταβλητών. Αρχικά ορίστηκε για πίνακες 2×2 και έπειτα με την εξέλιξη των υπολογιστών γενικεύτηκε σε μεγαλύτερους, αφού ο έλεγχος αυτός απαιτεί υπολογιστικά πολύπλοκους υπολογισμούς.

Έστω δύο μεταβλητές A, B δίτιμες με τιμές r_1, r_2 για την A και c_1, c_2 για την B και N το μέγεθος τυχαίου δείγματος με παρατηρήσεις της μορφής (r_i, c_j) . Προκύπτει επομένως ο πίνακας συνάφειας 4.1.

A, B	c_1	c_2	
r_1	x_{11}	x_{12}	m_1
r_2	x_{21}	x_{22}	m_2
	n_1	n_2	N

Πίνακας 4.1: Πίνακας συνάφειας 2×2

Τα x_{ij} , για $i = 1, 2, j = 1, 2$ είναι οι συχνότητες του κάθε κελιού και τα m_i, n_j τα αθροίσματα της κάθε γραμμής και στήλης αντίστοιχα.

Ορίζονται οι εξής υποθέσεις:

- Οι μεταβλητές A και B είναι ανεξάρτητες.
- Οι μεταβλητές A και B δεν είναι ανεξάρτητες.

Ως συνάρτηση ελέγχου θα χρησιμοποιηθεί ο συντελεστής συσχέτισης Pearson χ^2 , ο οποίος είναι κατάλληλος για έλεγχο της συσχέτισης σε πίνακες συνάφειας όπως έχει αναφερθεί και στο Κεφάλαιο 3.

Η συνάρτηση ελέγχου του Pearson χ^2 είναι:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(x_{ij} - \hat{x}_{ij})^2}{\hat{x}_{ij}}, \quad (4.8)$$

όπου $\hat{x}_{ij} = \frac{m_i n_j}{N}$ οι αναμενόμενες συχνότητες των κελιών, αν οι μεταβλητές ήταν ανεξάρτητες.

Επειδή ο έλεγχος για τον συντελεστή Pearson χ^2 χρησιμοποιεί ασυμπτωτικό αποτέλεσμα, για την χρήση του χρειάζεται να πληρούνται οι προϋποθέσεις που έχουν αναφερθεί. Στην περίπτωση που δεν πληρούνται οι προϋποθέσεις, αλλά και γενικά, είναι προτιμότερο να επιλεγθεί ένας έλεγχος τυχαιοποίησης, αφού έλεγχοι που χρησιμοποιούν ασυμπτωτικά αποτελέσματα δεν είναι τόσο ακριβείς.

Η συνάρτηση 4.8 επιλέγεται κυρίως όταν ο έλεγχος είναι αμφίπλευρος, δηλαδή όταν η εναλλακτική υπόθεση υποδηλώνει την απουσία ανεξαρτησίας γενικά, επειδή ο συντελεστής Pearson χ^2 υπολογίζει αποκλίσεις και προς τις δύο κατευθύνσεις.

Για την πραγματοποίηση του ελέγχου χρειάζεται να υπολογιστούν όλοι οι δυνατοί πίνακες συνάφειας, να υπολογιστεί η συνάρτηση ελέγχου για κάθε έναν και να βρεθεί πόσο ακραία είναι η παρατηρούμενη τιμή. Οι τιμές που θα πάρει η συνάρτηση ελέγχου ενδέχεται να είναι διακριτές και επαναλθανόμενες, παρόλ' αυτά θεωρείται ότι είναι συνέχης και κάτω από το ασυμπτωτικό αποτέλεσμα ισχύει $\chi^2 \sim \chi^2_{(r-1)(c-1)}$, όπου r, c το πλήθος των γραμμών και των στηλών του πίνακα αντίστοιχα.

Αποδεικνύεται ότι ο συνολικός αριθμός των δυνατών πινάκων είναι $\binom{N}{n_1}$. Το πλήθος των δυνατών πινάκων μπορεί να είναι πολύ μεγάλο, ειδικά όταν το N είναι μεγάλο. Υπάρχει η δυνατότητα εύρεσης του p -value χωρίς αναγραφή όλων των πινάκων και υπολογισμό της συνάρτησης ελέγχου για κάθε έναν.

Αφού τα περιθώρια αθροίσματα είναι σταθερά, αρκεί να οριστεί μόνο ένα στοιχείο του πίνακα ώστε να προκύψουν και τα υπόλοιπα. Επομένως αν οριστεί μόνο το x_{11} μπορούν να προκύψουν και τα υπόλοιπα και αποδυναμώνεται ότι οι φορές εμφάνισης αυτού του πίνακα είναι

$$\binom{m_1}{x_{11}} \binom{m_2}{n_1 - x_{11}}.$$

Συνεπώς η πιθανότητα εμφάνισης του κάθε πίνακα είναι

$$P(x_{11}) = \frac{\binom{m_1}{x_{11}} \binom{m_2}{n_1 - x_{11}}}{\binom{N}{n_1}}.$$

Όμως το άθροισμα τέτοιων πιθανοτήτων προϋποθέτει το άθροισμα πολλών παραγοντικών που είναι υπολογιστικά δύσκολο, ειδικά σε μεγάλες διαστάσεις πινάκων. Με χρήση επαναληπτικών και έξυπνων αλγορίθμων που υπάρχουν και στη βιβλιογραφία μπορούν να υπολογιστούν πολύπλοκες πιθανότητες, γενικά και ειδικά για τον υπολογισμό του p -value και να γίνουν επαναληπτικές πράξεις για να απλοποιηθούν τα παραπάνω παραγοντικά.

Στην περίπτωση που το πλήθος των διαφορετικών πινάκων συνάφειας είναι πολύ μεγάλο μπορεί να χρησιμοποιηθεί ο προσεγγιστικός έλεγχος τυχαιοποίησης

στον οποίο επιλέγεται ένα δείγμα από τους συνδυασμούς.

Η υλοποίηση του ελέγχου Fisher για πίνακες συνάφειας παρουσιάζεται στο υποκεφάλαιο 7.6.

4.5 Συμπεράσματα

Αξίζει να αναφερθούν τα παρακάτω συμπεράσματα :

- Η μόνη προϋπόθεση για την επιλογή της ελεγχοσυνάρτησης είναι να μπορεί να διακρίνει ανάμεσα στις δύο υποθέσεις, αφού κάθε δυνατή τιμή που μπορεί να πάρει υπολογίζεται κατά τον έλεγχο, χωρίς να είναι χρήσιμη η γνώση για την κατανομή της. Παρόλ' αυτά δύο διαφορετικές συναρτήσεις μπορούν να οδηγήσουν σε διαφορετική εκτίμηση του p -value, αφού η μία μπορεί να έχει δυνατότητα να διακρίνει ανάμεσα στις δύο υποθέσεις περισσότερο από την άλλη.
- Δεν χρειάζεται καμία υπόθεση για την κανονικότητα του πληθυσμού. Η μόνη υπόθεση που χρειάζεται ορίζεται από την μηδενική υπόθεση και είναι η απουσία οποιασδήποτε δομής. Θα πρέπει δηλαδή κάθε συνδυασμός να είναι πιθανός κάτω από τη μηδενική υπόθεση.
- Η ισχύς των ελέγχων τυχαιοποίησης είναι πολύ μεγαλύτερη από αυτή των παραμετρικών. Ακόμα και αν η επιλογή της ελεγχοσυνάρτησης είναι η ίδια. Αυτό οφείλεται στο γεγονός ότι δεν γίνεται καμία παραμετρική υπόθεση. Τα ασυμπτωτικά αποτελέσματα είναι λιγότερο αξιόπιστα. Επομένως, όταν είναι δυνατή η πραγματοποίησή τους, οι έλεγχοι τυχαιοποίησης, ακριβείς ή παραμετρικοί, προτιμώνται.
- Οι έλεγχοι τυχαιοποίησης μπορούν να βελτιώσουν ακόμα και τους μη παραμετρικούς. Παρόλο που δεν γίνεται κάποια υπόθεση περί κανονικότητας του πληθυσμού, στους μη παραμετρικούς χρησιμοποιούνται επίσης ασυμπτωτικά αποτελέσματα.
- Ένα μειονέκτημα των ελέγχων τυχαιοποίησης είναι πως δεν μπορούν να εφαρμοστούν για συγκεκριμένη μηδενική υπόθεση, αφού η λογική τους αφορά την τυχαιοποίηση και την επιλογή οποιουδήποτε πιθανού συνδυασμού κάτω από αυτήν. Στην περίπτωση που η μηδενική υπόθεση είναι συγκεκριμένη, μπορεί να γίνει χρήση των ελέγχων Monte Carlo, οι οποίοι θα αναφερθούν στο επόμενο κεφάλαιο.

Κεφάλαιο 5

Έλεγχοι Monte Carlo

5.1 Εισαγωγή

Η λογική των ελέγχων Monte Carlo βασίζεται στην προσομοίωση των τιμών από μια γνωστή κατανομή προκειμένου να ελεγχθεί αν μία τ.μ. ακολουθεί την κατανομή αυτή. Για αυτό το λόγο χρησιμοποιούνται όταν η μηδενική υπόθεση είναι συγκεκριμένη, π.χ. «Τα δεδομένα ακολουθούν διωνυμική κατανομή», επειδή μπορούν να προσομοιωθούν δείγματα από την κατανομή. Οι έλεγχοι αυτοί μπορούν να χρησιμοποιηθούν γενικά για την προσομοίωση των τιμών οποιασδήποτε κατανομής, όχι απαραίτητα για τη χρήση της σε ελέγχους υποθέσεων. Μέσω των προσομοιωμένων τιμών μπορεί να αναπαρασταθεί γραφικά η ζητούμενη κατανομή, αλλά και να υπολογιστούν συγκεκριμένα μεγέθη αυτής. Είναι σημαντικό να αναφερθεί ότι στους συγκεκριμένους ελέγχους απαιτείται η γνώση της κατανομής του πληθυσμού προκειμένου να γίνει η προσομοίωση.

5.2 Περιγραφή Ελέγχου

Όπως στους ελέγχους τυχαιοποίησης, έτσι και στους ελέγχους Monte Carlo υπάρχουν δείγματα στα οποία υπολογίζεται η τιμή της ελεγχουσυνάρτησης και έπειτα υπολογίζεται το p -value. Πρέπει όμως η μηδενική υπόθεση να είναι συγκεκριμένη και όχι απλά να ορίζει την απουσία δομής, αφού αυτή ορίζει την κατανομή από την οποία θα προσομοιωθούν τα δείγματα. Η διαδικασία που ακολουθείται είναι ίδια με τους ελέγχους τυχαιοποίησης, όπως αναφέρθηκε στο Κεφάλαιο 4. Η μόνη διαφορά έγκυται στον τρόπο με τον οποίο παράγονται τα δείγματα.

Σχηματικά τα βήματα των ελέγχων Monte Carlo είναι τα εξής:

- Βήμα 1: Ορισμός H_0 και H_1 .
- Βήμα 2: Επιλογή κατάλληλης ελεγχουσυνάρτησης T .
- Βήμα 3: Προσομοίωση k δειγμάτων μεγέθους n , όπου n το μέγεθος του αρχικού δείγματος, από την κατανομή που ορίζει η μηδενική υπόθεση.
- Βήμα 4: Υπολογισμός της T για κάθε δείγμα και του t_{obs} από το αρχικό.
- Βήμα 5: Υπολογισμός p-value ως εξής:

$$\text{p-value} = \frac{m + 1}{k + 1},$$

όπου m ο αριθμός των δειγμάτων για τα οποία ισχύει $T \geq t_{obs}$, $T \leq t_{obs}$ ή $|T| \geq t_{obs}$, ανάλογα με τη μορφή της εναλλακτικής υπόθεσης.

Η προσέγγιση του p-value βασίζεται στη λογική ότι αν αυτό είναι πολύ μεγάλο, τότε η παρατηρούμενη τιμή της ελεγχουσυνάρτησης είναι πολύ μακριά από τις προσομοιωμένες τιμές, οπότε το δεδομένο δείγμα δεν έχει προκύψει από την κατανομή που ορίζει η μηδενική υπόθεση. Αντίστροφα, αν το p-value είναι πολύ μικρό και άρα η μηδενική υπόθεση δεν απορρίπτεται, σημαίνει ότι η παρατηρούμενη τιμή δεν είναι ακραία και συνεπώς το δείγμα έχει προκύψει από την κατανομή που ορίζει η μηδενική υπόθεση.

5.3 Ενδιαφέροντα Παραδείγματα

5.3.1 Έλεγχος για τη Μέση Τιμή Ενός Πληθυσμού

Μία ενδιαφέρουσα περίπτωση της μεθόδου Monte Carlo είναι ο έλεγχος για τη μέση τιμή ενός πληθυσμού. Η μέθοδος αυτή ενδείκνυται όταν το δείγμα είναι πολύ μικρό, και επομένως δεν μπορούν να χρησιμοποιηθούν οι κλασσικοί έλεγχοι υποθέσεων. Για την εφαρμογή της μεθόδου Monte Carlo είναι απαραίτητη η γνώση της κατανομής του πληθυσμού, προκειμένου να προσομοιωθούν από αυτήν δείγματα.

Έστω ότι η μηδενική υπόθεση είναι της μορφής:

$$H_0 : \mu = \mu_0.$$

Τότε η ιδανική συνάρτηση ελέγχου θα ήταν η:

$$T = \bar{x} - \mu_0, \tag{5.1}$$

ανεξάρτητα από τη μορφή της εναλλακτικής υπόθεσης, ανεξάρτητα από το αν ο έλεγχος είναι μονόπλευρος ή αμφίπλευρος δηλαδή.

