



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ
ΤΜΗΜΑ ΙΑΤΡΙΚΗΣ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
«Μεθοδολογία Βιοϊατρικής Έρευνας,
Βιοστατιστική και Κλινική Βιοπληροφορική»**

**ΣΥΓΚΡΙΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΛΟΓΙΣΜΙΚΩΝ
(JMP, MINITAB, R, SPSS)
ΣΤΗΝ ΕΦΑΡΜΟΓΗ ΜΕΛΕΤΩΝ ΔΙΑΣΤΑΥΡΟΥΜΕΝΗΣ
ΜΕΤΑΒΑΣΗΣ (CROSSOVER STUDIES)
ΤΟΥ
ΔΗΜΗΤΡΙΟΥ Κ. ΤΣΑΦΗ**

Επιβλέπων: Αναπληρωτής Καθηγητής κ. Χρήστος Νάκας

ΕΡΓΑΣΙΑ
Που υποβλήθηκε στο Τμήμα Ιατρικής
της Σχολής Επιστημών Υγείας
του Πανεπιστημίου Θεσσαλίας
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη
Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική
Βιοπληροφορική

**Λάρισα
Ιανουάριος 2016**

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στη:

**Μεθοδολογία Βιοϊατρικής Έρευνας,
Βιοστατιστική και Κλινική Βιοπληροφορική**

που απονέμει το Τμήμα Ιατρικής της Σχολής Επιστημών Υγείας του Πανεπιστημίου Θεσσαλίας

Εγκρίθηκε τον Ιανουάριο 2016 από την εξεταστική επιτροπή:

Υπογραφές

Επιβλέπων καθηγητής:

Νάκας Χρήστος, Αναπληρωτής Καθηγητής Βιομετρίας

Μέλη εξεταστικής επιτροπής:

Νάκας Χρήστος, Αναπληρωτής Καθηγητής Βιομετρίας

Ζιντζαράς Ηλίας, Καθηγητής Βιομαθηματικών – Βιομετρίας

Ραχιώτης Γεώργιος, Επίκουρος Καθηγητής Επιδημιολογίας

Στη μνήμη του πατέρα
μου Κωνσταντίνου

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου κ. Χρήστο Νάκα ο οποίος - για δεύτερη φορά - μου έκανε την τιμή να επιβλέψει διπλωματική μου διατριβή. Αισθάνομαι για το φίλο Χρήστο απέραντη ευγνωμοσύνη και σεβασμό.

Θα ήθελα να εκφράσω τις βαθύτατες ευχαριστίες μου στη σύζυγό μου Φανή και στην κόρη μου Σουλτάνα για την υπομονή τους και στη μητέρα μου Σουλτάνα για την συμπαράστασή της.

Ευχαριστώ ακόμη την αδερφή μου Άννα, το γαμπρό μου Γιάννη και την ανιψιά μου Θεανώ οι οποίοι με φιλοξένησαν πολλά σαββατοκύριακα στη Λάρισα την περίοδο 2014 και 2015. Χωρίς τη δική τους βοήθεια δεν θα ήταν δυνατή η παρακολούθηση του συγκεκριμένου ΜΠΣ.

Τέλος, θέλω να ευχαριστήσω το συνάδελφο, φίλο και αδερφό Μηνά Χαραμαλή για την στήριξή του και για την δακτυλογράφηση μεγάλου μέρους του 1^{ου} και 2^{ου} κεφαλαίου.

ΠΕΡΙΛΗΨΗ

ΔΗΜΗΤΡΙΟΣ Κ. ΤΣΑΦΗΣ

ΣΥΓΚΡΙΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΛΟΓΙΣΜΙΚΩΝ (JMP, MINITAB, R, SPSS) ΣΤΗΝ ΕΦΑΡΜΟΓΗ ΜΕΛΕΤΩΝ ΔΙΑΣΤΑΥΡΟΥΜΕΝΗΣ ΜΕΤΑΒΑΣΗΣ

Ιανουάριος 2016

Οι μελέτες διασταυρούμενης μετάβασης αποτελούν τυχαίοποιημένες μελέτες με εφαρμογή κυρίως στη φαρμακοβιομηχανία (κλινικές μελέτες 1^{ης} και 2^{ης} φάσης) και στη βιομηχανία τροφίμων. Παρουσιάζουν ιδιορρυθμίες με σημαντικά πλεονεκτήματα αλλά και αρκετά μειονεκτήματα. Η σημαντικότερη παρενέργεια την οποία εμφανίζουν είναι η ύπαρξη του φαινομένου επίδρασης ακολουθίας. Η ύπαρξη στατιστικά σημαντικής επίδρασης ακολουθίας προκαλεί προβλήματα και δυσχεραίνει την στατιστική επεξεργασία των δεδομένων μιας μελέτης διασταυρούμενης μετάβασης.

Η παρούσα εργασία χωρίζεται σε δύο μέρη. Τα κεφάλαια 1, 2 και 3 παρουσιάζουν τη θεωρία που διέπει τις μελέτες διασταυρούμενης μετάβασης, τα πλεονεκτήματα και τα μειονεκτήματά τους καθώς και τις βασικές τους έννοιες. Ακόμη, γίνεται παρουσίαση του υποδείγματος των λατινικών τετραγώνων πάνω στο οποίο στηρίζονται οι περισσότεροι πολύπλοκες μελέτες διασταυρούμενης μετάβασης καθώς και του μεικτού γραμμικού μοντέλου. Εν συνεχεία, γίνεται διεξοδική παρουσίαση της πιο απλής μελέτης διασταυρούμενης μετάβασης (2X2 πείραμα) και περιγράφεται διεξοδικά το μαθηματικό της υπόδειγμα.

Το δεύτερο μέρος που αποτελείται από το κεφάλαια 4 και 5 αφορά στο πρακτικό μέρος της εργασίας. Με τη χρήση τεσσάρων στατιστικών λογισμικών (JMP, Minitab, R και SPSS) προσεγγίζεται μια 3X3 μελέτη διασταυρούμενης μετάβασης που τα δεδομένα της προέρχονται από test γευσίγνωσσίας σοκολάτας. Πραγματοποιείται σύγκριση των αποτελεσμάτων που δίνουν τα τέσσερα λογισμικά και αξιολόγησή τους. Τέλος, ακολουθεί μια ανάλυση ευαισθησίας προκειμένου να επιβεβαιωθούν (ή μη) τα αρχικά αποτελέσματα και συμπεράσματα.

Λέξεις κλειδιά: Μελέτες διασταυρούμενης μετάβασης, επίδραση ακολουθίας, λατινικά τετράγωνα, στατιστικό λογισμικό, ανάλυση δεδομένων, μεικτά μοντέλα, ανάλυση ευαισθησίας

ABSTRACT

DIMITRIOS K. TSAFIS

A COMPARISON OF STATISTICAL SOFTWARE (JMP, MINITAB, R, SPSS) FOR THE IMPLEMENTATION OF CROSSOVER STUDIES

JANUARY 2016

The crossover studies are randomized trials with application mainly in pharmaceuticals (clinical trials of 1st and 2nd phase) and the food industry. They have quirks with significant advantages and several disadvantages. The major side effect exhibited is the presence of the phenomenon carryover effect. The existence of the statistically significant carryover effect causes problems and complicates the statistical analysis of the data of a crossover study.

This paper is divided into two parts. Chapters 1, 2 and 3 illustrate the theory that governs the crossover studies, the advantages and disadvantages and their basic concepts. Moreover, the format of Latin squares, on which the more complex crossover studies and the mixed linear models are based, is presented. Then, a thorough presentation of the most simple crossover study transition (2X2 experiment) is made and its mathematical model is thoroughly described.

The second part that consists of Chapters 4 and 5 relates to the practical part of the paper. By using four statistical software (JMP, Minitab, R and SPSS) a 3X3 crossover study, whose data come from a chocolate tasting test, is approached. Following a comparison of the results given by the four software and evaluation. Finally, a sensitivity analysis is performed to confirm (or not) the initial results and conclusions.

Keywords: crossover studies, carryover effect, Latin squares, statistical software, data analysis, mixed models, sensitivity analysis

ΚΑΤΑΛΟΓΟΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

A/A	Τίτλος	Σελ.
	Αφιέρωση	I
	Ευχαριστίες	II
	Περίληψη	IV
	Abstract	VI
	Κατάλογος περιεχομένων	VII
	Κατάλογος πινάκων	X
	Κατάλογος γραφημάτων	XII
	Κατάλογος σχημάτων	XIII
	ΚΕΦΑΛΑΙΟ 1. Εισαγωγή στις Crossover δοκιμές	1
1-1	Εισαγωγή	1
1-2	Crossover δοκιμές	2
1-2-1	Ορισμός	2
1-2-2	Η αξία της τυχαιοποίησης (Randomization)	2
1-3	Πλεονεκτήματα – Μειονεκτήματα crossover δοκιμής	3
1-4	Περίοδος έκπλυσης (Washout period)	4
1-5	Είδη επιδράσεων σε crossover δοκιμές	5
1-6	Στατιστική ανάλυση	6
1-7	Χρήση crossover δοκιμών	7
	ΚΕΦΑΛΑΙΟ 2. Λατινικά τετράγωνα – Μεικτά μοντέλα	9
2-1	Πειραματικοί σχεδιασμοί	9
2-1-1	Εισαγωγή	9
2-1-2	Βασικές αρχές πειραματικού σχεδιασμού	11
2-1-3	Στάδια πειραματικού σχεδιασμού	13
2-2	Λατινικά τετράγωνα (Latin squares)	14
2-2-1	Ιστορική αναδρομή	14
2-2-2	Ορισμός και βασικές έννοιες	14
2-2-3	Είδη λατινικών τετραγώνων	16
2-2-4	Τυχαιοποίηση λατινικών τετραγώνων	17
2-2-5	Μαθηματικό υπόδειγμα λατινικών τετραγώνων	17
2-2-6	Παράδειγμα 5X5 λατινικού τετραγώνου	21
2-2-7	Μικρά λατινικά τετράγωνα	23
2-3	Στατιστικά λογισμικά	23
2-3-1	Γενικά	23