Η 5.1 είναι κατάλληλη συνάρτηση ελέγχου, αφού όταν ισχύει, η H_0 παίρνει τιμές κοντά στο 0, και επομένως είναι εύκολο να βρεθούν, αν υπάρχουν, ακραίες τιμές.

Η εναλλακτική υπόθεση H_1 θα καθορίσει αν θα απορριφθεί η H_0 βάσει ακραίων τιμών προς τα δεξιά, προς τα αριστερά ή και προς τις δύο κατευθύνσεις. Διακρίνονται οι εξής περιπτώσεις:

- Για $H_1: \mu < \mu_0$ αναζητούνται ακραίες τιμές στη δεξιά ουρά της T
- Για $H_1: \mu > \mu_0$ αναζητούνται ακραίες τιμές στην αριστερή ουρά της T
- Για $H_1: \mu \neq \mu_0$ αναζητούνται ακραίες τιμές και προς τις δύο κατευθύνσεις.

Για την πραγματοποίηση του ελέγχου απαιτείται αρχικά να προσομοιωθούν k δείγματα από την κατανομή του πληθυσμού. Στην περίπτωση που οι παράμετροι της κατανομής είναι άγνωστες, μπορούν να εκτιμηθούν είτε με εκτιμητήρια μέγιστης πιθανοφάνειας είτε με ροποεκτιμητήρια, από το δοσμένο δείγμα. Αφού προσομοιωθούν τα k δείγματα, υπολογίζεται η τιμή της T για κάθε δείγμα. Οι k τιμές της T αποτελούν δείγμα από την άγνωστη κατανομή αυτής της τ.μ., όταν η H_0 είναι αληθής. Τέλος αρκεί να υπολογιστεί το πλήθος των ακραίων τιμών, έστω m . Το ποιος τιμές θεωρούνται ακραίες καθορίζεται από τη μορφή της εναλλακτικής υπόθεσης όπως αναφέρθηκε παραπάνω.

Τελικά το p-value υπολογίζεται ως εξής:

$$\text{p-value} = \frac{m + 1}{k + 1}$$

Ο έλεγχος της μέσης τιμής με τη μέθοδο Monte Carlo υλοποιείται στο υποκεφάλαιο 7.2.

5.3.2 Monte Carlo Καλής Προσαρμογής

Οι έλεγχοι Monte Carlo καλής προσαρμογής χρησιμοποιούνται για να διαπιστωθεί αν ένα δοσμένο δείγμα ακολουθεί μια συγκεκριμένη κατανομή.

Για την πραγματοποίηση αυτού του ελέγχου απαιτείται προσομοίωση δειγμάτων από την κατανομή, η οποία ελέγχεται. Για να προσδιοριστεί η κατανομή χρειάζεται να προσδιοριστούν οι παράμετροι της. Αυτό μπορεί να γίνει είτε με τη μέθοδο των ροπών είτε με τη μέθοδο της Εκτιμητριας Μέγιστης Πιθανοφάνειας (Ε.Μ.Π.).

Ιδιαίτερη προσοχή απαιτεί ο ορισμός της εναλλακτικής υπόθεσης γιατί επηρεάζει την επιλογή της συνάρτησης ελέγχου. Είναι σημαντικό η συνάρτηση ελέγχου

να μπορεί να διαχωρίσει ανάμεσα στις δύο υποθέσεις.

Απ' τη μία θα μπορούσε να χρησιμοποιηθεί μια συνάρτηση που θα υπολόγιζε αποστάσεις των παρατηρήσεων από την κατανομή. Τέτοιες συναρτήσεις χρησιμοποιούν οι παραμετρικοί έλεγχοι, για παράδειγμα ο χ^2 καλής προσαρμογής.

Απ' την άλλη μπορούν να χρησιμοποιηθούν συναρτήσεις που χαρακτηρίζουν την κατανομή της H_0 και παίρνουν συγκεκριμένες τιμές διαφοροποιώντας την από όλες τις άλλες κατανομές ή τουλάχιστον από την H_1 . Η επιλογή τέτοιων συναρτήσεων απαιτεί πολύ καλή γνώση στατιστικής. Για παράδειγμα για μια κατανομή Poisson ισχύει ότι η μέση τιμή και η διακύμανση είναι ίσες. Αυτό δεν ισχύει για την Αρνητική Διωνυμική, όμως μπορεί να υπάρξει Κανονική κατανομή με μέση τιμή ίση με τη διακύμανση. Συνεπώς η επιλογή της συνάρτησης ελέγχου απαιτεί πολλή προσοχή τόσο στα χαρακτηριστικά της συνάρτησης όσο και στην εναλλακτική υπόθεση.

Αφού εκτιμηθούν οι παράμετροι της κατανομής και επιλεγεί η κατάλληλη συνάρτηση, προσομοιώνονται από αυτή την κατανομή, έστω k δείγματα, μεγέθους n , δηλαδή όσο το αρχικό δοσμένο δείγμα.

Για κάθε δείγμα υπολογίζεται η συνάρτηση ελέγχου. Στη συνέχεια καταμετρώνται σε πόσα από αυτά η τιμή T της συνάρτησης ελέγχου είναι μεγαλύτερη ή μικρότερη. Έστω ότι αυτά τα δείγματα είναι m . Η περιοχή απόρριψης της H_0 καθορίζεται από την εναλλακτική και τις τιμές που παίρνει η συνάρτηση ελέγχου στην κατανομή της εναλλακτικής.

Τέλος υπολογίζεται το p-value όπως αναφέρθηκε και στα προηγούμενα κεφάλαια, δηλαδή

$$\text{p-value} = \frac{m + 1}{k + 1}$$

Ο έλεγχος καλής προσαρμογής με τη μέθοδο Monte Carlo υλοποιείται στο υποκεφάλαιο 7.7.

5.4 Συμπεράσματα

Προκύπτουν τα εξής συμπεράσματα :

- Για την επιλογή της ελεγκοσυνάρτησης είναι σημαντική η καλή γνώση στατιστικής και η εμπειρία, αφού η μηδενική ή και οι δύο υποθέσεις υποδηλώνουν κάποια κατανομή. Συνεπώς θα πρέπει να γίνει χρήση κάποιου μέτρου με γνωστές ιδιότητες και πολλές φορές μοναδικές κάτω από τις υποθέσεις.

- Για την επιλογή του αριθμού k των επαναλήψεων και την εκτίμηση του p -value ισχύει ότι ισχύει και στους ελέγχους τυχαιοποίησης.
- Χρειάζεται η πρόσθετη πληροφορία για την κατανομή του πληθυσμού, ώστε να μπορεί να γίνει προσομοίωση από συγκεκριμένη κατανομή ή με συγκεκριμένα χαρακτηριστικά.
- Οι έλεγχοι Monte Carlo προτιμώνται σε ελέγχους καλής προσαρμογής, αφού η μηδενική υπόθεση κατευθύνει προς την κατανομή που θα προσομοιωθούν τα δείγματα.

Γενικά αν δεν είναι εφικτός ο υπολογισμός μια στατιστικής συνάρτησης από πληθυσμό με γνωστή κατανομή, τότε αυτή μπορεί να προσομοιωθεί και έπειτα αν χρειάζεται να υπολογισθούν διάφορα στατιστικά μεγέθη της, τυπικά σφάλματα, πιθανότητες της συνάρτησης, κλπ. Οι Monte Carlo μπορεί να απαιτούν την γνώση της κατανομής του πληθυσμού, παρόλ' αυτά μπορούν να ξεπεράσουν το πρόβλημα του μικρού ή μη αντιπροσωπευτικού δείγματος.

Κεφάλαιο 6

Bootstrap

6.1 Εισαγωγή

Η μέθοδος Bootstrap αποτελεί και αυτή μια μέθοδο τυχαιοποίησης, όπως η μέθοδος Monte Carlo, με διαφορά ότι αλλάζει ο τρόπος με τον οποίο παράγονται τα δείγματα. Η μέθοδος Bootstrap προτιμάται πολλές φορές σε σχέση με τους ελέγχους τυχαιοποίησης, αφού αυτοί μπορούν να εφαρμοστούν σε συγκεκριμένες υποθέσεις (απουσία δομής) και σε σχέση με τους ελέγχους Monte Carlo, αφού αυτοί προϋποθέτουν γνώση της κατανομής του πληθυσμού, ώστε να υλοποιηθεί η προσομοίωση. Η Bootstrap δεν προϋποθέτει γνώση της κατανομής του πληθυσμού, αφού η λογική της βασίζεται στην προέγγιση της άγνωστης κατανομής του πληθυσμού από την εμπειρική. Έτσι η μοναδική πληροφορία που χρειάζεται είναι το τυχαίο δείγμα, από το οποίο παράγονται και τα Bootstrap δείγματα. Λόγω του ότι χρησιμοποιεί αυτή την προσέγγιση, μπορεί να χρησιμοποιηθεί όχι μόνο σε ελέγχους υποθέσεων αλλά στην προσέγγιση οποιασδήποτε κατανομής είτε με πολύπλοκη συναρτησιακή μορφή, είτε για τον υπολογισμό διάφορων μεγεθών της κατανομής.

6.2 Περιγραφή Ελέγχου

Αρχικά πριν γίνει η περιγραφή του ελέγχου, πρέπει να οριστεί η εμπειρική κατανομή. Έστω ένα τυχαίο δείγμα μεγέθους n . Η εμπειρική κατανομή είναι η κατανομή που δίνει πιθανότητα $\frac{1}{n}$ σε κάθε παρατήρηση του δείγματος και 0 σε οποιαδήποτε άλλη τιμή, δηλαδή σε οποιαδήποτε τιμή εκτός δείγματος. Η κατασκευή της εμπειρικής κατανομής γίνεται μέσω της παραγωγής δειγμάτων με δειγματοληψία από το αρχικό με επανάθεση. Αυτό σημαίνει βέβαια ότι μία παρατήρηση μπορεί να επιλεγεί δύο ή περισσότερες φορές σε ένα δείγμα. Παρόλ' αυτά η πιθανότητα της

να επιλεγεί είναι $\frac{1}{n}$. Επιπλέον όσο αυξάνεται ο αριθμός των δειγμάτων τόσο θα πλησιάζει η συχνότητα της κάθε παρατήρησης αυτή που υποδηλώνει η εμπειρική κατανομή.

Αφού παραχθούν k Bootstrap δείγματα, ακολουθείται η διαδικασία που περιγράφηκε και στους ελέγχους τυχαιοποίησης. Ενδεικτικά τα βήματα της μεθόδου είναι τα εξής:

- Βήμα 1: Ορισμός H_0 και H_1 .
- Βήμα 2: Επιλογή κατάλληλης συνάρτησης ελέγχου T .
- Βήμα 3: Κατασκευή k δειγμάτων μεγέθους n , όπου n το μέγεθος του αρχικού δείγματος, με δειγματοληψία με επανάθεση από το αρχικό δείγμα.
- Βήμα 4: Υπολογισμός της T για κάθε δείγμα και του t_{obs} από το αρχικό.
- Βήμα 5: Υπολογισμός p-value ως εξής:

$$\text{p-value} = \frac{m + 1}{k + 1},$$

όπου m ο αριθμός των δειγμάτων για τα οποία ισχύει $T \geq t_{obs}$, $T \leq t_{obs}$ ή $|T| \geq t_{obs}$, ανάλογα με τη μορφή της εναλλακτικής υπόθεσης.

6.3 Προβλήματα της Μεθόδου

Η Bootstrap μπορεί να φαίνεται αρκετά κατάλληλη και απλή, αφού η μόνη πληροφορία που χρειάζεται ο ερευνητής για να την υλοποιήσει είναι το τυχαίο δείγμα, παρόλ' αυτά σε μερικές περιπτώσεις μπορεί να αποτύχει.

- Η μέθοδος αποτυγχάνει σε πολύ μικρά δείγματα. Αυτό συμβαίνει γιατί στα μικρά δείγματα το δειγματοληπτικό σφάλμα είναι αρκετά μεγάλο, και συνεπώς με την συγκεκριμένη μέθοδο και τα δείγματα που προκύπτουν θα το περιέχουν. Επιπλέον στην περίπτωση που οι παρατηρήσεις είναι λίγες, είναι πιθανό να επιλεγεί πολλές φορές το ίδιο δείγμα. Επομένως η προσέγγιση της πραγματικής κατανομής από την εμπειρική δεν θα είναι αντιπροσωπευτική.
- Η μέθοδος αποτυγχάνει επίσης όταν χρειάζεται να προσεγγιστεί κάποια ακραία τιμή. Αυτό συμβαίνει γιατί σε αυτές τις περιπτώσεις η δεδομένη πληροφορία αφορά μία συγκεκριμένη τιμή και άρα το δείγμα δεν είναι αντιπροσωπευτικό.

Παρόλ' αυτά στη βιβλιογραφία αναφέρονται τροποποιήσεις της μεθόδου, οι οποίες μπορούν να ξεπεράσουν αυτά τα προβλήματα. Για παράδειγμα χρησιμοποιώντας την παραμετρική μέθοδο Bootstrap εισάγονται πληροφορίες για την πραγμα-

τική κατανομή και έπειτα μπορεί να γίνει η διαδικασία που αναφέρθηκε στο υποκεφάλαιο 6.2. Σημειώνεται ότι οι έλεγχοι Monte Carlo είναι παραμετρικοί Bootstrap αφού εισάγουν την πρόσθετη πληροφορία της κατανομής του πληθυσμού.