2-3-2	JMP	24
2-3-3	MINITAB	25
2-3-4	R	25
2-3-5	SPSS	26
ΚΕΦΑΛΑΙΟ 3. 2Χ2 ή ΑΒ/ΒΑ και 3Χ3 Crossover δοκιμές		29
3-1	Εισαγωγή	29
3-2	Κέρδη και προβλήματα	31
3-3	Είδη επιδράσεων	32
3-3-1	Επίδραση θεραπείας (treatment effects)	32
3-3-2	Επίδραση περιόδου (period effects)	33
3-3-3	Επίδραση ακολουθίας (carry-over effects)	34
3-4	Γραφήματα – Στατιστική ανάλυση	34
3-4-1	Χρησιμότητα γραφημάτων	34
3-4-2	Γραφήματα και 2Χ2 crossover δοκιμές	35
3-4-3	Στατιστική ανάλυση	38
3-5	Μαθηματικό υπόδειγμα	39
3-5-1	Προσέγγιση με t-tests	41
3-5-2	Προσέγγιση μέσω ανάλυσης διακύμανσης	44
3-5-3	Προσέγγιση με Bayesian analysis	46
3-6	Παράδειγμα 2Χ2 crossover δοκιμής	46
3-7	3Χ3 crossover δοκιμή	52
3-7-1	Γενικά	52
3-7-2	Μεικτό μοντέλο	55
ΚΕΦΑΛΑΙΟ 4. Εφαρμογή: Test γευσιγνωσίας σοκολάτας		57
4-1	Το πείραμα	57
4-2	Στόχος	59
4-3	Αρχείο δεδομένων (Data file)	59
4-4	Επίδραση ακολουθίας (carry-over effects)	60
4-5	Έλεγχος για ελλείπουσες τιμές και ακραίες τιμές	62
4-6	Περιγραφική στατιστική	63
4-6-1	Εξαρτημένη μεταβλητή (Περιγραφικά μέτρα-Κατανομές-Κανονικότητα)	63
4-6-2	Ανεξάρτητες μεταβλητές (κατανομές)	67
4-7	Testing τα στατιστικά λογισμικά	69
4-8	Χρήσιμα γραφήματα για crossover δοκιμές	69

4-9	Έλεγχος επίδρασης carry-over	73
4-10	Πλήρες μοντέλο	79
4-11	Συμπέρασμα	87
ΚΕΦΑΛΑΙΟ 5. Ανάλυση ευαισθησίας (Sensitivity analysis)		90
5-1	Εισαγωγή	90
5-2	Ανάλυση ευαισθησίας και crossover δοκιμές	91
5-3	Ορισμός και χρησιμότητα ανάλυσης ευαισθησίας	91
5-4	Βασικές προσεγγίσεις ανάλυσης ευαισθησίας	92
5-5	Ανάλυση ευαισθησίας στο test γευσιγνωσίας σοκολάτας	93
5-6	Συμπέρασμα	100
Βιβλιογραφία		101
Παράρτημα		103

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Α/Α	Τίτλος πίνακα	Σελ.
Κεφάλαιο 2		
2-1	Μετρήσεις παραδείγματος αντοχής μηχανών	12
2-2	Εκτιμητές υποδείγματος λατινικών τετραγώνων	19
2-3	Ανάλυση διακύμανσης για λατινικά τετράγωνα με σταθερές επιδράσεις	20
2-4	Δεδομένα παραδείγματος πυραύλων	21
2-5	Ανάλυση διακύμανσης για παράδειγμα	22
Κεφάλαιο 3		
3-1	2X2 ή AB/BA crossover δοκιμή	30
3-2	Σταθερές και τυχαίες επιδράσεις για το πλήρες μοντέλο	41
3-3	Σταθερές επιδράσεις για το πλήρες μοντέλο	41
3-4	For-Sal / Sal – For crossover δοκιμή	46
3-5	Type III tests of fixed effects	49
3-6	Estimates of fixed effects	50
3-7	Type III tests of fixed effects	50
3-8	Estimates of fixed effects	51
3-9	Crossover σχεδιασμός 3 περιόδων και 3 παρεμβάσεων	52
3-10	Κωδικοποίηση επίδρασης carry-over σε 3X3 crossover δοκιμή	54
3-11	Κωδικοποίηση επίδρασης carry-over με τη χρήση ψευδομεταβλητών	54
Κεφάλαιο 4		
4-1	Επίδραση carry-over σε 3X3 crossover πείραμα	61
4-2	Κωδικοποίηση carry-over σε 3X3 crossover πείραμα	61
4-3	Έλεγχος για ελλείπουσες τιμές	62
4-4	Tests of Normality	64
4-5	Tests of Normality	65
4-6	Tests of Normality	66
4-7	Tests of Normality	67
4-8	F-ratios και p-values από Anova Table	69
4-9	Μέσες βαθμολογίες προτίμησης ως προς την περίοδο και την ομάδα δοκιμής	71
4-10	Μέσες βαθμολογίες προτίμησης ως προς την περίοδο και την σοκολάτα	72
4-11	Προσέγγιση μεικτού μοντέλου με 4 στατιστικά λογισμικά	74
4-12	Έλεγχος καλής προσαρμογής και επίδρασης carry-over	75
4-13	Εκτιμήσεις σταθερών (fixed) επιδράσεων	76
4-14	Εκτιμήσεις διακύμανσης από μεταβλητές μοντέλου	77
4-15	Μέσες τιμές επίδρασης carry-over (κατηγορίες)	78

4-16	Προσέγγιση πλήρους μεικτού μοντέλου με 4 στατιστικά λογισμικά	80
4-17	Έλεγχος καλής προσαρμογής μεικτού μοντέλου	81
4-18	Σταθερές επιδράσεις επεξηγηματικών μεταβλητών	82
4-19	Εκτιμήσεις διακύμανσης από μεταβλητές μοντέλου	83
4-20	Σύγκριση JMP, Minitab, R και SPSS	87
Κεφάλαιο 5		
5-1	Αρχικές και νέες τιμές μεταβλητής «Βαθμολογία προτίμησης»	95
5-2	Shapiro-Wilk Normality Test για «Βαθμολογία προτίμησης»	96
5-3	Βασικά περιγραφικά μέτρα αρχικών και νέων τιμών «Βαθμολογίας προτίμησης»	97
5-4	Προσαρμογή μεικτού μοντέλου σε αρχικές και νέες τιμές «Βαθμολογίας προτίμησης»	98

ΚΑΤΑΛΟΓΟΣ ΓΡΑΦΗΜΑΤΩΝ

Α/Α	Τίτλος γραφήματος	Σελ.
Κεφάλαιο 3		
3-1	2X2 crossover δοκιμή με επίδραση θεραπείας	33
3-2	2X2 crossover δοκιμή με επίδραση περιόδου	33
3-3	2X2 crossover δοκιμή με επίδραση carry-over	34
3-4	Period 2 vs Period 1 plot	35
3-5	Period 2 vs Period 1 plot with centroids	36
3-6	Profiles plots for outcome data	37
3-7	Τιμές PEF για ομάδες For/Sal και Sal/For	47
3-8	Profiles plots για ομάδες For/Sal και Sal/For	48
3-9	Means plot ως προς τις ομάδες και την περίοδο	48
3-10	Scatterplot Std residuals vs Predicted values	51
Κεφάλαιο 4		
4-1	Θηκόγραμμα προτίμησης	62
4-2	Probability plot μεταβλητής προτίμηση	63
4-3	Histogram προτίμηση ως προς αλληλουχία	64
4-4	Histogram προτίμηση ως προς το είδος σοκολάτας	65
4-5	Histogram προτίμηση ως προς την περίοδο δοκιμής	66
4-6	Histogram προτίμηση ως προς carry-over	67
4-7	Κυκλικά διαγράμματα 4 ανεξάρτητων μεταβλητών	68
4-8	Profiles plots γευσιγνωστών	70
4-9	Means plot προτίμησης ως προς την περίοδο και την ομάδα δοκιμής	72
4-10	Means plot προτίμησης ως προς την περίοδο και το είδος σοκολάτας	73
4-11	Residual plots βαθμολογίας προτίμησης	79
4-12	Least Squares Means Plot	86
4-13	Χρήση λογισμικών σε δημοσιεύσεις το 2014	89
Κεφάλαιο 5		
5-1	Ιστογράμματα αρχικών και νέων μεταβλητών	95

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

A/A	Τίτλος σχήματος	Σελ.
Κεφάλαιο 2		
2-1	6X6 Latin Square	15
2-2	4X4 Latin Square	15
2-3	5X5 Latin Square	15
2-4	Κανονικοποιημένο 4X4 Latin Square	16
Κεφάλαιο 3		
3-1	2X2 ή AB/BA crossover δοκιμή	30
3-2	Προσέγγιση 2 σταδίων για ανάλυση δεδομένων 2X2 crossover δοκιμής	39
Κεφάλαιο 4		
4-1	Εκτιμήσεις παραμέτρων (JMP και SPSS)	84
4-2	Εκτιμήσεις παραμέτρων (R και Minitab)	85
4-3	Least Square Means μεταβλητών	86

Κεφάλαιο 1

Εισαγωγή στις Crossover δοκιμές

(...παν δώρημα τέλειον, άνωθεν εστί καταβαίνον εκ σου του Πατρός των Φώτων...-
Απόσπασμα από Προσευχή Θείας Λειτουργίας Αγίου Χρυσοστόμου)

1

1-1. Εισαγωγή

Οι μελέτες διασταυρούμενης μετάβασης (στο εξής crossover δοκιμές) αποτελούν τυχαιοποιημένες μελέτες με κύριο πεδίο εφαρμογής τη φαρμακοβιομηχανία. Ειδικότερα, έχουν συχνή χρήση στις αρχικές φάσεις των κλινικών μελετών (φάσεις I και II) και από τον σωστό τους σχεδιασμό και εφαρμογή, εξαρτάται σε σημαντικό βαθμό η επιτυχία (ή μη) μιας κλινικής έρευνας.

Στο σημείο αυτό αξίζει να αναφερθεί ότι στόχος μιας κλινικής έρευνας είναι η τεκμηρίωση της ασφάλειας (safety) και της αποτελεσματικότητας (effectiveness) ενός ή περισσότερων υπό έρευνα δραστικών συστατικών – φαρμακευτικών σκευασμάτων. [1]

Επομένως, οι crossover δοκιμές όταν εφαρμόζονται στις κλινικές έρευνες είναι τυχαιοποιημένες μελέτες οι οποίες αποτελούν προγραμματισμένα ή

προσχεδιασμένα πειράματα που έχουν σκοπό να εξάγουν συμπεράσματα για το πώς οι – υπό έρευνα – θεραπευτικές αγωγές θα «συμπεριφερθούν» στο γενικό πληθυσμό των ασθενών που θα χρειαστούν θεραπεία μελλοντικά με βάση τα δεδομένα και τις πληροφορίες που προκύπτουν από ένα μικρό δείγμα αυτών. [2], [3]

1-2. Crossover δοκιμές

1-2-1. Ορισμός

Οι crossover δοκιμές είναι τυχαιοποιημένες μελέτες που - παρά το γεγονός ότι - δεν χρησιμοποιούνται συχνά, δεν είναι και σπάνιες. Σε μια crossover δοκιμή οι πειραματικές μονάδες (συνήθως ασθενείς) εκτίθενται σε μια σειρά (αλληλουχία) από διαφορετικές «θεραπείες». Σκοπός είναι να εντοπιστούν ενδεχόμενες στατιστικά σημαντικές διαφορές ανάμεσα σ' αυτές τις θεραπείες. Είναι προφανές ότι κάθε πειραματική μονάδα λειτουργεί ως μάρτυρας (control) του εαυτού της. [3]

Οι crossover δοκιμές διαφέρουν σημαντικά από τις μελέτες των παράλληλων ομάδων (parallel group analysis). Ο σχεδιασμός των parallel group μελετών είναι απλός και στηρίζεται στην τυχαιοποίηση του ασθενούς σε μια από τις δύο (ή περισσότερες) παράλληλες ομάδες της μελέτης με τη διαφορά ότι ο ασθενής παραμένει στην αρχική ομάδα μέχρι την ολοκλήρωση της έρευνας. [4]

1-2-2. Η αξία της τυχαιοποίησης (Randomization)

Τυχαιοποίηση (randomization) καλείται η διαδικασία στην οποία πραγματοποιείται κατανομή των ασθενών σε ομάδες - με τη βοήθεια μιας τυχαίας διαδικασίας. Μέσω της τυχαιοποίησης κάθε πειραματική μονάδα έχει την ίδια πιθανότητα να επιλεγεί στην ομάδα Α ή την ομάδα Β. Οι ομάδες αυτές διαφέρουν μόνο σε ότι αφορά τη θεραπευτική παρέμβαση. Με αυτό τον τρόπο επιτυγχάνεται η μείωση του σφάλματος κατανομής. Οι γνωστότερες μέθοδοι τυχαιοποίησης είναι οι ακόλουθες:

- α) Απλή τυχαιοποίηση
- β) Τυχαία μετατιθέμενες ομάδες (Blocks)
- γ) Μέθοδος μεροληπτικού νομίσματος (Biased coin method)
- δ) Μέθοδος ελαχιστοποίησης

Η επιτυχία μιας crossover δοκιμής - που ούτως ή άλλως αποτελεί είδος τυχαιοποιημένης μελέτης - εξαρτάται σε σημαντικό βαθμό από την ορθή και αμερόληπτη χρήση της επιλεγόμενης μεθόδου τυχαιοποίησης.

1-3. Πλεονεκτήματα – Μειονεκτήματα crossover δοκιμής

Τα βασικά πλεονεκτήματα της χρήσης crossover δοκιμών είναι τα ακόλουθα:

α) κάθε ασθενής λειτουργεί ως έλεγχος (control) του εαυτού του με αποτέλεσμα την αύξηση της ισχύος (power) της μελέτης δεδομένου ότι η μεταξύ των πειραματικών μονάδων (ασθενών) μεταβλητότητα είναι μικρότερη από εκείνη που παρατηρείται από τη σύγκριση διαφορετικών ασθενών. Ως εκ τούτου, απαιτείται μικρότερος αριθμός ασθενών – άρα μικρότερο δείγμα – σε σχέση με μια αντίστοιχη parallel group μελέτη. [4], [5]

β) κατά τη διάρκεια αλλά και μετά το πέρας της μελέτης δίνεται η δυνατότητα διερεύνησης των προτιμήσεων του ασθενούς. Για παράδειγμα, ο ασθενής μπορεί να ερωτηθεί για το ποια θεραπεία θεωρεί ότι είναι καλύτερη για τον ίδιο. [5]

Από την άλλη μεριά οι crossover μελέτες εμφανίζουν και αρκετά μειονεκτήματα:

α) δεν έχουν εφαρμογή σε περιπτώσεις στις οποίες – ο ασθενής έχει μεγάλη πιθανότητα είτε να αποβιώσει είτε να θεραπευτεί πλήρως από την ασθένεια. Συνεπώς, σε περιπτώσεις εξαιρετικά μολυσματικών ασθενειών ή επειγουσών καταστάσεων (π.χ οξεία πνευμονία) οι crossover δοκιμές δεν έχουν πεδίο εφαρμογής [5]

β) ύπαρξη αλληλεπίδρασης θεραπείας και περιόδου (treatment by period interaction) η οποία παρατηρείται όταν η επίδραση της θεραπείας δεν είναι σταθερή στη διάρκεια του χρόνου. Με απλά λόγια η ύπαρξη διαφορετικών συνθηκών σε διαφορετικές περιόδους μπορεί να έχει σημαντική επίδραση πάνω στους ασθενείς. [5] Η ύπαρξη σημαντικής αλληλεπίδρασης θεραπείας και περιόδου συνδέεται συχνά με την ύπαρξη του φαινομένου carry-over. [6], [7].

γ) ύπαρξη επίδρασης ακολουθίας (carry-over ή residual effects). Το φαινόμενο carry-over ερμηνεύεται ως η κατάσταση κατά την οποία η επίδραση της θεραπευτικής παρέμβασης μιας περιόδου εξακολουθεί να

υφίσταται και κατά την επόμενη περίοδο. Στα 2Χ2 crossover πειράματα η ύπαρξη σημαντικής carry-over επίδρασης οδηγεί αναπόφευκτα σε ύπαρξη σημαντικής αλληλεπίδρασης θεραπείας και περιόδου. Σε πιο πολύπλοκα crossover πειράματα αυτό μπορεί να μην ισχύει. [5], [7].

δ) εξαιτίας της πολυπλοκότητας των crossover δοκιμών - η οποία αυξάνεται στην περίπτωση ύπαρξης σημαντικής επίδρασης carry-over - καθίσταται δύσκολη η στατιστική επεξεργασία και η ανάλυση των δεδομένων. Αξίζει να σημειωθεί ότι η κατάσταση περιπλέκεται ακόμη περισσότερο αν η εξαρτημένη μεταβλητή (απόκριση) δεν είναι συνεχής μεταβλητή αλλά κατηγορική. [5], [7]

ε) η θεραπευτική παρέμβαση δεν ακολουθεί την κλινική πρακτική [7]

στ) υπάρχουν περιπτώσεις ασθενών οι οποίοι νιώθουν άβολα ή πολύ περισσότερο ενοχλούνται από το γεγονός ότι κατά τη διάρκεια μιας crossover μελέτης θα πρέπει να υποβληθούν σε περισσότερες της μιας θεραπευτικές παρεμβάσεις. Ακόμη, ο συνολικός χρόνος στον οποίο ένας ασθενής χρειάζεται να μείνει κάτω από παρακολούθηση είναι συχνά μεγάλος, πράγμα που λειτουργεί αποτρεπτικά για τη συμμετοχή του σε μια τέτοια μελέτη [5]

ζ) υπάρχουν ασθενείς που αποχωρούν από τη μελέτη πριν αυτή ολοκληρωθεί. Οι αποχωρήσεις ασθενών (drop outs) δημιουργούν σημαντικά προβλήματα σε μια crossover μελέτη τα οποία στις περισσότερες των περιπτώσεων είναι αξιόπεραστα. Για παράδειγμα, ένας ασθενής ο οποίος έχει υποβληθεί στη θεραπεία Α αποχωρεί από τη μελέτη, πράγμα που σημαίνει ότι δεν μπορεί να υπάρξει οποιοδήποτε πληροφορία από την εφαρμογή της θεραπείας Β ή Γ στο συγκεκριμένο ασθενή. [5]

1-4. Περίοδος έκπλυσης (wash out period)

Το χρονικό διάστημα μεταξύ δύο θεραπευτικών περιόδων καλείται περίοδος έκπλυσης (wash out period). Θεωρητικά η περίοδος έκπλυσης θα πρέπει να διαρκεί τόσο ώστε να εξλειφθεί η δράση της θεραπείας της 1^{ης} περιόδου από τον οργανισμό του ασθενή για να επιστρέψει στην προϋπάρχουσα της - αρχικής θεραπείας - κατάσταση, ώστε να είναι έτοιμος να δεχθεί τη θεραπεία της 2^{ης} περιόδου. Η περίοδος έκπλυσης εφαρμόζεται για να αποφευχθεί η επίδραση carry-over. Το χρονικό διάστημα της περιόδου έκπλυσης αποτελεί

απόφαση του ερευνητή σύμφωνα με την εμπειρία του και τη βιβλιογραφία. [4], [5].

1-5. Είδη επιδράσεων σε crossover δοκιμές

Στις crossover μελέτες είναι σημαντικό να προσδιοριστούν οι παράγοντες που επηρεάζουν σημαντικά την απόκριση του ασθενούς. Όταν λέμε απόκριση του ασθενούς εννοούμε στην ουσία τις τιμές που λαμβάνει η εξαρτημένη μεταβλητή που αποτελεί τη μεταβλητή ενδιαφέροντος για την οποία σχεδιάστηκε το πείραμα.

- **Επίδραση θεραπείας (treatment effects)**

Πρόκειται για την επίδραση της θεραπευτικής παρέμβασης αυτής καθεαυτής στις πειραματικές μονάδες (ασθενείς) κατά τη διάρκεια της περιόδου που αυτή εφαρμόζεται. Αποτελεί και το αντικείμενο μελέτης της έρευνας. [5],[7]

- **Επίδραση περιόδου (period effects)**

Η επίδραση περιόδου επηρεάζει μια crossover δοκιμή συνολικά. Έστω ότι σε ασθενείς δίνονται δύο παρόμοιες θεραπείες σε δύο χρονικές περιόδους. Είναι πιθανό οι μετρήσεις της εξαρτημένης μεταβλητής (απόκρισης) στην 1^η περίοδο να είναι σημαντικά μικρότερες ή μεγαλύτερες σε σχέση με εκείνες της 2^{ης} περιόδου. Γι' αυτό το λόγο στις crossover δοκιμές έχουμε συνήθως δύο ομάδες ασθενών που λαμβάνουν θεραπευτικές παρεμβάσεις με διαφορετική σειρά. [5],[7]

- **Επίδραση ακολουθίας (carry-over or residual effect)**

Η επίδραση ακολουθίας ή απλά επίδραση carry-over επηρεάζει την επίδραση της θεραπευτικής παρέμβασης (treatment effects). Επίδραση carry-over έχουμε όταν η αρχική θεραπεία συνεχίζει να ενεργεί και στη 2^η περίοδο αλλοιώνοντας με αυτόν τον τρόπο την επίδραση της 2^{ης} θεραπείας. Η ενσωμάτωση της περιόδου έκπλυσης (wash out period) στο πείραμα μπορεί να εξαφανίσει ή τουλάχιστον να περιορίσει σημαντικά την επίδραση carry-over. Η ύπαρξη σημαντικής επίδρασης carry-over σ' ένα πείραμα δεν είναι επιθυμητή και ο στόχος όλων των ερευνητών είναι η εξάλειψή του. [5],[7]