6.4 Ενδιαφέρουσες Περιπτώσεις

6.4.1 Έλεγχος της Μέσης Τιμής Ενός Πληθυσμού

Έστω ότι χρειάζεται να ελεγχθεί η μέση τιμή μ ενός πληθυσμού, αν είναι ίση με μ_0 δεδομένου ενός δείγματος X_i , μεγέθους n , άγνωστης κατανομής. Η μηδενική υπόθεση θα είναι η $H_0: \mu = \mu_0$, ενώ η εναλλακτική θα διαφέρει ανάλογα με το είδος του ελέγχου, μονόπλευρος ή αμφίπλευρος. Διακρίνονται οι εξής περιπτώσεις:

- $H_1: \mu > \mu_0$
- $H_1: \mu < \mu_0$
- $H_1: \mu \neq \mu_0$,

Σε αυτή την περίπτωση προφανώς δεν μπορεί να χρησιμοποιηθεί ο έλεγχος Monte Carlo, όπως εξηγήθηκε στο Κεφάλαιο 5, αφού προϋποθέτει γνώση της κατανομής του πληθυσμού, για να προσομοιωθούν από αυτήν τα απαραίτητα δείγματα. Η δημιουργία δειγμάτων θα γίνει μόνο από το δοσμένο δείγμα. Η δημιουργία των δειγμάτων βασίζεται στην τυχαία δειγματοληψία με επανάθεση. Προκύπτουν λοιπόν νέα δείγματα που, όταν το πλήθος τους είναι πολύ μεγάλο, μπορούν να προσομοιώσουν χαρακτηριστικές ποσότητες για το αρχικό δείγμα και επομένως για τον πληθυσμό, από τον οποίο προέρχεται το δείγμα.

Πρέπει να δοθεί προσοχή στο γεγονός ότι μπορεί να χρειαστεί να τροποποιηθεί το αρχικό δείγμα ανάλογα με την δοσμένη μηδενική υπόθεση. Αν το αρχικό δείγμα δεν έχει μέση τιμή με μεγάλη απόκλιση από την μ_0 τότε μετακινούνται όλες οι παρατηρήσεις, έτσι ώστε να έχει μέση τιμή μ_0 , για να προσομοιωθούν τα δείγματα από δείγμα με την συγκεκριμένη μέση τιμή. Είναι σημαντικό να αναφερθεί πως μια τέτοια μετακίνηση θα προκαλέσει αλλαγή στη μέση τιμή αλλά όχι στα υπόλοιπα χαρακτηριστικά του δείγματος, για παράδειγμα στη διακύμανση.

Σχετικά με την επιλογή της συνάρτησης ελέγχου T , είναι σημαντικό να ελεγχθεί μια συνάρτηση που παίρνει γνωστές τιμές όταν ισχύει η μηδενική υπόθεση. Στο συγκεκριμένο παράδειγμα επιλέγεται η

$$T = |\bar{x} - \mu_0|$$

για αμφίπλευρο έλεγχο, και η

$$T = \bar{x} - \mu_0$$

για μονόπλευρο έλεγχο. Οι περιοχές απόρριψης διαμορφώνονται ανάλογα με τη μορφή της εναλλακτικής υπόθεσης. Είναι προφανές ότι όταν ισχύει η H_0 , η T παίρνει τιμές κοντά στο 0.

Επειδή η T προσεγγίζεται από τα προσομοιωμένα δείγματα, και δεν χρειάζεται γνώση της κατανομής της για να οριστεί η περιοχή απόρριψης, μπορεί να επιλεγεί οποιαδήποτε συνάρτηση, είτε απλή είτε σύνθετη.

Αφού υπολογιστεί η T για κάθε δείγμα υπολογίζεται το p-value και απορρίπτεται ή όχι η H_0 βάσει όσων αναφέρθηκαν στο υποκεφάλαιο 6.2.

Ο έλεγχος για τη μέση τιμή ενός πληθυσμού με τη μέθοδο *Bootstrap* υλοποιείται στο υποκεφάλαιο 7.2.

6.4.2 Έλεγχος για Γραμμική Συσχέτιση Δύο Μεταβλητών

Η μέθοδος *Bootstrap* μπορεί να εφαρμοστεί και σε έλεγχο σχετικά με τη γραμμική συσχέτιση δύο μεταβλητών. Για την υλοποίηση του υπάρχουν δύο προσεγγίσεις.

Η πρώτη μοιάζει με τους ελέγχους τυχαιοποίησης. Η λογική της βασίζεται στο γεγονός ότι όλοι οι συνδυασμοί των παρατηρήσεων των δύο μεταβλητών έχουν την ίδια πιθανότητα να εμφανιστούν, αν η μηδενική υπόθεση είναι αληθής. Η διαφορά στη μέθοδο *bootstrap* με τη μέθοδο της τυχαιοποίησης είναι πως δεν μένει σταθερή η μία μεταβλητή και ανακατεύονται οι παρατηρήσεις της δεύτερης, αλλά μένει σταθερή η μία μεταβλητή και στη δεύτερη πραγματοποιείται τυχαία δειγματοληψία με επανάθεση και δημιουργούνται τα νέα ζεύγη παρατηρήσεων.

Ο υπολογισμός του p-value είναι ίδιος με τους ελέγχους τυχαιοποίησης, αφού υπολογιστεί για κάθε νέο δείγμα ο συντελεστής συσχέτισης που έχει επιλεγεί σαν συνάρτηση ελέγχου.

Η δεύτερη προσέγγιση προϋποθέτει την υπόθεση ότι η σχέση που συνδέει τις μεταβλητές X, Y είναι η

$$Y = aX + \beta. \tag{6.1}$$

Έστω (X, Y) τυχαίο δείγμα μεγέθους n .

Απαιτείται αρχικά η εκτίμηση, με χρήση κάποιου μοντέλου, των παραμέτρων a, β . Έστω οι εκτιμώμενες παράμετροι \hat{a} και $\hat{\beta}$. Στη συνέχεια υπολογίζονται οι

ποσότητες

$$e_i = Y_i - \hat{\alpha}x_i - \hat{\beta},$$

που ονομάζονται κατάλοιπα. Δημιουργούνται τα bootstrap δείγματα ως εξής: Οι παρατηρήσεις x_i παραμένουν ίδιες και οι y_i γίνονται $y_i^* = \hat{\alpha}x_i - \hat{\beta} + e_i^*$, όπου τα e_i^* προκύπτουν με δειγματοληψία με επανάθεση από τα e_i .

Για τα παραπάνω δείγμα μπορεί να πραγματοποιηθεί στατιστική συμπερασματολογία σχετικά με την γραμμική συσχέτιση των μεταβλητών X, Y όπως αναφέρθηκε παραπάνω.

6.5 Συμπεράσματα

Συνοψίζοντας, οι έλεγχοι Bootstrap είναι και αυτοί έλεγχοι τυχαιοποίησης. Προτιμώνται από τους ελέγχους Monte Carlo γιατί δεν απαιτείται η γνώση κάποιας κατανομής για τον πληθυσμό. Οι απαραίτητες πληροφορίες δίνονται αποκλειστικά και μόνο από το παρατηρούμενο δείγμα. Οι έλεγχοι Bootstrap μπορούν να εφαρμοστούν για την προσέγγιση οποιασδήποτε συνάρτησης, όχι μόνο σε ελέγχους υποθέσεων.

Παρόλα αυτά χρειάζεται προσοχή στο μέγεθος του δείγματος, το οποίο δεν πρέπει να είναι μικρό, γιατί τότε δεν θα δίνεται επαρκής πληροφορία στον έλεγχο.

Κεφάλαιο 7

Εφαρμογές στην R

7.1 Εισαγωγή

Στο παρόν κεφάλαιο θα παρουσιαστεί η υλοποίηση των περισσότερων μεθόδων που αναφέρθηκαν στα προηγούμενα κεφάλαια, κατηγοριοποιημένες ανάλογα με το είδος του ελέγχου. Για την υλοποίηση χρησιμοποιείται η γλώσσα προγραμματισμού R και το λογισμικό RStudio. Αυτή η γλώσσα προγραμματισμού είναι κατάλληλη τόσο για στατιστικούς υπολογισμούς όσο και για την οπτικοποίηση των δεδομένων.

Τα δεδομένα που χρησιμοποιούνται στο παρόν κεφάλαιο είναι πλασματικά και δεν προκύπτουν από κάποια πειραματική διαδικασία. Τα δεδομένα χρησιμοποιούνται μόνο για να υλοποιηθούν οι στατιστικές μέθοδοι. Παρόλα αυτά, οι κώδικες που χρησιμοποιούνται μπορούν να εφαρμοστούν σε πραγματικά δεδομένα.

7.2 Έλεγχος για τη Μέση Τιμή Ενός Πληθυσμού

Έστω το εξής πρόβλημα: «Σε ένα νοσοκομείο μετρήθηκε ο αριθμός των επειγόντων περιστατικών ανά ημέρα σε διάστημα 100 ημερών. Ζητείται να ελεγχθεί αν κατά μέσο όρο τα επείγοντα περιστατικά σε αυτό το νοσοκομείο είναι 1 ή περισσότερα.»

Ορίζονται οι εξής υποθέσεις:

- $H_0: \mu = 1$
- $H_1: \mu > 1$

Η μεταβλητή, για την οποία μελετάται η συμπεριφορά της είναι η X : αριθμός επειγόντων περιστατικών ανά ημέρα στο συγκεκριμένο νοσοκομείο. Επιλέγεται

επίπεδο σημαντικότητας $\alpha = 0.05$.

Προκειμένου να χρησιμοποιηθούν στατιστικές μέθοδοι και συνάρτησεις είναι απαραίτητη η βιβλιοθήκη BSDA στην R, η οποία εγκαθίσταται στο τερματικό παράθυρο με την εντολή 7.1:

```
1 install.packages("BSDA")
```

Κώδικας 7.1: Εγκατάσταση βιβλιοθήκης BSDA

και εισάγεται στον κώδικα με την εντολή 7.2, στην αρχή του προγράμματος.

```
1 library(BSDA)
```

Κώδικας 7.2: Ενσωμάτωση της BSDA στον κώδικα

Τα δεδομένα που συλλέχθηκαν, ομαδοποιήθηκαν και παρουσιάζονται στον Πίνακα 7.1:

Αριθμός επειγόντων περιστατικών	Αριθμός ημερών
0	34
1	27
2	18
3	13
4	5
5	3

Πίνακας 7.1: Αριθμός επειγόντων περιστατικών ενός νοσοκομείου σε διάστημα 100 ημερών

Αρχικά δημιουργείται ένα διάγραμμα των παρατηρήσεων και επισημαίνεται η μέση τιμή του δείγματος για την καλύτερη εποπτεία των δεδομένων. Το Διάγραμμα 7.1 είναι το διάγραμμα των δεδομένων του Πίνακα 7.1 με επισημασμένη τη μέση τιμή.

Ο πρώτος έλεγχος που πραγματοποιείται είναι ο παραμετρικός για μεγάλο δείγμα, αφού ο αριθμός των παρατηρήσεων είναι 100 (>30), μετά από υπόθεση κανονικότητας της μεταβλητής.

```
1 #parametric
2 z_test <- z.test(data, mu = m0, sigma.x = sd)
```

Κώδικας 7.4: Παραμετρικός έλεγχος για έναν πληθυσμό

Μετά την εκτέλεση της εντολής 7.4 υπολογίζεται το p-value, το οποίο είναι 0.00654585. Συνεπώς η μηδενική υπόθεση απορρίπτεται. Υπενθυμίζεται ότι η μηδενική υπόθεση απορρίπτεται όταν $p\text{-value} < \alpha$.

Ο επόμενος έλεγχος που υλοποιείται είναι η μέθοδος Monte Carlo. Για τη χρήση της μεθόδου απαιτείται η κατανομή που ακολουθεί η μεταβλητή, προκειμένου να προσομοιωθούν από αυτή δείγματα. Είναι γνωστό στη στατιστική ότι όταν μία μεταβλητή αποτυπώνει τον αριθμό των γεγονότων ανά συγκεκριμένο χρονικό διάστημα ακολουθεί Poisson. Για τον ορισμό της σ.μ.π. της Poisson απαιτείται ο προσδιορισμός της παραμέτρου της. Η εκτιμήτρια μέγιστης πιθανοφάνειας της παραμέτρου της Poisson είναι ο δειγματικός μέσος. Συνεπώς θα χρησιμοποιηθεί η τιμή του, που έχει υπολογιστεί από το δοσμένο δείγμα.