- **Αλληλεπίδραση θεραπείας και περιόδου (treatment by period interaction)**

Οι σπουδαιότεροι λόγοι στους οποίους υφίσταται αλληλεπίδραση θεραπείας και περιόδου είναι οι παρακάτω (για 2Χ2 crossover δοκιμή):

α) αν η περίοδος έκπλυσης είναι μικρή, η απόκριση της 2^{ης} περιόδου μπορεί να επηρεαστεί από τη θεραπευτική παρέμβαση της 1^{ης} όπως και από εκείνη της 2^{ης} περιόδου

β) ακόμη και αν η περίοδος έκπλυσης είναι ικανοποιητική, υπάρχουν περιπτώσεις όπου η επίδραση της θεραπείας της 1^{ης} περιόδου συνεχίζει να υφίσταται και στη 2^η περίοδο. Αυτό οφείλεται συχνότατα στον ψυχολογικό παράγοντα

γ) αν υπάρχει ισχυρή επίδραση περιόδου (period effects) οι διαφορετικές τιμές απόκρισης στις θεραπευτικές παρεμβάσεις της 1^{ης} και της 2^{ης} περιόδου μπορούν να επηρεάσουν τα τελικά συμπεράσματα της μελέτης. [5],[7]

Στο σημείο πρέπει να τονιστεί ότι η αλληλεπίδραση μεταξύ μεταβλητών γενικά θα πρέπει να διαχωριστεί από την επίδραση carry-over που συχνά παρατηρείται στις cross-over μελέτες. Ωστόσο στις 2Χ2 crossover δοκιμές η αλληλεπίδραση θεραπείας και περιόδου ταυτίζεται με την ύπαρξη επίδρασης carryover. Σε περισσότερο πολύπλοκα πειράματα όμως θα πρέπει να εξετάζεται χωριστά η αλληλεπίδραση θεραπείας και περιόδου με την ύπαρξη επίδρασης carryover (δεν υπάρχει ταύτιση).

1-6. Στατιστική ανάλυση

Η ανάλυση δεδομένων σε μια crossover μελέτη δεν αποτελεί εύκολη διαδικασία – το αντίθετο μάλιστα. Έχει αναφερθεί λεπτομερώς ότι ύπαρξη στατιστικά σημαντικής επίδρασης carry-over καθιστά τη στατιστική ανάλυση προβληματική και δύσκολη.

Είναι συνεπώς ζωτικής σημασίας, η ανάλυση να ξεκινά από το αν υφίσταται στατιστικά σημαντική επίδραση carry-over ως προς τις τιμές της εξαρτημένης μεταβλητής (απόκριση). Σε περίπτωση που υπάρχει σημαντική επίδραση carry-over τότε δεν έχει νόημα να ελεγχθεί η επίδραση περιόδου και η επίδραση θεραπείας χωριστά. Κατασκευάζουμε και παρατηρούμε ένα διάγραμμα μέσων τιμών για την απόκριση και εν συνεχεία αναλύουμε τα δεδομένα μόνο της 1^{ης} περιόδου σαν να πρόκειται για κλασσική μελέτη

παράλληλων ομάδων (parallel group analysis). Αυτή η προσέγγιση μπορεί να προκαλέσει μεγάλη αύξηση του σφάλματος τύπου I.

Η ύπαρξη σημαντικής επίδρασης carry-over συνεπάγεται και σημαντική ύπαρξη αλληλεπίδρασης θεραπείας και περιόδου. Σε αρκετά crossover πειράματα – εξαιτίας κακής τυχαιοποίησης – προκύπτει λανθασμένα ύπαρξη σημαντικής επίδρασης carry-over. Αυτό συμβαίνει εξαιτίας της υψηλής συσχέτισης μεταξύ των διαφορών στις μετρήσεις της απόκρισης της 1^{ης} περιόδου (1^η μεταβλητή) και της αλληλεπίδρασης θεραπείας και περιόδου (2^η μεταβλητή). Το παραπάνω προκαλεί στατιστικά σημαντική διαφορά στις διαφορές των μετρήσεων της απόκρισης της 1^{ης} περιόδου. Άρα, η ύπαρξη σημαντικής επίδρασης carry-over προκαλεί αρκετά και σημαντικά προβλήματα σ' ένα crossover πείραμα. [4], [8]

Σε περίπτωση απουσίας σημαντικής επίδρασης carry-over τότε για ένα 2X2 crossover πείραμα όπου η απόκριση είναι συνεχής μεταβλητή με τα δεδομένα να ακολουθούν την κανονική κατανομή, το test που προτείνεται είναι ο έλεγχος paired t test. Αν τα δεδομένα δεν ακολουθούν τον κανονικό νόμο τότε χρησιμοποιείται το κατάλληλο μη παραμετρικό test. Για μεγαλύτερα crossover πειράματα χρησιμοποιούνται γενικά γραμμικά μοντέλα π.χ μεικτά γραμμικά μοντέλα (mixed linear models). Ακόμη, για τον ίδιο σκοπό μπορεί να χρησιμοποιηθεί και η Μπεϋζιανή (Bayesian) προσέγγιση ανάλυσης δεδομένων. [4], [8]

Είναι σημαντικό να τονιστεί ότι οι ασθενείς που αποχωρούν μετά την 1^η θεραπευτική περίοδο (drop outs) δεν συμμετέχουν στην στατιστική ανάλυση για την αποφυγή συστηματικών σφαλμάτων. [4]

1-7. Χρήση crossover δοκιμών

Οι crossover δοκιμές χρησιμοποιούνται συχνά στα πρώτα στάδια κλινικών ερευνών π.χ φάση I και φάση II καθώς και σε προκλινικές δοκιμές (φάση 0) που διεξάγονται σε πειραματόζωα. Ειδικότερα έχουν χρήση σε:

- α) Φαρμακοκινητικές και φαρμακοδυναμικές μελέτες
- β) Κλινική φαρμακολογία
- γ) Μελέτες βιοϊσοδυναμίας
- δ) Μελέτες στη βιομηχανία τροφίμων

ε) Μελέτες για τον προσδιορισμό της μέγιστης ανεκτής δόσης για το υπό έρευνα φαρμακευτικό σκεύασμα (maximum tolerated dose).

Σπανιότερα χρησιμοποιούνται σε φάσεις III και IV κλινικών ερευνών π.χ σε μελέτες για τη διερεύνηση προτιμήσεων ασθενών κλπ. Αξίζει να σημειωθεί ότι οι crossover μελέτες έχουν εφαρμογή κυρίως σε χρόνια νοσήματα όπως το άσθμα, η επιληψία, η υψηλή αρτηριακή πίεση, η πολλαπλή σκλήρωση, οι διαταραχές ύπνου, η αρθρίτιδα κλπ. Αντίθετα οι crossover δοκιμές δεν έχουν εφαρμογή σε επείγουσες ιατρικές καταστάσεις (π.χ οξύ έμφραγμα μυοκαρδίου, πνευμονία κλπ).

Κεφάλαιο 2

Λατινικά Τετράγωνα - Μεικτά Μοντέλα

(Μετά το παιχνίδι, ο βασιλιάς και το πιόνι πάνε στο ίδιο κουτί – Ιταλική παροιμία)

2-1. Πειραματικοί σχεδιασμοί

2-1-1. Εισαγωγή

Ως πείραμα ορίζεται μια δοκιμή (test) ή μια σειρά δοκιμών κατά τις οποίες πραγματοποιούνται συγκεκριμένες αλλαγές σ' ένα σύνολο μεταβλητών που περιγράφουν διαδικασίες ή συστήματα. Σκοπός ενός πειράματος είναι να παρατηρήσουμε και εν συνεχεία να αναγνωρίσουμε τα αίτια για τις μεταβολές που υπάρχουν στα αποτελέσματα.

Είναι φανερό ότι σ' ένα πείραμα, το υπό έρευνα σύστημα βρίσκεται κάτω από τον πλήρη έλεγχο του ερευνητή. Τούτο σημαίνει ότι τα υποκείμενα ή το υλικό που ερευνάται, η φύση των παρεμβάσεων καθώς και οι μέθοδοι μετρήσεων που χρησιμοποιούνται καθορίζονται απολύτως από τον ερευνητή.

Για παράδειγμα, έστω ότι έχουμε μια τυχαιοποιημένη κλινική μελέτη όπου οι πειραματικές μονάδες είναι ασθενείς ή υγιείς εθελοντές οι οποίοι κατανέμονται

με τυχαίο τρόπο σε δύο ομάδες. Αξίζει να σημειωθεί ότι οι ασθενείς πληρούν τα κριτήρια επιλογής που προβλέπει το πρωτόκολλο και έχουν υπογράψει έντυπο συγκατάθεσης. Στην 1^η ομάδα ασθενών χορηγείται ένα νέο φαρμακευτικό σκεύασμα T και στη 2^η ομάδα ασθενών χορηγείται εικονικό φάρμακο P. Οι ασθενείς παρακολουθούνται για μια συγκεκριμένη χρονική περίοδο κατά την οποία λαμβάνονται διάφορες μετρήσεις που αφορούν στην μεταβλητή ενδιαφέροντος η οποία καλείται εξαρτημένη μεταβλητή (απόκριση). [10]

Επομένως, ένα πείραμα μπορεί να πραγματοποιηθεί για:

- α) να καθορίσει τις βασικές αιτίες της μεταβλητότητας της εξαρτημένης μεταβλητής (απόκρισης)
- β) να εντοπίσει της συνθήκες που προκαλούν τη μέγιστη ή ελάχιστη τιμή στην εξαρτημένη μεταβλητή (απόκριση)
- γ) να συγκρίνει τις τιμές στην εξαρτημένη μεταβλητή (απόκριση) που προκαλούνται από μεταβολές στις επεξηγηματικές μεταβλητές και
- δ) να κατασκευάσει ένα μαθηματικό μοντέλο για να υπάρχει η δυνατότητα για μελλοντικές προβλέψεις. [11]

10

Πειραματικός σχεδιασμός (experimental design) είναι το αποτέλεσμα μιας λεπτομερούς διαδικασίας που πραγματοποιείται κατά τη διάρκεια ενός πειράματος. Η επιλογή του πειραματικού σχεδιασμού εξαρτάται κυρίως από τον αριθμό των μεταβλητών (συχνά καλούνται και συντελεστές) που εξετάζονται. Ένας συντελεστής μπορεί να είναι συνεχής (π.χ ύψος, θερμοκρασία, αρτηριακή πίεση) ή διακριτός (π.χ επίπεδο ικανοποίησης, φύλο κλπ). Οι δυνατές τιμές που μπορεί να λάβει ένας συντελεστής (μεταβλητή) ονομάζονται επίπεδα.

Κάθε πείραμα αποτελείται από ένα σύνολο πειραματικών μονάδων που ονομάζονται πεδία ή δοκιμές (π.χ ασθενείς, ζώα, αγροτεμάχια κλπ). Οι συγκρίσεις μεταξύ των διαφορετικών εφαρμογών που ακολουθούνται μεταξύ των επιπέδων των μεταβλητών και μεταξύ των συνδυασμών των εφαρμογών, γίνονται με βάση τις πληροφορίες που συλλέγονται - μετά από κάθε εφαρμογή του πειράματος- για τις επιδράσεις που υφίσταται το πειραματικό αντικείμενο. [11]

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

2-1-2. Βασικές αρχές πειραματικού σχεδιασμού

Ο στατιστικός σχεδιασμός ενός πειράματος έχει ως στόχο τη στατιστική επεξεργασία των δεδομένων (data) που προκύπτουν (από το πείραμα). Αποτέλεσμα αυτής της επεξεργασίας είναι η εξαγωγή χρήσιμων και αντικειμενικών συμπερασμάτων γύρω από την υπόθεση που ερευνάται. Η στατιστική προσέγγιση σ' ένα πειραματικό σχεδιασμό είναι απαραίτητο στοιχείο για την επιτυχία του. Στην περίπτωση που τα δεδομένα ενός πειράματος υπόκεινται σε πειραματικά λάθη, οι στατιστικές μέθοδοι αποτελούν την μόνη αντικειμενική προσέγγιση για τον εντοπισμό τους. Συνεπώς, ο σχεδιασμός και η στατιστική ανάλυση των δεδομένων αποτελούν δύο κύριες προϋποθέσεις για την τέλεση ενός πειράματος. [10],[11]

Υπάρχουν τρεις βασικές αρχές στον πειραματικό σχεδιασμό. Αυτές είναι η επαναληψιμότητα (replication), η τυχαιοποίηση (randomization) και η ομαδοποίηση (blocking).

Με τον όρο επαναληψιμότητα εννοείται η δυνατότητα επανάληψης του βασικού πειράματος. Η επαναληψιμότητα έχει δύο σημαντικές ιδιότητες. Πρώτον, επιτρέπει στον ερευνητή να εκτιμήσει το σφάλμα που προκύπτει από το πείραμα. Αυτή η εκτίμηση του σφάλματος αποτελεί μέγεθος με το οποίο μπορεί ο ερευνητής να αποφανθεί αν οι παρατηρούμενες διαφορές στα δεδομένα είναι στατιστικά σημαντικές. Δεύτερον, αν για παράδειγμα, ο δειγματικός μέσος χρησιμοποιείται για την εκτίμηση της επίδρασης ενός παράγοντα του πειράματος τότε η δυνατότητα επανάληψης (του πειράματος) επιτρέπει στον ερευνητή να πετύχει μεγαλύτερη ακρίβεια στην εκτίμηση αυτής της επίδρασης. [9],[10],[11]

Η τυχαιοποίηση αποτελεί τον ακρογωνιαίο λίθο πάνω στον οποίο στηρίζεται η χρήση των στατιστικών μεθόδων στον πειραματικό σχεδιασμό. Είναι πολύ σημαντικό σ' ένα πείραμα (π.χ σε μια κλινική μελέτη) να πραγματοποιείται κατανομή των ασθενών σε δύο ή περισσότερες ομάδες με τέτοιο τρόπο ώστε αυτές να είναι σχεδόν ίδιες (παρόμοιες) όσον αφορά πιθανούς συγχυτικούς παράγοντες. Ο στόχος αυτός επιτυγχάνεται με την τυχαιοποίηση στην οποία κάθε υποκείμενο κατανέμεται σε μια ομάδα ανεξάρτητα από τα υπόλοιπα υποκείμενα και επιπλέον δεν επηρεάζεται από τις προτιμήσεις του ερευνητή.

Με την τυχαιοποίηση αποφεύγονται τα συστηματικά σφάλματα (systematic bias), τα σφάλματα επιλογής (selection bias), τα τυχαία σφάλματα (accidental bias) και τα σφάλματα που οφείλονται στον ίδιο τον ερευνητή. Υπάρχουν αρκετές μέθοδοι τυχαιοποίησης, κάποιες από τις οποίες αναφέρθηκαν σε αδρές γραμμές στο πρώτο κεφάλαιο. [9],[10],[11]

Η ομαδοποίηση (blocking) αποτελεί τεχνική με την οποία βελτιώνεται η ακρίβεια των συγκρίσεων που πραγματοποιούνται μεταξύ των παραγόντων που ερευνώνται. Σε ένα πείραμα - είναι σύνηθες - η μεταβλητότητα που προκύπτει από ένα «ενοχλητικό» παράγοντα (nuisance factor) να επηρεάζει το αποτέλεσμα. Στην περίπτωση που ο παράγοντας αυτός είναι γνωστός και ελέγξιμος, χρησιμοποιούμε το blocking ως τεχνική για να ελαττώσουμε ή να εξαφανίσουμε (πολύ δύσκολο) τη μεταβλητότητα που προκύπτει από τον «ενοχλητικό» παράγοντα. [9],[10],[11]

Για παράδειγμα, έστω ότι επιθυμούμε να ελέγξουμε την αντοχή 4 μηχανών (παραμβάσεις - treatments). Για να το πετύχουμε, τοποθετούμε στην κάθε μηχανή 4 μεταλλικές ράβδους (blocks). Από το μέγεθος της παραμόρφωσης την οποία θα υποστούν οι ράβδοι μετρείται η αντοχή των μηχανών. Ο σχεδιασμός του πειράματος δίνεται στον ακόλουθο πίνακα:

Πίνακας 2-1

Μετρήσεις παραδείγματος αντοχής μηχανών

Μηχανές (Παραμβάσεις)	Ράβδοι (Blocks)			
	1 ^η	2 ^η	3 ^η	4 ^η
1 ^η	Y ₁₁	Y ₁₂
2 ^η	Τιμές Απόκρισης (Μέγεθος παραμόρφωσης)	
3 ^η		
4 ^η	Y ₄₄

Αν χρησιμοποιήσουμε ανάλυση διακύμανσης ως προς ένα παράγοντα (one way ANOVA) τότε από το σύνολο των ράβδων που έχουμε, θα επιλέγαμε

τυχαία τη σειρά με την οποία οι ράβδοι θα δοκιμάζονταν στην κάθε μηχανή. Τούτο όμως δημιουργεί ένα σοβαρό πρόβλημα. Αν οι ράβδοι διαφέρουν έστω και λίγο μεταξύ τους (π.χ λόγω διαφορετικής θερμοκρασίας κατασκευής) τότε στη μεταβλητότητα που ίσως υπάρχει μεταξύ των μηχανών θα προστεθεί και η μεταβλητότητα ανάμεσα στις ράβδους, αλλοιώνοντας μ' αυτό τον τρόπο τα αποτελέσματα. Για να αντιμετωπιστεί η μεταβλητότητα ανάμεσα στις ράβδους, χρησιμοποιούμε τυχαίοποιημένο σχεδιασμό κατά ομάδες (blocks), σύμφωνα με τον οποίο επιθυμούμε η κάθε μηχανή i σε κάθε σειρά δοκιμών j να δοκιμάζεται με την ίδια ράβδο. Για να γίνει αυτό κόβουμε την κάθε ράβδο σε 4 – όσες και οι μηχανές – κομμάτια. Κάθε block (ράβδος) χρησιμοποιείται τόσες φορές όσες και οι παρεμβάσεις (μηχανές). Κάθε κομμάτι j της ράβδου που θα δοκιμαστεί στην i μηχανή επιλέγεται τυχαία. [9], [10]

Είναι σημαντικό να τονιστεί ότι οι τρεις βασικές αρχές του πειραματικού σχεδιασμού επαναληψιμότητα, τυχαίοποίηση και ομαδοποίηση (blocking) αποτελούν μέρος κάθε πειράματος.

2-1-3. Στάδια πειραματικού σχεδιασμού

Σε αδρές γραμμές τα βήματα για τον σχεδιασμό ενός πειράματος είναι τα ακόλουθα:

- α) Παρουσίαση και έκθεση του προβλήματος
- β) Επιλογή παραγόντων – επεξηγηματικών μεταβλητών
- γ) Επιλογή εξαρτημένης μεταβλητής (απόκρισης)

Τα τρία πρώτα βήματα αφορούν στο λεγόμενο προπειραματικό σχεδιασμό. Στην πράξη τα βήματα (β) και (γ) λαμβάνουν χώρα συγχρόνως.

- δ) Επιλογή πειραματικού σχεδιασμού
- ε) Πραγματοποίηση του πειράματος – Μετρήσεις
- στ) Στατιστική ανάλυση των δεδομένων
- ζ) Συμπεράσματα και συστάσεις. [9]

Ένας συνηθισμένος πειραματικός σχεδιασμός που χρησιμοποιείται συχνά από τους ερευνητές είναι τα λατινικά τετράγωνα (Latin squares). [9]

2-2. Λατινικά Τετράγωνα (Latin Squares)

2-2-1. Ιστορική αναδρομή

Το 1779 ο Ελβετός μαθηματικός Leonhard Euler μελέτησε το μαθηματικό γρίφο των 36 αξιωματικών. 36 αξιωματικοί επιλεγμένοι από 6 διαφορετικούς βαθμούς και 6 διαφορετικά τάγματα στρατού έπρεπε να τοποθετηθούν σ' ένα τετράγωνο έτσι ώστε σε κάθε οριζόντια και κάθετη γραμμή να υπάρχουν 6 αξιωματικοί από κάθε βαθμό και τάγμα. Ο Euler καταγράφοντας μόνο τους βαθμούς των αξιωματικών δημιούργησε ένα λατινικό τετράγωνο. Ομοίως, καταγράφοντας μόνο τα τάγματα δημιούργησε ακόμη ένα λατινικό τετράγωνο. Όταν ο Euler τοποθέτησε τα δύο λατινικά τετράγωνα που δημιούργησε το ένα πάνω στο άλλο διαπίστωσε ότι τα ζευγάρια δεν ήταν όλα διαφορετικά μεταξύ τους. Συνεπώς δεν ήταν ορθογώνια.