Στον κώδικα 7.5 προσομοιώνονται από την Poisson 99 δείγματα, μεγέθους 100 με παράμετρο 1.37.

```
1 #Create 99 samples from Poisson, parameter = 1.37
2 l <- 1.37
3 num_samplesMC <- 99
4 samplesMC <- replicate(num_samplesMC, rpois(n, l))
```

Κώδικας 7.5: Δημιουργία 99 δειγμάτων από την Poisson

Είναι απαραίτητο να οριστεί μία συνάρτηση ελέγχου που θα διαχωρίζει ανάμεσα στις δύο υποθέσεις. Επιλέγεται η

$$T = \bar{x} - l,$$

η οποία, όταν ισχύει η μηδενική υπόθεση παίρνει τιμές κοντά στο 0, ενώ η μηδενική υπόθεση απορρίπτεται όταν οι τιμές της είναι πολύ μεγάλες. Η παρατηρούμενη τιμή της συνάρτησης ελέγχου t_{obs} , δηλαδή η τιμή που προκύπτει από το δείγμα είναι $t_{obs} = 0.37$. Στη συνέχεια για κάθε ένα από τα παραπάνω 99 προσομοιωμένα δείγματα υπολογίζεται η τιμή της T και το πλήθος των τιμών αυτών που είναι μεγαλύτερες από την παρατηρούμενη. Προκύπτει, λοιπόν, το p-value, με τη χρήση του Κώδικα 7.6, το οποίο είναι ίσο με 0.49.

```

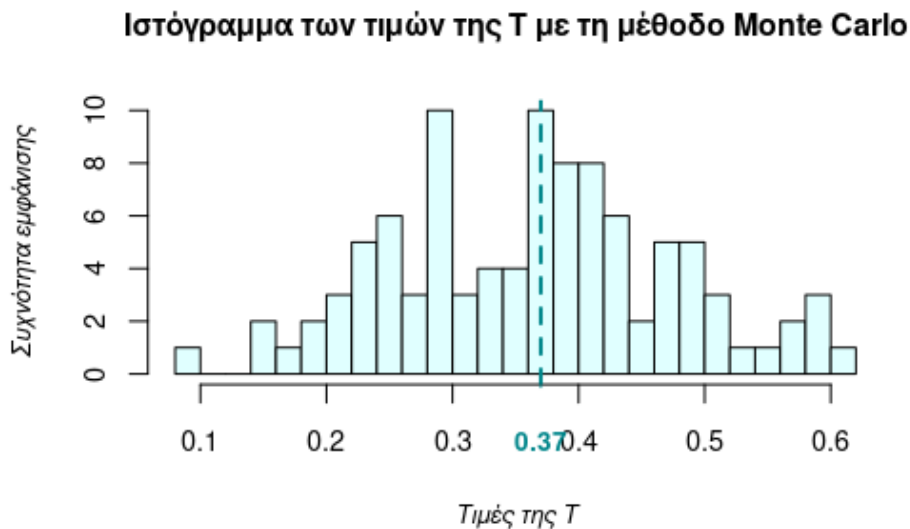
1 #Compute T for each sample
2 tobs <- m-1
3 result <- lapply(1:ncol(samplesMC), function(i){
4     mean(samplesMC[, i]) - 1
5 })
6 #Compute p-value
7 t_values_MC <- unlist(result)
8 greater_obs_MC <- sum(t_values_MC > tobs)
9 pValueMC <- (greater_obs_MC + 1)/(num_samplesMC + 1)

```

Κώδικας 7.6: Υπολογισμός T και p-value για τα δείγματα της Poisson

Συνεπώς με τη χρήση αυτής της μεθόδου η μηδενική υπόθεση δεν απορρίπτεται, σε επίπεδο σημαντικότητας $\alpha = 0.05$.

Είναι επίσης βοηθητική η χρήση ενός ιστογράμματος με τις τιμές της συνάρτησης ελέγχου, προκειμένου να υπάρξει εποπτεία της συμπεριφοράς της. Το Διάγραμμα 7.2 παρουσιάζει τις τιμές της συνάρτησης ελέγχου που προέκυψαν από τα προσομοιωμένα δείγματα με τη μέθοδο Monte Carlo, αλλά και την παρατηρούμενη τιμή της.



Σχήμα 7.2: Ιστόγραμμα των τιμών της συνάρτησης ελέγχου από την προσομοίωση δειγμάτων με Monte Carlo για 1 πληθυσμό

Παρατηρείται πως με χρήση περισσότερων δειγμάτων, π.χ. 999, το p-value δεν διαφοροποιείται δραματικά.

Η τελευταία μέθοδος που θα υλοποιηθεί είναι η Bootstrap. Για την μέθοδο Bootstrap δεν απαιτείται η γνώση της κατανομής που ακολουθεί η μεταβλητή. Αρκεί η δημιουργία, έστω 99, δειγμάτων, μεγέθους 100, καθένα από τα οποία προκύπτει από την τυχαία επιλογή με επανάθεση των παρατηρήσεων από το αρχικό δείγμα. Η δημιουργία των δειγμάτων Bootstrap φαίνεται στον κώδικα 7.7.

```
1 #Define number of bootstrap samples
2 num_samples_bootstrap <- 99
3 #Create new samples
4 new_samples_bootstrap <- replicate(num_samples_bootstrap, sample(data
5 , n, replace = TRUE))
6 result_bootstrap <- lapply(1:ncol(new_samples_bootstrap), function(i
7 ) {
8   mean(new_samples_bootstrap[, i]) - 1
9 })
```

Κώδικας 7.7: Bootstrap για τη μέση τιμή ενός πληθυσμού

Η επιλογή της συνάρτησης, η παρατηρούμενη τιμή της και το p-value υπολογίζονται όπως στη μέθοδο Monte Carlo. Διαφοροποιείται μόνο ο τρόπος δημιουργίας των δειγμάτων. Προκύπτει λοιπόν p-value ίσο με 0.43 όπως φαίνεται στον Κώδικα 7.8, και επομένως η μηδενική υπόθεση δεν απορρίπτεται.

```
1 #Compute p-value
2 t_values_bootstrap <- unlist(result_bootstrap)
3 greater_obs_bootstrap <- sum(t_values_bootstrap > 0.37)
4 pValueBootstrap <- (greater_obs_bootstrap + 1)/(num_samples_bootstrap
5 + 1)
```

Κώδικας 7.8: Υπολογισμός p-value για τα bootstrap δείγματα για τη μέση τιμή ενός πληθυσμού

Το διάγραμμα των τιμών της συνάρτησης ελέγχου που προέκυψαν από τα προσομοιωμένα δείγματα με τη μέθοδο Monte Carlo, αλλά και η παρατηρούμενη τιμή της συνάρτησης ελέγχου είναι το Διάγραμμα 7.3.

Η μέθοδος Monte Carlo και η μέθοδος Bootstrap θα μπορούσαν να επαναληφθούν πολλές φορές προκειμένου να προκύψουν πολλές τιμές για το p-value στο οποίο θα μπορούσε να γίνει στατιστική συμπερασματολογία.



Σχήμα 7.3: Ιστογράμμα των τιμών της συνάρτησης ελέγχου από την προσομοίωση δειγμάτων με τη μέθοδο Bootstrap για 1 πληθυσμό

Είναι λογικό η απόρριψη ή μη της μηδενικής υπόθεσης να διαφοροποιείται ανάμεσα στον παραμετρικό έλεγχο και στους ελέγχους Monte Carlo και Bootstrap, αφού για την χρήση του παραμετρικού ελέγχου παίρνονται σαν δεδομένο υποθέσεις κανονικότητας, οι οποίες δεν είναι γνωστό αν πληρούνται. Συνεπώς το αποτέλεσμα των μεθόδων Monte Carlo και Bootstrap μπορεί να θεωρηθεί πιο αξιόπιστο.

Συνοπτικά τα αποτελέσματα για τον έλεγχο της μέσης τιμής ενός πληθυσμού παρουσιάζονται στον πίνακα 7.2.

Μέθοδος	p-value	Απόφαση
z-test	0.00655	Απόρριψη της H_0
Monte Carlo	0.49	Όχι απόρριψη της H_0
Bootstrap	0.43	Όχι απόρριψη της H_0

Πίνακας 7.2: Σύνοψη αποτελεσμάτων για τον έλεγχο της μέσης τιμής ενός πληθυσμού

7.3 Έλεγχος της Διαφοράς των Μέσων Δύο Πληθυσμών

Έστω το εξής πρόβλημα: «Δύο αντικαρκινικά φάρμακα A και B χορηγήθηκαν σε 6 ασθενείς το καθένα, και μετρήθηκε το ποσοστό εξάλειψης των καρκινικών κυττάρων. Τίθεται το ερώτημα αν τα δύο φάρμακα είναι ισοδύναμα.»

Η μεταβλητή η οποία μελετάται είναι η X : ποσοστό εξάλειψης καρκινικών κυττάρων. Επιλέγεται επίπεδο σημαντικότητας $\alpha = 0.05$. Οι πληθυσμοί, οι οποίοι εξετάζονται, είναι οι πληθυσμοί που ακολουθούν τα φάρμακα A και B .

Τα δεδομένα που συλλέχθηκαν παρουσιάζονται στον Πίνακα 7.3.

Αρχικά ορίζονται οι δύο υποθέσεις:

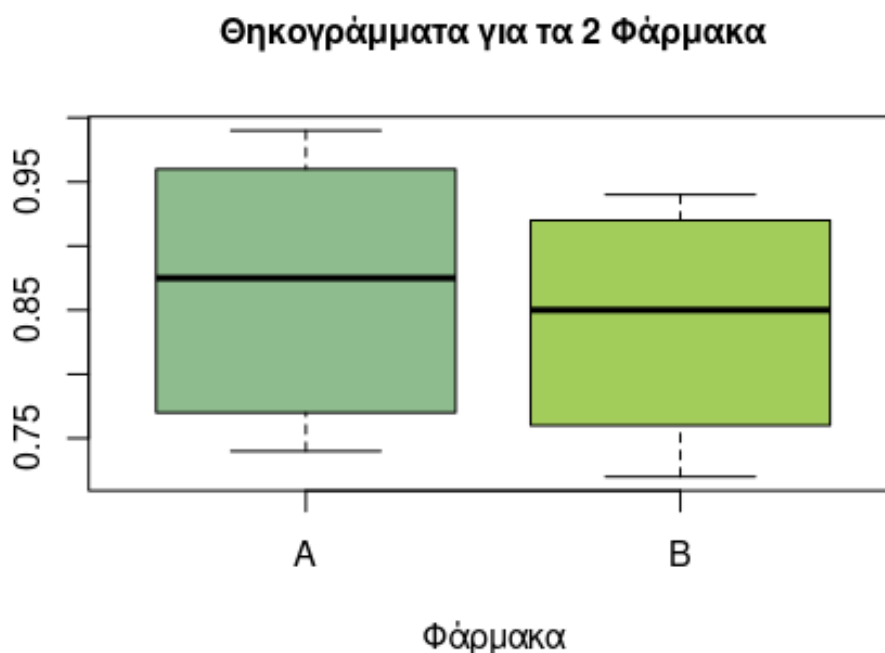
- $H_0: \mu_A = \mu_B$
- $H_1: \mu_A \neq \mu_B$,

όπου μ_A, μ_B οι μέσες τιμές των κατανομών των πληθυσμών που ακολουθούν τα φάρμακα A και B αντίστοιχα.

Φάρμακο A	0.96	0.81	0.77	0.99	0.94	0.74
Φάρμακο B	0.87	0.92	0.94	0.83	0.72	0.76

Πίνακας 7.3: Ποσοστό εξάλειψης καρκινικών κυττάρων για δύο τύπους φαρμάκων

Προκειμένου να ελεγχθούν τα δύο φάρμακα εποπτικά δημιουργείται το Διάγραμμα 7.4:



Σχήμα 7.4: Απεικόνιση των δεδομένων των δύο φαρμάκων

Το Διάγραμμα 7.4 είναι ένα θηκογράμμα (boxplot) και η μαύρη οριζόντια γραμμή εντός των κουτιών δείχνει τη διάμεσο. Δεν μπορεί να προκύψει άμεσο συμπέρασμα για τη μέση τιμή από ένα θηκογράμμα, παρόλα αυτό είναι γνωστό πως όταν αυτό είναι συμμετρικό, η διάμεσος και η μέση τιμή είναι περίπου ίσες.

Στη συγκεκριμένη περίπτωση τα δύο θηκογράμματα είναι συμμετρικά, και όπως προκύπτει από τις διαμέσους, οι δύο τιμές είναι αρκετά κοντά η μία στην άλλη.

Στη συνέχεια εισάγονται τα δεδομένα με τη χρήση του Κώδικα 7.9:

```
1 A <- c(0.96,0.81,0.77,0.99,0.94,0.74)
2 B <- c(0.87,0.92,0.94,0.83,0.72,0.76)
```

Κώδικας 7.9: Εισαγωγή δεδομένων φαρμάκων

Ο πρώτος έλεγχος που πραγματοποιείται είναι ο παραμετρικός, υπό την υπόθεση της ανεξαρτησίας, όπως φαίνεται τον Κώδικα 7.10, από τον οποίο προκύπτει p-value ίσο με 0.6279.

```

1 #parametric t-test
2 t_test <- t.test(A,B)

```

Κώδικας 7.10: Παραμετρικός έλεγχος δύο πληθυσμών

Συνεπώς σύμφωνα με την παραπάνω τιμή, η μηδενική υπόθεση δεν απορρίπτεται.

Ο επόμενος έλεγχος είναι ο μη παραμετρικός Wilcoxon Rank Sum test όπως φαίνεται στον Κώδικα 7.11, από τον οποίο υπολογίζεται p-value ίσο με 0.513.

```

1 #non parametric
2 wilcoxon_test <- wilcox.exact(A,B)

```

Κώδικας 7.11: Μη παραμετρικός έλεγχος δύο πληθυσμών

Και σε αυτόν τον έλεγχο η μηδενική υπόθεση δεν απορρίπτεται.

Έπειτα πραγματοποιείται ένας έλεγχος τυχαιοποίησης. Ο έλεγχος είναι προσεγγιστικός. Οι δυνατοί συνδυασμοί των 12 παρατηρήσεων σε 2 ομάδες των 6 παρατηρήσεων η κάθε μια είναι $\binom{12}{6} = 5544$. Επιλέγονται έστω 99 πιθανοί συνδυασμοί. Η επιλογή της συνάρτησης ελέγχου προκύπτει από τη φύση των υποθέσεων. Μια συνάρτηση ελέγχου της μορφής

$$T = |\bar{x}_A - \bar{x}_B|,$$

όπου \bar{x}_A , \bar{x}_B οι δειγματικοί μέσοι των δειγμάτων A και B αντίστοιχα, είναι κατάλληλοι, αφού αν ισχύει η μηδενική υπόθεση παίρνει τιμές πολύ κοντά στο 0 ενώ μεγάλες τιμές υποδηλώνουν απόκλιση από τη μηδενική υπόθεση.