Ο Euler κατόρθωσε να κατασκευάσει ζεύγη ορθογώνιων λατινικών τετραγώνων για κάθε τάξη περιττού αριθμού και για κάθε τάξη που διαιρείται με το 4, αλλά δεν κατάφερε να κατασκευάσει κανένα ζεύγος λατινικών τετραγώνων τάξης $4k+2, k \geq 0$ όπου k ακέραιος αριθμός. Ως εκ τούτου ο Euler υπέθεσε ότι δεν υπάρχουν λατινικά τετράγωνα 6^{ης} τάξης. Τον 20^ο αιώνα οι Bose, Shrikhande και Parker (1960) απέδειξαν ότι η εικασία του Euler ήταν λανθασμένη για όλες τις τάξεις n της μορφής $n = 4k + 2$, εκτός $n = 2$ ή $n = 6$ δίνοντας μια μέθοδο κατασκευής ζευγών ορθογωνίων λατινικών τετραγώνων για όλες αυτές τις τάξεις. Η απόδειξη αυτών υπήρξε μακροσκελής και εμπειρείχε έννοιες από τη θεωρία πειραματικού σχεδιασμού. Η έρευνα για τη δομή και τις ιδιότητες των λατινικών τετραγώνων συνεχίζεται. Ένα πολύ γνωστό είδος λατινικού τετραγώνου αποτελεί το παιχνίδι sudoku (ειδικό λατινικό τετράγωνο). [12]

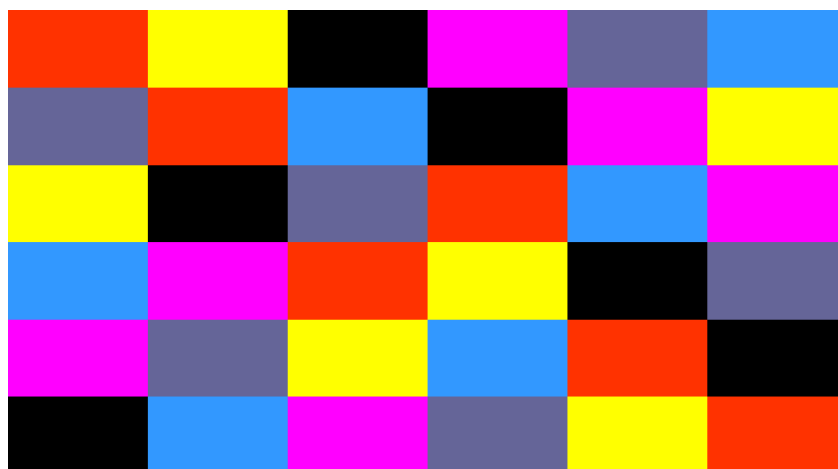
2-2-2. Ορισμός και βασικές έννοιες

Ένα λατινικό τετράγωνο τάξης p είναι μια μήτρα (πίνακας) $p \times p$ με ακριβώς p διαφορετικά σύμβολα όπου κάθε σύμβολο εμφανίζεται μια φορά σε κάθε γραμμή και μια φορά σε κάθε στήλη. [12]

Ένα λατινικό τετράγωνο έχει τη μορφή:

Σχήμα 2-1

6X6 Latin Square



Παραδείγματα λατινικών τετραγώνων τάξης 4 και τάξης 5 δίνονται στα ακόλουθα σχήματα:

Σχήμα 2-2

4X4 Latin Square

1	2	3	4
4	1	2	3
3	4	1	2
2	3	4	1

Σχήμα 2-3

5X5 Latin Square

1	2	3	4	5
2	1	5	3	4
3	4	1	5	2
4	5	2	1	3
5	3	4	2	1

15

Ένας απλός τρόπος κατασκευής λατινικού τετραγώνου είναι η δημιουργία της 1^{ης} γραμμής τοποθετώντας τα σύμβολα σε τυχαία σειρά και εν συνεχεία η μετακίνηση της 1^{ης} γραμμής προς τα δεξιά (ή αριστερά) κατά 1,2,... $p-1$ θέσεις ώστε να δημιουργηθούν και οι υπόλοιπες $p-1$ γραμμές. [13].

Ένα λατινικό τετράγωνο λέγεται μειωμένο ή κανονικοποιημένο αν η πρώτη του γραμμή και στήλη είναι σε φυσική σειρά. Η κανονικοποίηση ενός λατινικού τετραγώνου πραγματοποιείται με αντιμεταθέσεις των γραμμών και των στηλών του. Αν - για παράδειγμα στο λατινικό τετράγωνο τάξης 4 (σχήμα 2-2) – αντιμεταθέσουμε τη 2^η γραμμή με την 4^η προκύπτει:

Σχήμα 2-4

Κανονικοποιημένο 4X4 Latin Square

1	2	3	4
2	3	4	1
3	4	2	2
4	1	2	3

Το λατινικό τετράγωνο τάξης 5 του σχήματος 2-3 είναι ήδη κανονικοποιημένο.

Το πλήθος των λατινικών τετραγώνων τάξης p είναι $p!(p-1)!$ φορές το πλήθος των μειωμένων λατινικών τετραγώνων. Όσο αυξάνεται η τάξη p τόσο αυξάνεται και το πλήθος των μειωμένων λατινικών τετραγώνων. [14]

2-2-3. Είδη λατινικών τετραγώνων

Υπάρχουν διάφορα είδη λατινικών τετραγώνων και παρακάτω θα περιγράψουμε εν συντομία ορισμένα που παρουσιάζουν ιδιαίτερο ενδιαφέρον:

α) Διαγώνιο λατινικό τετράγωνο τάξης p είναι ένα λατινικό τετράγωνο στο οποίο κανένα σύμβολο δεν εμφανίζεται πάνω από μια φορά σε οποιαδήποτε από τις δύο κύριες διαγωνίους του [15]

β) Δύο λατινικά τετράγωνα τάξης p καλούνται ορθογώνια αν η υπέρθεσή τους αποτελείται από p^2 διαφορετικά ζεύγη συμβόλων ή παρόμοια αν η υπέρθεσή τους περιλαμβάνει όλους τους δυνατούς συνδυασμούς των συμβόλων των 2 λατινικών τετραγώνων. [12]

γ) Τέλειο λατινικό τετράγωνο τάξης p^2 είναι ένα διαγώνιο λατινικό τετράγωνο τάξης p^2 στο οποίο κάθε σύμβολο εμφανίζεται μία φορά σε κάθε κύριο υποτετράγωνο. Το υποτετράγωνο $S_{i,j}$ ενός λατινικού τετραγώνου τάξης p^2 ορίζεται ως ένα $p \times p$ τετράγωνο του οποίου το πάνω αριστερά κελί έχει συντεταγμένες (i, j) . Το παιχνίδι Sudoku είναι ένα τέλειο λατινικό τετράγωνο τάξης 9. [15]

2-2-4. Τυχαιοποίηση λατινικών τετραγώνων

Έχουμε ήδη αναφέρει ότι ένα $p \times p$ λατινικό τετράγωνο είναι ένα τετράγωνο που περιλαμβάνει p γραμμές και p στήλες. Το πλήθος των αποτελεσμάτων p^2 συμβολίζεται συνήθως με κεφαλαία γράμματα του λατινικού αλφαβήτου A, B, C κ.ο.κ και κάθε ένα από αυτά απαντάται μία (και μόνο μία) φορά σε κάθε γραμμή και σε κάθε στήλη. Από τη χρήση των γραμμμάτων του λατινικού αλφαβήτου πήραν και την ονομασία τους τα λατινικά τετράγωνα.

Για τη χρήση των λατινικών τετραγώνων σ' ότι αφορά στον πειραματικό σχεδιασμό και σε σχέση με την τυχαιοποίηση έχουμε να επισημάνουμε τα ακόλουθα:

- α) Κάθε γραμμή και κάθε στήλη αποτελεί μια πλήρη ομάδα (block)
- β) Σε κάθε γραμμή και στήλη τυχαιοποιούνται ανεξάρτητα οι παρεμβάσεις. Οι παρεμβάσεις συμβολίζονται με κεφαλαία γράμματα A, B, C κ.λπ.
- γ) Με τη βοήθεια ειδικών πινάκων επιλέγουμε ένα σχέδιο λατινικού τετραγώνου
- δ) Τυχαιοποιούμε τις γραμμές
- ε) Τυχαιοποιούμε τις στήλες
- στ) Τυχαιοποιούμε τις παρεμβάσεις
- ζ) Πραγματοποιούμε τον τελικό σχεδιασμό του πειράματος

2-2-5. Μαθηματικό υπόδειγμα λατινικών τετραγώνων

Η χρήση του υποδείγματος των λατινικών τετραγώνων έχει ως σκοπό την ελάττωση του πειραματικού σφάλματος και την αύξηση της ευαισθησίας του πειράματος. Ταυτόχρονα επιχειρείται ο έλεγχος των δύο γνωστών πηγών μεταβλητότητας (γραμμές και στήλες) με στόχο τον περιορισμό ή ακόμη καλύτερα την απομάκρυνση της επίδρασης της μεταβλητότητας αυτών των πηγών. [16]

Θεωρούμε ένα $p \times p$ λατινικό τετράγωνο το οποίο περιλαμβάνει p γραμμές και p στήλες. Το στατιστικό μοντέλο ενός λατινικού τετραγώνου έχει τη μορφή:

$$Y_{ijk} = \mu + \alpha_i + \tau_j + b_k + e_{ijk} \quad \begin{array}{l} i = 1, 2, \dots, p \\ j = 1, 2, \dots, p \\ k = 1, 2, \dots, p \end{array} \quad (2-1)$$

όπου Y_{ijk} = η παρατήρηση της i γραμμής και της k στήλης για τον j συνδυασμό

μ = ο συνολικός μέσος

α_i = το αποτέλεσμα της i γραμμής

τ_j = το αποτέλεσμα του j συνδυασμού

b_k = το αποτέλεσμα της k στήλης και

e_{ijk} = τα τυχαία σφάλματα που ακολουθούν $N(0, \sigma^2)$ [16]

Από το μαθηματικό υπόδειγμα ενός λατινικού τετραγώνου παρατηρούμε ότι αυτό περιλαμβάνει:

α) p παρεμβάσεις (treatments)

β) δύο blocking μεταβλητές, όπου η κάθε μία περιλαμβάνει p τάξεις

γ) κάθε γραμμή και κάθε στήλη περιλαμβάνει το σύνολο των παρεμβάσεων. [17]

Δεν υπάρχει αλληλοσυσχέτιση μεταξύ γραμμών, στηλών και παρεμβάσεων (treatments) οπότε το υπόδειγμα είναι αθροιστικό.