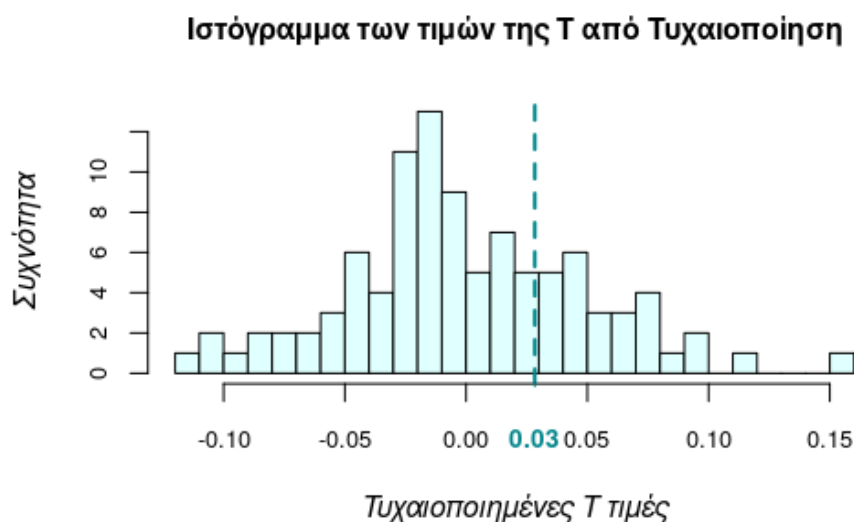
Ο Κώδικας 7.12 παρουσιάζει τον υπολογισμό της συνάρτησης ελέγχου των δοσμένων δειγμάτων, την δημιουργία 99 συνδυασμών, την κατασκευή της συνάρτησης που υπολογίζει για κάθε συνδυασμό την συνάρτηση ελέγχου και τον υπολογισμό του p-value.

Από τον παραπάνω έλεγχο προκύπτει p-value ίσο με 0.27, και επομένως η μηδενική υπόθεση δεν απορρίπτεται.

Στο Διάγραμμα 7.5 απεικονίζονται οι τιμές της συνάρτησης ελέγχου που προέκυψαν από τα προσομοιωμένα δείγματα με τη μέθοδο της τυχαιοποίησης, αλλά και η παρατηρούμενη τιμή της ελεγχουσυνάρτησης.


```
1  # Observed difference in means
2  observed_diff <- mean(A) - mean(B)
3  # Combine the data
4  combined_data <- c(A, B)
5  # Number of permutations
6  n_permutations <- 99
7
8  # Function to calculate the difference in means for permuted data
9  perm_diff <- function(data, n1, n2) {
10  permuted_data <- sample(data)
11  groupA <- permuted_data[1:n1]
12  groupB <- permuted_data[(n1 + 1):(n1 + n2)]
13  return(mean(groupA) - mean(groupB))
14  }
15
16  set.seed(123) # For reproducibility
17  rand_diffs <- replicate(n_permutations, perm_diff(combined_data,
18  length(A), length(B)))
19
20  # Calculate p-value
21  greater_obs_rand <- sum(rand_diffs > observed_diff)
22  pValueRand <- (greater_obs_rand + 1) / (n_permutations + 1)
```

Κώδικας 7.12: Έλεγχος τυχαιοποίησης δύο πληθυσμών και υπολογισμός p-value



Σχήμα 7.5: Ιστογράμμα των τιμών της συνάρτησης ελέγχου με τη μέθοδο της τυχαιοποίησης για 2 πληθυσμούς

Ο τελευταίος έλεγχος που πραγματοποιείται γίνεται με τη μέθοδο Bootstrap. Η λογική του βασίζεται στην τυχαία επιλογή με επανάθεση από τα δύο δείγματα, με την κάθε παρατήρηση να διατηρεί το δείγμα από το οποίο προέρχεται. Η υλοποίηση του και ο υπολογισμός του p-value φαίνεται στον Κώδικα 7.13.

```

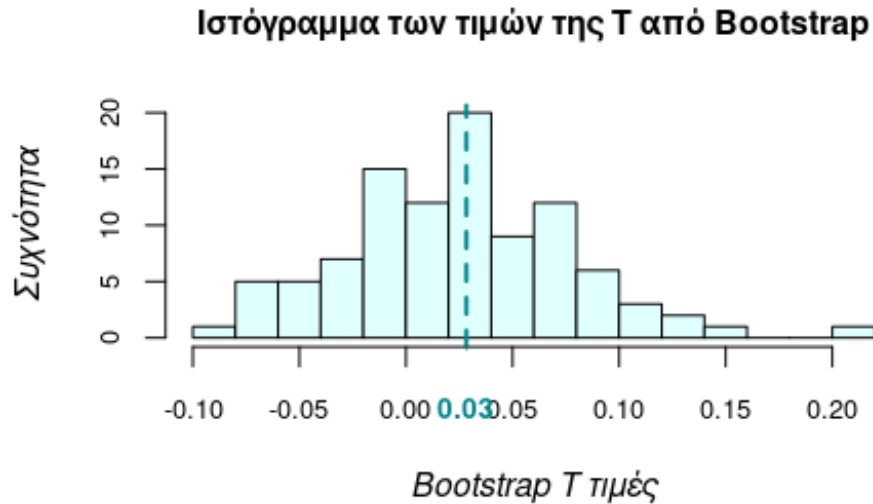
1  # Number of bootstrap samples
2  n_bootstrap <- 99
3
4  # Function to calculate the difference in means
5  bootstrap_diff <- function(data1, data2) {
6  sample1 <- sample(data1, length(data1), replace = TRUE)
7  sample2 <- sample(data2, length(data2), replace = TRUE)
8  return(mean(sample1) - mean(sample2))
9  }
10 # Generate bootstrap samples
11 bootstrap_diffs <- replicate(n_bootstrap, bootstrap_diff(A, B))
12
13 # Calculate p-value
14 greater_obs_boot <- sum(bootstrap_diffs > observed_diff)
15 pValueBoot <- (greater_obs_boot + 1) / (n_bootstrap + 1)

```

Κώδικας 7.13: Έλεγχος Bootstrap δύο πληθυσμών και υπολογισμός p-value

Στο Διάγραμμα 7.6 απεικονίζονται οι τιμές της συνάρτησης ελέγχου που προέκυψαν

από τα προσομοιωμένα δείγματα με τη μέθοδο Bootstrap και η παρατηρούμενη τιμή της ελεγχουσυνάρτησης.



Σχήμα 7.6: Ιστόγραμμα των τιμών της συνάρτησης ελέγχου με τη μέθοδο Bootstrap για 2 πληθυσμούς

Συνοπτικά τα αποτελέσματα για τον έλεγχο της διαφοράς των μέσων τιμών δύο πληθυσμών παρουσιάζονται στον Πίνακα 7.4.

Μέθοδος	p-value	Απόφαση
t-test	0.6279	Όχι απόρριψη της H_0
Wilcoxon Rank Sum	0.513	Όχι απόρριψη της H_0
Τυχαιοποίησης	0.27	Όχι απόρριψη της H_0
Bootstrap	0.48	Όχι απόρριψη της H_0

Πίνακας 7.4: Σύνοψη αποτελεσμάτων για τον έλεγχο της διαφοράς των μέσων τιμών δύο πληθυσμών

Το γεγονός ότι οι δύο πρώτοι έλεγχοι, παραμετρικός και μη καταλήγουν στο ίδιο συμπέρασμα με τον έλεγχο τυχαιοποίησης και τον έλεγχο bootstrap, οι οποίοι είναι πιο αξιόπιστοι, είναι λογικό, αφού και από το Διάγραμμα 7.4 δεν προκύπτει σημαντική διαφορά στα δύο δείγματα και στις μέσες τιμές τους.

7.4 Έλεγχος της Διαφοράς των Μέσων Τριών Πληθυσμών

Έστω το εξής πρόβλημα: «Σε τρεις τύπους κυττάρων A, B και Γ μετρήθηκε το ποσοστό απορρόφησης ακτινοβολίας σε 12 ανθρώπους, 4 για το κάθε είδος κυττάρου. Ζητάται να μελετηθεί αν το ποσοστό απορρόφησης της ακτινοβολίας είναι ίδιο σε κάθε τύπο.» Τα δεδομένα παρουσιάζονται στον Πίνακα 7.5.

A	0.57	0.73	0.69	0.71
B	0.46	0.55	0.32	0.49
Γ	0.78	0.62	0.87	0.93

Πίνακας 7.5: Ποσοστό απορρόφησης ακτινοβολίας για τρεις τύπους κυττάρων

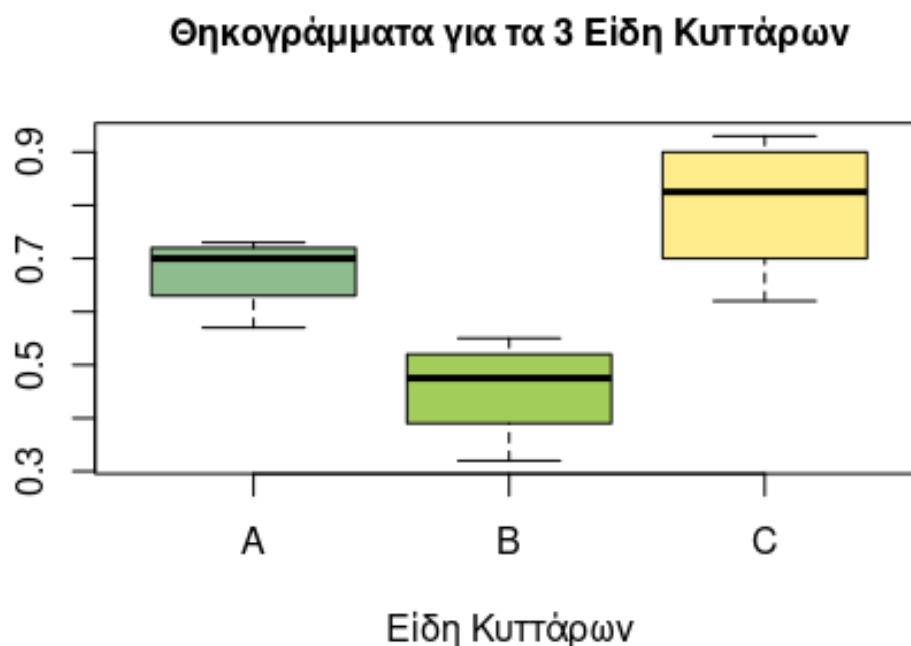
Η μεταβλητή, η οποία μελετάται είναι η X : ποσοστό απορρόφησης ακτινοβολίας. Οι 3 πληθυσμοί είναι οι πληθυσμοί από τους οποίους προέρχονται τα 3 είδη κυττάρων.

Ορίζονται αρχικά οι δύο υποθέσεις του ελέγχου:

- H_0 : Όλοι οι πληθυσμιακοί μέσοι είναι ίσοι
- H_1 : Τουλάχιστον ένας πληθυσμιακός μέσος είναι διαφορετικός από τους άλλους.

Ορίζεται επίσης το επίπεδο σημαντικότητας $\alpha = 0.05$.

Αρχικά δημιουργείται το Διάγραμμα 7.7, δηλαδή το διάγραμμα των παρατηρήσεων για την καλύτερη εποπτεία των δεδομένων.



Σχήμα 7.7: Απεικόνιση των δεδομένων των τριών τύπων κυττάρων

Τα θηκογράμματα του Διαγράμματος 7.7 είναι ασύμμετρα, και μάλιστα αριστερόστροφα ασύμμετρα. Σε ένα αριστερόστροφο ασύμμετρο θηκογράμμα η μέση τιμή είναι μικρότερη από τη διάμεσο. Επομένως και στα τρία διαγράμματα η μέση τιμή του κάθε δείγματος είναι μικρότερη από την οριζόντια μαύρη γραμμή. Είναι εμφανές ότι οι μέσες τιμές των δειγμάτων διαφέρουν αρκετά η μία από την άλλη.

Στη συνέχεια εισάγονται τα δεδομένα στο πρόγραμμα, και υπολογίζονται οι μέσες τιμές και οι διασπορές των δειγμάτων. Έπειτα προετοιμάζονται τα δεδομένα σε δομή data frame ώστε να μπορέσουν να υλοποιηθούν οι ακόλουθοι έλεγχοι. Η διαδικασία αυτή παρουσιάζεται στον Κώδικα 7.14.

```

1  A <- c(0.57,0.73,0.69,0.71)
2  B <- c(0.46,0.55,0.32,0.49)
3  C <- c(0.78,0.62,0.87,0.93)
4  mA <- mean(A)
5  mB <- mean(B)
6  mC <- mean(C)
7  varA <- var(A)
8  varB <- var(B)
9  varC <- var(C)
10 # Combine the data into a single vector

```

```

11 values <- c(A,B,C)
12 # Create a factor variable to indicate the group
13 groups <- factor(rep(c("A", "B", "C"), each = length(A)))
14 # Create a data frame
15 data <- data.frame(Group = groups, Value = values)

```

Κώδικας 7.14: Εισαγωγή δεδομένων κυττάρων και υπολογισμός χαρακτηριστικών μεγεθών των δειγμάτων

Ο πρώτος έλεγχος που πραγματοποιείται είναι ο παραμετρικός ANOVA αφού οι πληθυσμοί που ελέγχονται είναι 3. Η υλοποίηση του φαίνεται στον Κώδικα 7.15.

```

1 # Perform the ANOVA
2 anova_result <- aov(Value ~ Group, data = data)

```

Κώδικας 7.15: Παραμετρικός έλεγχος τριών πληθυσμών

Μετά την πραγματοποίηση του παραπάνω ελέγχου προκύπτει p-value ίσο με 0.00368 και επομένως η μηδενική υπόθεση απορρίπτεται.

Ο επόμενος έλεγχος είναι ο παραμετρικός Kruskal-Wallis και υλοποιείται όπως φαίνεται στον Κώδικα 7.16.

```

1 #Perform Kruskal–Wallis
2 kruskalwallis_result <- kruskal.test(values ~ groups, data = data)

```

Κώδικας 7.16: Μη παραμετρικός έλεγχος τριών πληθυσμών

Από τον έλεγχο αυτόν προκύπτει p-value ίσο με 0.0154 και επομένως η μηδενική υπόθεση απορρίπτεται ξανά.

Ο τελευταίος έλεγχος που πραγματοποιείται είναι το τυχαιοποιημένο ANOVA σε 1000 τυχαιοποιημένα δείγματα. Η υλοποίηση του φαίνεται στον Κώδικα 7.17.

```

1 # Perform permutation ANOVA
2 perm_test <- oneway_test(values ~ groups, data = data, distribution
  = approximate(nresample = 1000))

```

Κώδικας 7.17: Έλεγχος τυχαιοποίησης για τρεις πληθυσμούς

Με τον παραπάνω έλεγχο τυχαιοποίησης προκύπτει p-value ίσο με 0.001 και επομένως η μηδενική υπόθεση απορρίπτεται.

Όπως φαίνεται και από το Διάγραμμα 7.7 είναι λογική η απόρριψη της μηδενικής υπόθεσης.

Συνοπτικά τα αποτελέσματα για τον έλεγχο της διαφοράς μεταξύ τριών πληθυσμών παρουσιάζονται στον Πίνακα 7.6.

Μέθοδος	p-value	Απόφαση
ANOVA	0.00368	Απόρριψη της H_0
Kruskal-Wallis	0.0154	Απόρριψη της H_0
Τυχαιοποιημένη ANOVA	0.001	Απόρριψη της H_0

Πίνακας 7.6: Σύνοψη αποτελεσμάτων για τον έλεγχο της διαφοράς μεταξύ τριών πληθυσμών

7.5 Γραμμική Συσχέτιση

Έστω το παρακάτω πρόβλημα: «Μετρήθηκε η συστολική και η διαστολική πίεση ενός ασθενούς σε διάστημα 7 ημερών. Τίθεται το ερώτημα αν τα δύο είδη πίεσης σχετίζονται μεταξύ τους γραμμικά.»

Οι δύο μεταβλητές, οι οποίες εξετάζονται αν σχετίζονται γραμμικά είναι η X : συστολική πίεση και Y : διαστολική πίεση.

Ορίζονται οι δύο υποθέσεις:

- H_0 : Δεν υπάρχει γραμμική συσχέτιση μεταξύ των X και Y
- H_1 : Υπάρχει γραμμική συσχέτιση μεταξύ των X και Y

Επιλέγεται επίσης επίπεδο σημαντικότητας $\alpha = 0.05$.

Τα δεδομένα φαίνονται στον Πίνακα 7.7:

Ημέρα	1	2	3	4	5	6	7
Συστολική	131	162	124	149	161	133	156
Διαστολική	87	103	81	95	102	84	99

Πίνακας 7.7: Μετρήσεις Συστολικής και Διαστολικής Πίεσης σε διάστημα 7 ημερών

Εισάγονται αρχικά τα δεδομένα, όπως φαίνεται στον Κώδικα 7.18:

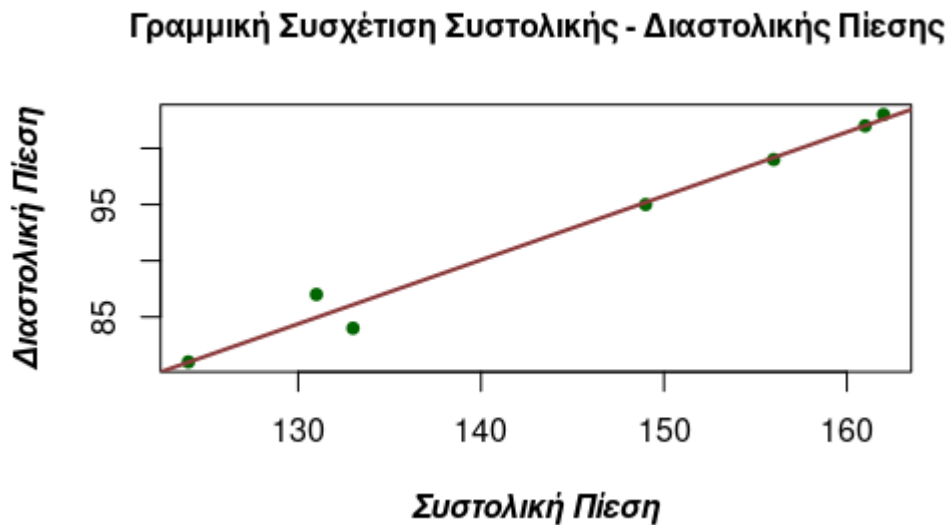
```

1 systolic <- c(131,162,124,149,161,133,156)
2 diastolic <- c(87,103,81,95,102,84,99)

```

Κώδικας 7.18: Εισαγωγή δεδομένων συστολικής και διαστολικής πίεσης

Προκειμένου να διαπιστωθεί γραφικά αν η σχέση των δύο μεταβλητών είναι γραμμική δημιουργείται το Διάγραμμα 7.8. Στο Διάγραμμα 7.8 παρουσιάζονται τα δεδομένα με τη μορφή σημείων και το γραμμικό μοντέλο με τη μορφή μίας ευθείας, το οποίο δημιουργήθηκε βάσει των παρατηρήσεων.



Σχήμα 7.8: Διάγραμμα Συστολικής - Διαστολικής Πίεσης

Είναι φανερό ότι η ευθεία του γραμμικού μοντέλου οριακά διαπερνά όλα τα σημεία του δείγματος. Επομένως, εποπτικά, μπορεί να προκύψει το συμπέρασμα ότι η σχέση τους είναι γραμμική και μάλιστα θετική, δηλαδή όσο αυξάνεται το ένα μέγεθος, αυξάνεται και το άλλο.

Υπολογίζεται αρχικά ο συντελεστής συσχέτισης Pearson, όπως φαίνεται στον Κώδικα 7.19.

```

1 #Calculate Pearson's Correlation Coefficient
2 corPearson <- cor(systolic , diastolic , method="pearson")

```

Κώδικας 7.19: Υπολογισμός συντελεστή συσχέτισης Pearson

Από τον παραπάνω κώδικα υπολογίζεται ο συντελεστής συσχέτισης Pearson ίσος με 0.99, το οποίο υποδηλώνει ισχυρή γραμμική συσχέτιση. Υπολογίζεται επι-

πλέον το p-value, το οποίο είναι ίσο με 0.0000154. Συνεπώς η μηδενική υπόθεση απορρίπτεται.

Ο επόμενος συντελεστής συσχέτισης, που υπολογίζεται και φαίνεται στον Κώδικα 7.20, είναι ο μη παραμετρικός Spearman. Υπολογίζεται ότι είναι ίσος με 0.964, το οποίο υποδηλώνει ξανά ισχυρή γραμμική συσχέτιση. Υπολογίζεται επίσης το p-value ίσο με 0.002778 και επομένως η μηδενική υπόθεση απορρίπτεται.

```
1 #Calculate Spearman's Correlation Coefficient
2 corSpearman <- cor.test(systolic , diastolic , method="spearman")
```

Κώδικας 7.20: Υπολογισμός συντελεστή συσχέτισης Spearman

Ο επόμενος έλεγχος που πραγματοποιείται είναι ο έλεγχος τυχαιοποίησης, και πιο συγκεκριμένα ένας προσεγγιστικός έλεγχος τυχαιοποίησης για τον οποίο θα χρησιμοποιηθεί ως συνάρτηση ελέγχου ο συντελεστής συσχέτισης Pearson και θα δημιουργηθούν 99 συνδυασμοί από τους 120. Όλοι οι δυνατοί συνδυασμοί προκύπτουν αν βρεθούν όλες οι δυνατές βάδες ζευγών από τα δεδομένα. Για να υπολογιστούν αρκεί να μείνει σταθερή η μία μεταβλητή και να βρεθούν όλες οι διατάξεις των 6 παρατηρήσεων της δεύτερης μεταβλητής. Όλες διατάξεις των 6 παρατηρήσεων υπολογίζονται ίσες με $6! = 120$. Για κάθε διάταξη υπολογίζεται ο συντελεστής συσχέτισης Pearson. Έπειτα υπολογίζεται πόσοι από τους 99 συνδυασμούς που επιλέχθηκαν τυχαία από τους 120 έχουν συντελεστή συσχέτισης που αποκλίνει από τον παρατηρούμενο συντελεστή συσχέτισης που υπολογίστηκε παραπάνω και υπολογίζεται το p-value. Η παραπάνω διαδικασία παρουσιάζεται στον Κώδικα 7.21.

```
1 # Randomization test.
2 nreps = 99 # Note: 99 combinations
3 N <- length(systolic)
4 r <- numeric(nreps)
5 counter = 0
6 set.seed <- 1086 #This will force each run to come to the same
7 result.
8 for (i in 1:nreps) {
9   randDiastolic <- sample(diastolic , N, replace = FALSE)
10  r[i] <- cor(systolic , randDiastolic)
11  if(r[i] >= corPearson) counter = counter + 1
}
```

Κώδικας 7.21: Έλεγχος τυχαιοποίησης για γραμμική συσχέτιση

Υπολογίζεται p-value οριακά ίσο με 0 και επομένως η μηδενική υπόθεση απορρίπτεται.

Ο τελευταίος έλεγχος είναι ένα έλεγχος Bootstrap με δειγματοληψία με επανάθεση. Η υλοποίηση του παρουσιάζεται στον Κώδικα 7.22:

```

1 # Data as dataframe
2 data <- data.frame(
3   x = c(87, 103, 81, 95, 102, 84, 99),
4   y = c(131, 162, 124, 149, 161, 133, 156)
5 )
6 # Number of samples
7 num_samples <- 999
8 # Vector in order to save the Pearson's coefficient correlations
9 pearson_correlations <- numeric(num_samples)
10 # Create samples and compute correlation
11 set.seed(456)
12 for (i in 1:num_samples) {
13   # Create random sample with replacement
14   sample_indices <- sample(1:nrow(data), size = nrow(data), replace =
15     TRUE)
16   sample_data <- data[sample_indices, ]
17   # Calculate Pearson's correlation coefficient
18   pearson_correlations[i] <- cor(sample_data$x, sample_data$y)
19 }

```

Κώδικας 7.22: Έλεγχος bootstrap για γραμμική συσχέτιση

Στη συνέχεια υπολογίζεται το p-value με τη χρήση των εντολών 7.23.

```

1 #Initialize a counter
2 bootstrap_counter=0
3 #Count how many coefficients are bigger than the observed
4 for (i in 1:length(pearson_correlations)) {
5   if(pearson_correlations[i] >= corPearson) bootstrap_counter =
6     bootstrap_counter + 1
7 }
8 #Calculate p-value
9 pBootstrap <- (bootstrap_counter+1)/(num_samples)

```

Κώδικας 7.23: Υπολογισμός p-value των δειγμάτων bootstrap για γραμμική συσχέτιση

Προκύπτει λοιπόν $p\text{-value}=0.655$ και επομένως η μηδενική υπόθεση δεν μπορεί να απορριφθεί.

Συνοπτικά τα αποτελέσματα για τον έλεγχο της διαφοράς των μέσων τιμών δύο πληθυσμών παρουσιάζονται στον πίνακα 7.8.

Μέθοδος	p-value	Απόφαση
Pearson	0.0000154	Απόρριψη της H_0
Spearman	0.002778	Απόρριψη της H_0
Τυχαιοποίησης	≈ 0	Απόρριψη της H_0
Bootstrap	0.655	Όχι απόρριψη της H_0

Πίνακας 7.8: Σύνοψη αποτελεσμάτων για τον έλεγχο της γραμμικής συσχέτισης

7.6 Συσχέτιση σε Πίνακες Συνάφειας

Έστω το εξής πρόβλημα: «Εξετάστηκαν 12 άτομα για το αν είναι ή όχι καπνιστές και το αν έχουν περάσει ή όχι καρκίνο του πνεύμονα. Ζητάται να εξεταστεί αν υπάρχει συσχέτιση ανάμεσα στον καρκίνο του πνεύμονα και το κάπνισμα.» Τα δεδομένα παρουσιάζονται στον Πίνακα 7.9, ο οποίος είναι ένας πίνακας συνάφειας.