Τα σύνολα ως προς τις γραμμές, τις στήλες, τις παρεμβάσεις και το γενικό (σύνολο) καθώς και οι αντίστοιχοι μέσοι τους δίνονται από τις σχέσεις:

$$\begin{aligned} Y_{i..} &= \sum_j Y_{ijk} & \bar{Y}_{i..} &= \frac{Y_{i..}}{p} \\ Y_{.j.} &= \sum_i Y_{ijk} & \bar{Y}_{.j.} &= \frac{Y_{.j.}}{p} \\ Y_{..k} &= \sum_{i,j} Y_{ijk} & \bar{Y}_{..k} &= \frac{Y_{..k}}{p} \\ Y_{...} &= \sum_i \sum_j Y_{ijk} & \bar{Y}_{...} &= \frac{Y_{...}}{p^2} \end{aligned}$$

Ακόμη οι εκτιμητές ελαχίστων τετραγώνων και μεγίστης πιθανοφάνειας των παραμέτρων ενός μοντέλου λατινικών τετραγώνων είναι:

Πίνακας 2-2

Εκτιμητές υποδείγματος λατινικών τετραγώνων

Παράμετρος	Εκτιμητής
$\mu_{...}$	$\hat{\mu} = \bar{Y}_{...}$
α_i	$\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$
τ_j	$\hat{\tau}_j = \bar{Y}_{.j.} - \bar{Y}_{...}$
b_k	$\hat{b}_k = \bar{Y}_{..k} - \bar{Y}_{...}$

Οι εκτιμώμενες από το μοντέλου τιμές δίνονται από τη σχέση:

$$\hat{Y}_{ijk} = \bar{Y}_{i..} + \bar{Y}_{.j.} + \bar{Y}_{..k} - 2\bar{Y}_{...} \quad (2-2)$$

και τα υπόλοιπα από τη σχέση:

$$e_{ijk} = Y_{ijk} - \hat{Y}_{ijk} \quad (2-3)$$

Το συνολικό άθροισμα των τετραγώνων (total sum of squares) των N παρατηρήσεων δίνεται από τη σχέση:

$$SS_T = SS_{ROWS} + SS_{COLUMNS} + SS_{TREATS} + SS_E \text{ με } p^2 - 1 \text{ βαθμούς ελευθερίας} \quad (2-4)$$

Οι βαθμοί ελευθερίας (degrees of freedom) του συνολικού αθροίσματος τετραγώνων προκύπτουν από τη σχέση:

$$p^2 - 1 = p - 1 + p - 1 + p - 1 + (p - 2)(p - 1) \quad (2-5)$$

Τα αθροίσματα τετραγώνων (sum of squares) των γραμμών, των στηλών, των παρεμβάσεων και των σφαλμάτων όταν διαιρεθούν με τους ανάλογους βαθμούς ελευθερίας μας δίνουν τα αντίστοιχα μέσα τετράγωνα (mean squares). Τα μέσα τετράγωνα χρησιμοποιούνται για των υπολογισμό των F στατιστικών. Η υπολογιστική διαδικασία της ανάλυσης διακύμανσης δίνεται από τον ακόλουθο πίνακα:

Πίνακας 2-3

Ανάλυση διακύμανσης για λατινικά τετράγωνα με σταθερές επιδράσεις

Πηγή μεταβλητότητας	Άθροισμα τετραγώνων	Βαθμοί ελευθερίας	Μέσο τετράγωνο	Στατιστικό F
Γραμμές	$SS_{ROW} = p \sum_i (\bar{Y}_{i..} - \bar{Y}_{...})^2$	$p - 1$	$MS_{ROW} = \frac{SS_{ROW}}{p - 1}$	$F_{ROW} = \frac{MS_{ROW}}{MS_E}$
Στήλες	$SS_{COL} = p \sum_j (\bar{Y}_{.j.} - \bar{Y}_{...})^2$	$p - 1$	$MS_{COL} = \frac{SS_{COL}}{p - 1}$	$F_{COL} = \frac{MS_{COL}}{MS_E}$
Μεταβλητές (Παρεμβάσεις)	$SS_{TREAT} = p \sum_j (\bar{Y}_{.k.} - \bar{Y}_{...})^2$	$p - 1$	$MS_{TREAT} = \frac{SS_{TREAT}}{p - 1}$	$F_{TREAT} = \frac{MS_{TREAT}}{MS_E}$
Σφάλμα	$SS_E = \sum_i \sum_j (Y_{ijk} - \bar{Y}_{i..} - \bar{Y}_{.j.} - \bar{Y}_{.k.} + 2\bar{Y}_{...})^2$	$(p-1)(p-2)$	$MS_E = \frac{SS_E}{(p-1)(p-2)}$	
Σύνολο	$SS_T = SS_{ROW} + SS_{COL} + SS_{TREATS} + SS_E$	$p^2 - 1$		

Για να ελεγχθεί η επίδραση των παρεμβάσεων (treatments) σ' ένα λατινικό τετράγωνο ελέγχουμε την υπόθεση:

$$H_o : \tau_j = 0$$

$$\text{vs} \quad \mu\epsilon \quad j = 1, 2, \dots, p$$

$$H_1 : \tau_j \neq 0$$

Κάνοντας χρήση του στατιστικού $F_{TREAT} = \frac{MS_{TREAT}}{MS_E}$ ελέγχουμε την παραπάνω

υπόθεση.

-Αν $F_{TREAT} \leq F_{1-\alpha, p-1, (p-1)(p-2)}$ τότε η H_o δεν απορρίπτεται

-Αν $F_{TREAT} > F_{1-\alpha, p-1, (p-1)(p-2)}$ τότε η H_o απορρίπτεται

Η καταλληλότητα του μοντέλου ελέγχεται με τη βοήθεια της ανάλυσης υπολοίπων. Σ' ένα λατινικό τετράγωνο τα υπολοιπά δίνονται από τη σχέση (2-3). Με τη βοήθεια γραφημάτων και αριθμητικών μεθόδων ελέγχεται:

α) η υπόθεση της κανονικότητας των υπολοίπων

β) η υπόθεση των ίσων διακυμάνσεων (ομοσκεδαστικότητα) των υπολοίπων ως προς τις μεταβλητές (treatments), τις γραμμές και τις στήλες και

γ) η υπόθεση της μη αλληλεπίδρασης (ανεξαρτησίας) ανάμεσα στα treatments, τις γραμμές και τις στήλες κάτι το οποίο υποθέσαμε και προηγουμένως.

Η ισχύς του στατιστικού F_{TREAT} για την επίδραση των παρεμβάσεων σ' ένα λατινικό τετράγωνο δίνεται από τη σχέση:

$$\Phi = \frac{1}{\sigma} \sqrt{\sum \tau_j^2} \quad [16] \quad (2-6)$$

2-2-6. Παράδειγμα 5X5 λατινικού τετραγώνου

Έστω ότι έχουμε 5 διαφορετικά είδη φόρμουλας που χρησιμοποιούνται στους προωστήρες πυραύλων. Κάθε φόρμουλα είναι κατασκευασμένη από διαφορετικό χειριστή (με διαφορετική εμπειρία και ικανότητα) και είναι φτιαγμένη από διαφορετικές παρτίδες υλικού. Τονίζεται ότι κάθε φόρμουλα συνδέεται με ένα χειριστή και απαντάται ακριβώς μία φορά σε κάθε παρτίδα.

Τα δεδομένα δίνονται στον ακόλουθο πίνακα:

Πίνακας 2-4

Δεδομένα παραδείγματος

Παρτίδες	Χειριστές				
	1ος	2ος	3ος	4ος	5ος
1 ^η	24	20	19	24	24
2 ^η	17	24	30	27	36
3 ^η	18	38	26	27	21
4 ^η	26	31	26	23	22
5 ^η	22	30	20	29	31

Το συνολικό άθροισμα τετραγώνων (total sum of squares) δίνεται από τη σχέση:

$$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N} = 680 - \frac{(10)^2}{25} = 676$$

Το άθροισμα τετραγώνων ως προς τις παρτίδες είναι:

$$SS_{Part} = \frac{1}{p} \sum_{i=1}^p y_{i..}^2 - \frac{y_{...}^2}{N} = 68$$

Το άθροισμα τετραγώνων ως προς τους χειριστές είναι:

$$SS_{Chir} = \frac{1}{p} \sum_{k=1}^p y_{..k}^2 - \frac{y_{...}^2}{N} = 150$$

Το άθροισμα τετραγώνων ως προς τις φόρμουλες (treatments) είναι:

$$SS_{Form} = \frac{1}{p} \sum_{j=1}^p y_{.j.}^2 - \frac{y_{...}^2}{N} = 330$$

Το άθροισμα τετραγώνων των σφαλμάτων ισούται με:

$$SS_E = SS_T - SS_{Part} - SS_{Chir} - SS_{Form} = 128$$

Με τη βοήθεια των παραπάνω ο πίνακας διακύμανσης του 5Χ5 λατινικού τετραγώνου γίνεται:

Πίνακας 2-5

Ανάλυση διακύμανσης για παράδειγμα

Πηγή μεταβλητότητας	Άθροισμα τετραγώνων	Βαθμοί ελευθερίας	Μέσο τετράγωνο	Στατιστικό F
Γραμμές (Παρτίδες)	68	4	17	
Στήλες (Χειριστές)	150	4	37,5	
Παρεμβάσεις (Φόρμουλες)	330	4	82,5	7,73
Σφάλμα	128	12	10,67	
Σύνολο	676	24		

Από την τιμή του στατιστικού $F = 7.73$ και συγκρίνοντάς το με την τιμή $F_{0.05,4,12}$ παρατηρούμε ότι $F > F_{0.05,4,12}$ οπότε η υπόθεση H_0 απορρίπτεται. Επομένως, υφίσταται στατιστικά σημαντική διαφορά στο μέσο ποσοστό καύσης ανάμεσα στους διαφορετικούς συνδυασμούς πυραύλων εκτόξευσης. Τονίζεται ότι για να αποφανθούμε για την καταλληλότητα του μοντέλου είναι απαραίτητη η ανάλυση υπολοίπων. [9]

2-2-7. Μικρά λατινικά τετράγωνα

Μικρά λατινικά τετράγωνα ονομάζονται εκείνα που έχουν μικρό αριθμό γραμμών και στηλών. Για παράδειγμα, ένα 3X3 ή 4X4 λατινικό τετράγωνο θεωρείται μικρό. Ένα μειονέκτημα των μικρών λατινικών τετραγώνων είναι ότι δίνουν ένα σχετικά μικρό αριθμό βαθμών ελευθερίας των σφαλμάτων. Για παράδειγμα, ένα 3X3 λατινικό τετράγωνο έχει δύο βαθμούς ελευθερίας, ένα 4X4 λατινικό τετράγωνο έχει έξι βαθμούς ελευθερίας κλπ. Γενικά, όταν γίνεται χρήση μικρών λατινικών τετραγώνων πραγματοποιούμε η επανάληψεις αυτών προκειμένου ν' αυξηθούν οι βαθμοί ελευθερίας των σφαλμάτων. Υπάρχουν αρκετές μέθοδοι επανάληψης ενός μικρού λατινικού τετραγώνου. Αξίζει να σημειωθεί ότι η μέθοδος επανάληψης που επιλέγεται επηρεάζει τη μεταβλητότητα του λατινικού τετραγώνου.