	Απουσία καρκίνου	Υπαρξη καρκίνου	Σύνολο
Μη καπνιστής	3	2	5
Καπνιστής	2	5	7
Σύνολο	5	7	12

Πίνακας 7.9: Δεδομένα συσχέτισης καπνίσματος - καρκίνου του πνεύμονα

Οι δύο μεταβλητές, οι οποίες εξετάζονται αν σχετίζονται, είναι η X : Κάπνισμα και Y : Καρκίνος του πνεύμονα. Ορίζονται οι δύο υποθέσεις:

- H_0 : Δεν υπάρχει συσχέτιση καπνίσματος - καρκίνου του πνεύμονα
- H_1 : Υπάρχει συσχέτιση καπνίσματος - καρκίνου του πνεύμονα

Ορίζεται επίσης επίπεδο σημαντικότητας $\alpha = 0.05$. Τα δεδομένα του Πίνακα 7.9 εισάγονται σαν πίνακας στο πρόγραμμα όπως φαίνεται στον Κώδικα 7.24:

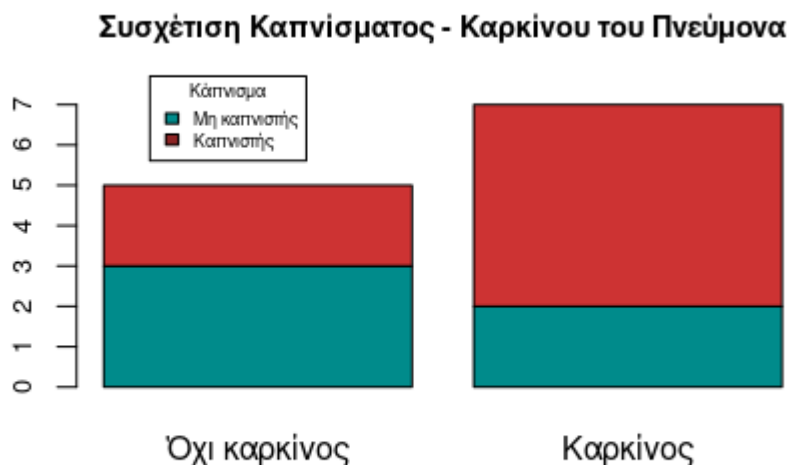
```

1 #Insert data
2 matrix_data <- matrix(
3 c(3,2,2,5),
4 nrow = 2,
5 ncol = 2,
6 byrow = TRUE
7 )
8 # Converting the matrix to a table
9 data <- as.table(matrix_data)

```

Κώδικας 7.24: Εισαγωγή δεδομένων για τη συσχέτιση καρκίνου-καπνίσματος με τη μορφή πίνακα συνάφειας

Προκειμένου να διαπιστωθεί γραφικά αν υπάρχει συσχέτιση μεταξύ των ποιοτικών μεταβλητών δημιουργείται το Διάγραμμα 7.9, στο οποίο παρουσιάζεται η συσχέτιση των ποιοτικών μεταβλητών «Κάπνισμα» και «Καρκίνος του Πνεύμονα».



Σχήμα 7.9: Διάγραμμα συσχέτισης Καπνίσματος - Καρκίνου του Πνεύμονα

Ο παραμετρικός έλεγχος χ^2 για πίνακες συνάφειας θα μπορούσε να χρησιμοποιηθεί για τον έλεγχο της συσχέτισης μεταξύ ποιοτικών μεταβλητών, όμως το δείγμα είναι απαγορευτικά μικρό. Παρόλαυτα μπορεί να υπολογιστεί η τιμή του ελέγχου και το αντίστοιχο p-value όπως φαίνεται στον Κώδικα 7.25.

```

1 #Perform Pearson 's Chi-squared test
2 pearson_chisq <- chisq.test(data, correct = FALSE)

```

Κώδικας 7.25: Υπολογισμός συντελεστή συσχέτισης Pearson χ^2

Με την παραπάνω εντολή προκύπτει $\chi^2 = 1.1853$ και p-value= 0.276.

Κάπνισμα(Κ: μη καπνιστής, Κ: καπνιστής)	Κ	Κ	Κ	Κ	Κ	Κ	Κ	Κ	Κ	Κ	Κ	Κ
Καρκίνος(0: απουσία, 1: ύπαρξη)	0	0	0	1	1	0	0	1	1	1	1	1

Πίνακας 7.10: Διαφορετική μορφή δεδομένων συσχέτισης καπνίσματος - καρκίνου του πνεύμονα

Για την πραγματοποίηση του ελέγχου τυχαιοποίησης Fisher για πίνακες συνάφειας είναι απαραίτητο να υπολογιστεί ο συντελεστής Pearson χ^2 , ο οποίος επιλέγεται ως συνάρτηση ελέγχου, προκειμένου να διαπιστωθεί πόσο ακραία είναι η παρατηρούμενη τιμή που υπολογίστηκε παραπάνω. Τα δεδομένα θα μπορούσαν να αναπαρασταθούν όπως φαίνονται στον Πίνακα 7.10. Από την μορφή των δεδομένων όπως παρουσιάζονται στον Πίνακα 7.10 φαίνεται ότι όλοι οι δυνατοί συνδυασμοί που μπορούν να προκύψουν από το δείγμα είναι ίσοι με όλους τους διαφορετικούς δυνατούς τρόπους που μπορούν να τοποθετηθούν τα πέντε 0 στις 12 θέσεις. Αυτοί είναι ίσοι με $\binom{12}{5} = 792$, και υπολογίζονται με τη χρήση του Κώδικα 7.26.

```

1 # Load necessary library
2 library(combinat)
3 # Function to generate all combinations of 12 seats with 5 0s and 7
  1s
4 generate_combinations <- function(n, k) {
5   combinations <- combn(n, k)
6   result <- apply(combinations, 2, function(x) {
7     sequence <- rep(1, n)
8     sequence[x] <- 0
9     paste(sequence, collapse = "")
10  })
11  return(result)
12  }
13 # Generate combinations for 12 seats with 5 0s
14 combinations <- generate_combinations(12, 5)

```

Κώδικας 7.26: Δημιουργία των δυνατών συνδυασμών του πίνακα συνάφειας

Από τους παραπάνω συνδυασμούς, επειδή τα δεδομένα είναι μόνο 0 και 1 και κάποιοι συνδυασμοί επαναλαμβάνονται, προκύπτουν οι παρακάτω πίνακες συνάφειας:

$$A = \begin{bmatrix} 0 & 5 \\ 5 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & 4 \\ 4 & 2 \end{bmatrix}, C = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}, D = \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix}, E = \begin{bmatrix} 4 & 1 \\ 1 & 6 \end{bmatrix}, F = \begin{bmatrix} 5 & 0 \\ 0 & 7 \end{bmatrix}.$$

Είναι φανερό ότι αρκεί να οριστεί το στοιχείο X_{11} κάθε πίνακα και τα υπόλοιπα προκύπτουν από τα αθροίσματα γραμμών και στηλών.

Για να υπολογιστεί η συχνότητα εμφάνισης του κάθε πίνακα απαιτούνται γνώσεις συνδυαστικής. Κάθε φορά χρειάζεται να βρεθούν οι δυνατοί συνδυασμοί των μηδενικών που ορίζονται από το X_{11} στις πρώτες 5 θέσεις και οι δυνατοί συνδυασμοί των υπόλοιπων μηδενικών στις επόμενες 7 θέσεις. Για την εύρεση των συνολικών φορών εμφάνισης κάθε πίνακα πολλαπλασιάζονται τα αποτελέσματα των δύο παραπάνω συνδυασμών, βάσει του πολλαπλασιαστικού νόμου.

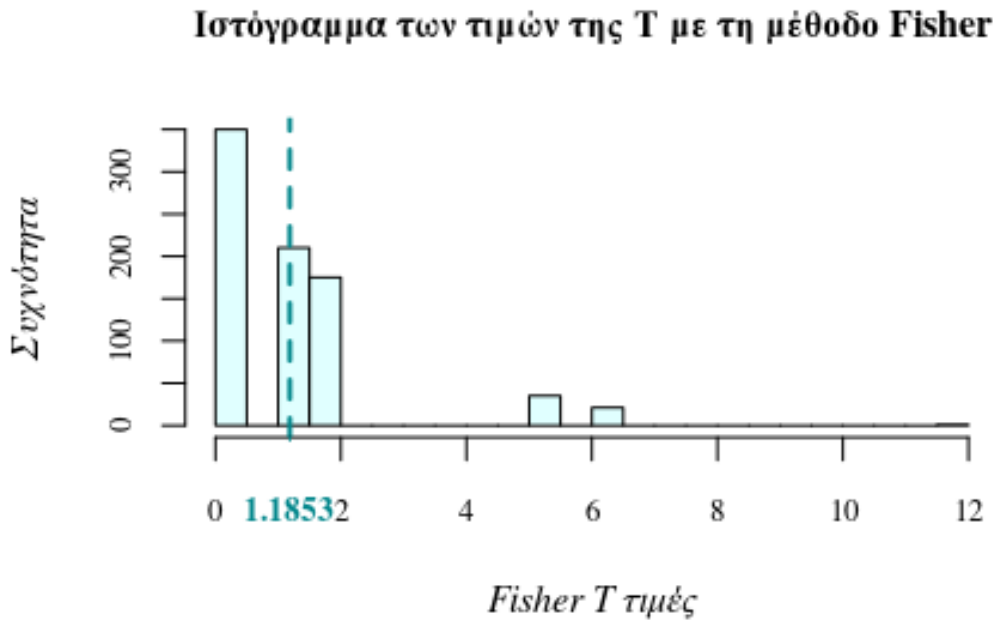
Στη συνέχεια μπορεί να υπολογιστεί η τιμή της συνάρτησης ελέγχου για κάθε πίνακα. Όλα τα παραπάνω συνοψίζονται στον Πίνακα 7.11. Για τον υπολογισμό του

Πίνακας	Συχνότητα	Τιμή Συνάρτησης Ελέγχου
A	21	6.1224
B	175	1.6555
C	350	0.0098
D	210	1.1853
E	35	5.182
F	1	12

Πίνακας 7.11: Συχνότητα εμφάνισης κάθε δυνατού πίνακα συνάφειας και τιμή Pearson

p-value αρκεί να βρεθούν πόσοι πίνακες έχουν ίση ή πιο ακραία τιμή συνάρτησης ελέγχου από την παρατηρούμενη. Υπολογίζεται ότι p-value = $\frac{442}{792} = 0.558$. Επομένως σε επίπεδο σημαντικότητας $\alpha=0.05$ η μηδενική υπόθεση δεν απορρίπτεται.

Οι τιμές της συνάρτησης ελέγχου που προέκυψαν από τους δυνατούς πίνακες συνάφειας με τη μέθοδο Fisher, αλλά και η παρατηρούμενη τιμή της συνάρτησης ελέγχου παρουσιάζονται στο Διάγραμμα 7.10.



Σχήμα 7.10: Ιστόγραμμα των τιμών της συνάρτησης ελέγχου για τους διαφορετικούς δυνατούς πίνακες συνάφειας με τη μέθοδο Fisher

7.7 Έλεγχος Καλής Προσαρμογής

Έστω το εξής πρόβλημα: «Ζητάται να ελεγχθεί αν ο αριθμός των επειγόντων περιστατικών ανά ημέρα ενός νοσοκομείου ακολουθεί κατανομή Poisson.» Τα δεδομένα αφορούν τον αριθμό των επειγόντων περιστατικών σε διάστημα 100 ημερών, ομαδοποιήθηκαν και παρουσιάζονται στον Πίνακα 7.12.

Η μεταβλητή, η οποία μελετάται, είναι η X : αριθμός επειγόντων περιστατικών νοσοκομείου.

Για την πραγματοποίηση του ελέγχου ορίζεται η μηδενική υπόθεση λόγω της διατύπωσης του προβλήματος ως « H_0 : Ο αριθμός των περιστατικών ακολουθεί κατανομή Poisson». Η εναλλακτική θα ήταν « H_1 : Ο αριθμός των περιστατικών δεν ακολουθεί κατανομή Poisson». Για να μπορεί η συνάρτηση ελέγχου να διαχωρίσει ανάμεσα στις δύο υποθέσεις θα οριστεί εναλλακτική υπόθεση με συγκεκριμένη κατανομή. Επομένως οι υποθέσεις του ελέγχου είναι:

- H_0 : Η μεταβλητή ακολουθεί κατανομή Poisson
- H_1 : Η μεταβλητή ακολουθεί Αρνητική Διωνυμική κατανομή

Αριθμός επειγόντων περιστατικών	Αριθμός ημερών
0	53
1	27
2	13
3	4
4	2
5	1

Πίνακας 7.12: Αριθμός επειγόντων περιστατικών ενός νοσοκομείου σε διάστημα 100 ημερών

Ορίζεται επίπεδο σημαντικότητα $\alpha = 0.05$.

Αρχικά εισάγονται τα δεδομένα και υπολογίζονται οι απαραίτητες ποσότητες από το δείγμα, όπως φαίνεται στον Κώδικα 7.27.

```

1  #Insert data
2  data <- c(rep(0,53),
3  rep(1,27),
4  rep(2,13),
5  rep(3,4),
6  rep(4,2),
7  rep(5,1))
8  #Calculate mean, number of observations and variance of the dataset
9  x_mean <- mean(data)
10 n <- length(data)
11 var_mean <- var(data)
12 #Calculate the observed value of T
13 tobs <- var_mean/x_mean

```

Κώδικας 7.27: Εισαγωγή δεδομένων για έλεγχο κατανομής και υπολογισμός χαρακτηριστικών μεγεθών του δείγματος

Στο Διάγραμμα 7.11 παρουσιάζονται οι συχνότητες του Πίνακα 7.12 και οι αναμενόμενες συχνότητες των τιμών της μεταβλητής «Αριθμός επειγόντων περιστατικών» αν αυτή ακολουθούσε την κατανομή Poisson.



Σχήμα 7.11: Παρατηρούμενες και Αναμενόμενες Τιμές

Ο έλεγχος που θα πραγματοποιηθεί είναι ένας έλεγχος Monte Carlo καλής προσαρμογής. Για την πραγματοποίηση του απαιτείται να επιλεχθεί μια συνάρτηση ελέγχου που θα διαχωρίζει ανάμεσα στις δύο υποθέσεις και θα παίρνει μοναδικές τιμές όταν ισχύει η μηδενική υπόθεση. Στην συγκεκριμένη περίπτωση θα μπορούσε να είναι η

$$T = \frac{s^2}{\bar{x}},$$

όπου s^2 η διασπορά του δείγματος και \bar{x} ο δειγματικός μέσος. Η παραπάνω συνάρτηση παίρνει τιμές κοντά στο 1, όταν η κατανομή είναι η Poisson και μεγαλύτερες από 1 όταν η κατανομή είναι η Αρνητική Διωνυμική. Η συγκεκριμένη συνάρτηση T ονομάζεται δείκτης διασποράς.

Ορίζεται επίσης η απαραίτητη συνάρτηση για την προσομοίωση των δειγμάτων από την κατανομή Poisson και για τον υπολογισμό της συνάρτησης ελέγχου για κάθε δείγμα. Ο ορισμός της συνάρτησης υλοποιείται με τον Κώδικα 7.28.

Έπειτα ορίζεται ο αριθμός των δειγμάτων που θα προσομοιωθούν από την κατανομή Poisson, έστω 999. Υπενθυμίζεται ότι η εκτιμήτρια μέγιστης πιθανοφάνειας για την Poisson είναι ο δειγματικός μέσος. Στη συνέχεια καλείται η παραπάνω συνάρτηση, με παράμετρο τη μέση τιμή του δείγματος και αποθηκεύεται η τιμή της συνάρτησης ελέγχου για κάθε δείγμα. Έπειτα υπολογίζεται το p-value. Τα παραπάνω υπολογίζονται στον κώδικα 7.29.

```
1 #Function that simulates k samples from Poisson with parameter l and
  length n
2 poisson.clt <- function(k,n,l)
3 {
4   Sn <- rep(NA,k)
5   for(i in 1:k)
6   {
7     x <- rpois(n,l)
8     Sn[i] <- var(x)/mean(x)
9   }
10  return(Sn)
11 }
```

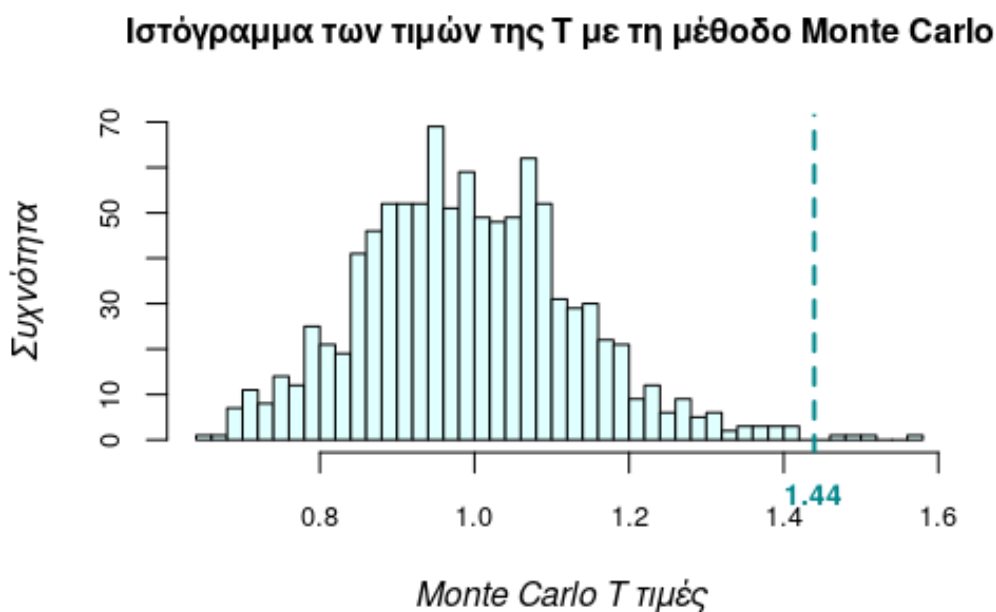
Κώδικας 7.28: Δημιουργία συνάρτησης για προσομοίωση από την Poisson

```
1 #Number of samples
2 times <- 999
3 total <- length(data)
4 #Simulation of 999 Poisson samples with parameter mean of the
  observed dataset
5 T <- poisson.clt(times, total, x_mean)
6 #Calculation of p-value
7 p <- (sum(T>tobs)+1)/(times+1)
```

Κώδικας 7.29: Δημιουργία προσομοιωμένων δειγμάτων βάσει του δοσμένου, υπολογισμός της T και του p-value

Προκύπτει p-value ίσο με 0.004 και επομένως η μηδενική υπόθεση απορρίπτεται.

Οι παραπάνω τιμές από την υλοποίηση της μεθόδου Monte Carlo καλής προσαρμογής, αλλά και η παρατηρούμενη τιμή της συνάρτησης ελέγχου φαίνονται στο Διάγραμμα 7.12.



Σχήμα 7.12: Ιστόγραμμα των τιμών της συνάρτησης ελέγχου της μεθόδου Monte Carlo καλής προσαρμογής

7.8 Συμπεράσματα

Συνοψίζοντας είναι φανερό ότι με τη χρήση των υπολογιστών, και συγκεκριμένα με την R, μπορούν να οπτικοποιηθούν εύκολα κάθε είδους δεδομένα, να βρεθεί η σχέση που μπορεί να έχουν διαφορετικά σύνολα δεδομένων, καθώς και να γίνουν στατιστικοί έλεγχοι, που λόγω πολλών και δύσκολων υπολογισμών, με χαρτί και μολύβι είναι αδύνατο.

Είναι σημαντικό να επισημανθεί ότι οι αποφάσεις που προκύπτουν από τους παραμετρικούς και μη παραμετρικούς ελέγχους είναι λιγότερο αξιόπιστες, αφού βασίζονται σε υποθέσεις των πληθυσμών, που για τα δεδομένα που αναφέρθηκαν σε αυτό το κεφάλαιο δεν είναι γνωστές. Για αυτό το λόγο, μπορεί να οδηγήσουν και σε διαφορετικές αποφάσεις από τους ελέγχους τυχαιοποίησης, όπως στο Υποκεφάλαιο 7.2.

Κεφάλαιο 8

Συμπεράσματα

Από την παρούσα διπλωματική εργασία προκύπτουν τα εξής συμπεράσματα:

- Για την εφαρμογή των παραμετρικών ελέγχων απαιτούνται υποθέσεις για την κατανομή του πληθυσμού ή των πληθυσμών της μεταβλητής ή των μεταβλητών αντίστοιχα. Οι υποθέσεις είναι απαραίτητες για να χρησιμοποιηθεί ασυμπτωτικό αποτέλεσμα στην συνάρτηση ελέγχου.
- Οι μη παραμετρικοί έλεγχοι χρησιμοποιούν άλλες ιδιότητες των δεδομένων, και όχι την κατανομή που ακολουθούν. Παρόλα αυτά και σε αυτό το είδος των ελέγχων χρησιμοποιούνται ασυμπτωτικά αποτελέσματα. Είναι σημαντικό όμως να αναφερθεί ότι οι μη παραμετρικού έλεγχοι έχουν εφαρμογή και σε πολύ μικρά δείγματα. Όμως όταν οι υποθέσεις περί κανονικότητας πληρούνται, προτιμώνται οι παραμετρικοί έλεγχοι.
- Στους ελέγχους τυχαιοποίησης, μοναδική προϋπόθεση είναι η H_0 να ορίζει την απουσία δομής στα δεδομένα. Δεν χρειάζεται η H_0 να είναι συγκεκριμένη αφού η βάση των ελέγχων τυχαιοποίησης είναι το γεγονός ότι όλοι οι συνδυασμοί είναι ισοπίθανοι κάτω από τη μηδενική υπόθεση. Είναι σημαντικό να αναφερθεί ότι η ισχύς των ελέγχων τυχαιοποίησης είναι μεγαλύτερη από τους παραμετρικούς ή μη παραμετρικούς ελέγχους, και επομένως η απόφαση στην οποία οδηγούν είναι πιο αξιόπιστη.
- Οι έλεγχοι Monte Carlo προτιμώνται σε ελέγχους καλής προσαρμογής, αφού η μηδενική υπόθεση κατευθύνει προς την κατανομή από την οποία προσομοιώνονται τα δείγματα.
- Για την πραγματοποίηση των ελέγχων με τη μέθοδο Bootstrap δεν απαιτείται

γνώση της κατανομής από την οποία προέρχεται η μεταβλητή που εξετάζεται, όπως και στους ελέγχους τυχαιοποίησης. Όλες οι απαραίτητες πληροφορίες προκύπτουν από το δοσμένο δείγμα. Για αυτό το λόγο οι συγκεκριμένοι έλεγχοι αποτυγχάνουν σε πολύ μικρά δείγματα, αφού δεν υπάρχει επαρκής πληροφορία.

- Μοναδικό κριτήριο για την επιλογή της συνάρτησης ελέγχου στους ελέγχους τυχαιοποίησης, στους Monte Carlo και τους Bootstrap είναι να μπορεί να διακρίνει ανάμεσα στις δύο υποθέσεις. Δεν είναι απαραίτητα η γνώση της κατανομής της, αφού οι τιμές της προκύπτουν από τα προσομοιωμένα δείγματα. Επισημαίνεται ότι στους ελέγχους Monte Carlo η επιλογή της συνάρτησης ελέγχου είναι πιο δύσκολη, απαιτείται πολύ καλή γνώση της στατιστικής και των ιδιοτήτων των μέτρων που επιλέγονται, αφού οι δύο υποθέσεις υποδηλώνουν συγκεκριμένες κατανομές, και η συνάρτηση ελέγχου πρέπει να ξεχωρίζει ανάμεσα στις δύο.

Γενικά οι Έλεγχοι Τυχαιοποίησης, Monte Carlo και Bootstrap έχουν χρήση όχι μόνο σε ελέγχους υποθέσεων, αλλά γενικά σε οποιοδήποτε πρόβλημα χρειάζεται να εκτιμηθεί κάποια μεταβλητή με άγνωστη κατανομή και να υπολογιστούν έπειτα στατιστικά μεγέθη. Στην παρούσα διπλωματική αναφέρονται γενικά τα βήματα των μεθόδων και οι εφαρμογές τους σε μερικά μόνο είδη των ελέγχων υποθέσεων.

Παρόλα αυτά στη βιβλιογραφία μπορεί κανείς να βρει περισσότερες εφαρμογές, π.χ. εκτίμηση διαστημάτων εμπιστοσύνης με τη μέθοδο Bootstrap, ανάλυση χρονοσειρών με τη μέθοδο Bootstrap αλλά και χρήση της μεθόδου Bootstrap για εύρεση των κύριων συνιστωσών, ιδιοτιμών και ιδιοδιανυσμάτων.

Τέλος είναι σημαντικό να αναφερθεί η τεράστια σημασία που έχει η χρήση των υπολογιστών στην υλοποίηση αυτών των μεθόδων. Άλλωστε και η ιδέα για την ανάπτυξη τους βασίζεται στην ανάπτυξη των υπολογιστών. Συγκεκριμένα η R είναι μια ιδανική γλώσσα για στατιστική συμπερασματολογία και για οπτικοποίηση των δεδομένων.

Βιβλιογραφία

- Φουσκάκης, Δ. & Κοκολάκης, Γ. (2009). *Στατιστική Θεωρία & Εφαρμογές*. Αθήνα: Εκδόσεις Συμεών.
- Φουσκάκης, Δ. (2021). *Ανάλυση Δεδομένων με Χρήση της R*. Αθήνα: Εκδόσεις Τσότρας.
- Καρλής, Δ. (2008). *Υπολογιστική Στατιστική*. Αθήνα: Εκδόσεις Οικονομικού Πανεπιστημίου Αθηνών.
- Gibbons, J. D. & Chakraborti, S. (1992). *Nonparametric Statistical Inference*. New York: Marcel Dekker, Inc.
- Good, F. (2005). *Permutation, Parametric, and Bootstrap Tests of Hypotheses*. Springer Science+Business Media, Inc.
- Ostertagova, E. & Ostertag, O. (2013). *Methodology and Application of One-way ANOVA*. Americal Journal of Mechanical Engineering. Science and Education Publishing.
- Douglas, R. W. (2004). *A Student's Guide to Statistics for Analysis of Cross Tabulations*. World Cultures.
- Howard Beasley, W. & Rodgers, J. L. (2012). *Bootstrapping and Monte Carlo Methods*.
- Lobo M. & Guntur, R. D. (2018). *Spearman's rank correlation analysis on public perception toward health partnership projects between Indonesia and Australia in East Nusa Tenggara Province*. Journal of Physics.
- Obilor, E. I. & Amadi, E. C. (2018). *Test for Significance of Pearson's Correlation Coefficient (r)*. International Journal of Innovative Mathematics, Statistics & Energy Policies.
- Zhang, Y. & Zhao, Q. (2023). *What is a Randomization Test?*. Journal of the American Statistical Association.

Συντομογραφίες - Αρκτικόλεξα - - Ακρωνύμια

τ.μ.	Τυχαία μεταβλητή
σ.π.π.	Συνάρτηση Πυκνότητας Πιθανότητας
σ.μ.π.	Συνάρτηση Μάζας Πιθανότητας
ε.σ.	Επίπεδο Σημαντικότητας
τ.δ.	Τυχαίο Δείγμα
μ.τ.	Μέση Τιμή
Κ.Ο.Θ.	Κεντρικό Όριακό Θεώρημα
β.ε.	Βαθμοί Ελευθερίας
Ε.Μ.Π.	Εκτιμήτρια Μέγιστης Πιθανοφάνειας