Το υπόδειγμα των λατινικών τετραγώνων έχει εφαρμογή και σε πειράματα επαναλαμβανόμενων μετρήσεων όπου υπάρχουν p παρεμβάσεις σε p περιόδους. Ένα τέτοιο παράδειγμα αποτελεί και μια 3X3 crossover δοκιμή που παρουσιάζεται σε επόμενο κεφάλαιο. Μια 3X3 crossover δοκιμή αποτελεί ορθογώνιο λατινικό τετράγωνο. [9]

2-3. Στατιστικά λογισμικά

2-3-1. Γενικά

Η ανάπτυξη της τεχνολογίας και της πληροφορικής είχε – μεταξύ άλλων – και ως αποτέλεσμα τη δημιουργία στατιστικών προγραμμάτων με τα οποία δόθηκε η δυνατότητα στους - ανά τον κόσμο - στατιστικούς και ερευνητές να διαχειρίζονται χιλιάδες δεδομένα και να πραγματοποιούν πολύπλοκους μαθηματικούς υπολογισμούς μέσα σε λίγα δευτερόλεπτα. Στη σημερινή εποχή

κυκλοφορούν αρκετές δεκάδες στατιστικά λογισμικά τέσσερα (4) από τα οποία παρουσιάζονται εν συντομία παρακάτω. Με τα συγκεκριμένα στατιστικά προγράμματα πραγματοποιήθηκε στο κεφάλαιο 4 η ανάλυση δεδομένων του test γευσιγνωσίας σοκολάτας.

2-3-2. JMP

Το JMP (προφέρεται JUMP) είναι ένα στατιστικό λογισμικό που αναπτύχθηκε από την JMP Business η οποία αποτελεί εταιρεία του Ινστιτούτου SAS. Δημιουργήθηκε τη δεκαετία του '80 χρησιμοποιώντας το περιβάλλον γραφικών που εισήγαγε νωρίτερα η Apple Macintosh. Έκτοτε έχει βελτιωθεί και είναι διαθέσιμο και από το λειτουργικό των Windows. Το JMP χρησιμοποιείται εκτενώς σε εφαρμογές όπως ο στατιστικός ποιοτικός έλεγχος, η μηχανική, ο σχεδιασμός πειραμάτων και εν γένει η επιστημονική έρευνα. Αξίζει να σημειωθεί ότι το συγκεκριμένο λογισμικό διατίθεται σήμερα σε πέντε (5) προϊόντα, ήτοι JMP, JMP Pro, JMP Clinical, JMP Genomics και JMP Graph Builder App (for iPad).

Όπως αναφέρθηκε το JMP αναπτύχθηκε τη δεκαετία του '80 από τον John Sall και μια ομάδα προγραμματιστών με την αξιοποίηση του περιβάλλοντος γραφικών που πριν λίγο καιρό είχε παρουσιάσει η Apple Macintosh. Το 1991 παρουσιάστηκε η έκδοση JMP 2.0 στην οποία προστέθηκαν διαδραστικά γραφικά και άλλες λειτουργίες. Το 1999 κυκλοφόρησε η έκδοση JMP 3.0 και τρία (03) χρόνια μετά ή έκδοση JMP 4.0. Στην έκδοση αυτή υπήρχε η δυνατότητα εισαγωγής δεδομένων από μια μεγάλη γκάμα πηγών δεδομένων. Επιπλέον προστέθηκαν νέες λειτουργίες που αφορούσαν στην ανάλυση χρονολογικών σειρών, τεχνικές εξομάλυνσης κλπ. Με την έκδοση JMP 5.0 του 2005 προστέθηκαν λειτουργίες σχετικά με την εξόρυξη δεδομένων και τα νευρωνικά δίκτυα. Αργότερα την ίδια χρονιά παρουσιάστηκε η έκδοση JMP 6.0 και το 2007 εμφανίστηκε η έκδοση JMP 7.0. Το 2009 παρουσιάστηκε η έκδοση JMP 8.0 στα 64 bit με ένα νέο περιβάλλον γραφικών και το 2010 εκδόθηκε το JMP 9.0 που έδινε τη δυνατότητα χρησιμοποίησης και της γλώσσας προγραμματισμού R. Το 2012 κυκλοφόρησε η έκδοση JMP 10.0 που είχε σημαντικές βελτιώσεις σε ζητήματα εξόρυξης δεδομένων, μοντελοποίησης και μεθόδων προβλέψεων.

Το λογισμικό JMP εστιάζει στη διερευνητική ανάλυση δεδομένων και στην οπτικοποίηση αυτών. Δίνει ιδιαίτερο βάρος στη γραφική αναπαράσταση των δεδομένων παρέχοντας μεγάλο αριθμό και ποικιλία γραφημάτων σχεδόν για κάθε στατιστική εργασία. Αποτελεί κοινή παραδοχή ότι οι προγραμματιστές του συγκεκριμένου λογισμικού συνεχίζουν να εργάζονται και να προσθέτουν σε κάθε νέα έκδοση καινοτόμες και μοντέρνες στατιστικές μεθόδους καθώς και βελτιωμένες μεθόδους γραφικής αναπαράστασης των δεδομένων.

Το λογισμικό JMP αποτελεί πανίσχυρο εργαλείο για στατιστικές αναλύσεις κάθε μορφής και παράγει γραφήματα άριστης ποιότητας. Ταυτόχρονα η «παραθυρική» προσέγγιση - μέσω των μενού – των διαφόρων λειτουργιών του το καθιστά εξαιρετικά φιλικό για το χρήστη. [28],[29]

2-3-3. MINITAB

Το Minitab αποτελεί στατιστικό πακέτο το οποίο αναπτύχθηκε στο Πανεπιστήμιο της Πενσυλβάνια το έτος 1972 από τους ερευνητές (προγραμματιστές και στατιστικούς) Barbara F. Ryan, Thomas A. Ryan Jr. και Brian L. Joiner. Το Minitab παρουσιάστηκε ως μια «ελαφριά» έκδοση του προγράμματος OMNITAB που δημιουργήθηκε από την εταιρεία ανάπτυξης λογισμικού NIST. Το λογισμικό Minitab διατίθεται από την εταιρεία Minitab Inc.

Παρά το γεγονός ότι μέχρι σήμερα έχουν κυκλοφορήσει αρκετές εκδόσεις του συγκεκριμένου λογισμικού δεν σημειώθηκε σημαντική βελτίωση σε σχέση με την αρχική έκδοση. Το Minitab χρησιμοποιείται κυρίως στον στατιστικό ποιοτικό έλεγχο καθώς και σε άλλες στατιστικές εφαρμογές. Είναι φιλικό στο χρήστη δεδομένου ότι όλες οι διαδικασίες πραγματοποιούνται «παραθυρικά» μέσω των μενού. Τα περιβάλλον στο οποίο δίνονται τα αποτελέσματα (output) υστερεί από άποψη παρουσίασης σε σχέση με άλλα στατιστικά πακέτα.

Η τελευταία έκδοση του Minitab είναι η 17 και διατίθεται σε οκτώ (08) γλώσσες αγγλικά, γαλλικά, γερμανικά, ιαπωνικά, κορεάτικα, πορτογαλικά, κινέζικα και ισπανικά. Εξαιτίας του σχετικά χαμηλού του κόστους χρησιμοποιείται από αρκετά εκπαιδευτικά ιδρύματα και φοιτητές. [27]

2-3-4. R

Η γλώσσα προγραμματισμού S που δημιουργήθηκε από τα Bell Laboratories και συγκεκριμένα από τον John Chambers και την ομάδα του αποτελεί τη βάση

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

της R. Η R αναπτύχθηκε ως γλώσσα προγραμματισμού για ανάλυση στατιστικών δεδομένων και δημιουργίας γραφημάτων από τους Ross Ihaka και Robert Gentleman. Οι Ihaka και Ross είναι καθηγητές στο Τμήμα Στατιστικής του Πανεπιστημίου του Auckland στη Νέα Ζηλανδία.

Η R αποτελεί μια εφαρμογή ανοικτού κώδικα (open source) ή αλλιώς μια γλώσσα προγραμματισμού με ενσωματωμένο περιβάλλον, που περιέχει πλήθος συναρτήσεων για στατιστικούς ελέγχους και δίνει τη δυνατότητα δημιουργίας ποικίλων γραφημάτων.

Η R παρουσιάζει πολλά πλεονεκτήματα τα οποία συνοψίζονται στα ακόλουθα:

α) είναι μια εφαρμογή ανοικτού κώδικα πράγμα που σημαίνει ότι δίνει τη δυνατότητα σε οποιονδήποτε να δημιουργήσει νέα πακέτα (packages) γράφοντας κώδικα στον οποίο μπορεί να έχει πρόσβαση κάθε ενδιαφερόμενος.

β) λειτουργεί στα συνήθη λειτουργικά συστήματα όπως Windows, Linux κλπ

γ) διατίθεται δωρεάν (καθένας μπορεί να το «κατεβάσει» από τη διεύθυνση <http://www.r-project.org>)

δ) προσαρμόζεται στις ανάγκες κάθε ερευνητή

ε) υπάρχει πλήθος εγχειριδίων που έχουν γραφεί για την R, στην ιστοσελίδα της υπάρχουν πολλές πληροφορίες γύρω από αυτή και υπάρχουν ομάδες χρηστών που παρέχουν κάθε πληροφορία για την αντιμετώπιση προβλημάτων γύρω από την R.

Ένα μειονέκτημα της R είναι ότι απαιτεί χρήση εντολών (κώδικα) γι' αυτό και προτιμάται από έμπειρους χρήστες οι οποίοι έχουν μια σχετικά ικανοποιητική εξοικείωση με τον προγραμματισμό. Αντίθετα χρήστες με μικρή εμπειρία ή πολύ περισσότερο αρχάριοι αποφεύγουν να ασχοληθούν με την R. Γενικά η R θεωρείται ένα εξαιρετικό λογισμικό για την επεξεργασία και ανάλυση στατιστικών δεδομένων. [18]

2-3-5. SPSS

Το στατιστικό πακέτο SPSS (Statistical Package for social sciences) αποτελεί ένα λογισμικό που χρησιμοποιείται ευρέως για στατιστική ανάλυση δεδομένων κυρίως στις κοινωνικές επιστήμες. Εκτεταμένη χρήση του γίνεται ακόμη από εταιρείες έρευνας αγοράς, από κυβερνητικούς οργανισμούς, πανεπιστήμια,

ερευνητικά κέντρα, μεμονωμένους ερευνητές (στον τομέα της εκπαίδευσης, της υγείας κλπ) και φοιτητές.

Το συγκεκριμένο λογισμικό παρουσιάστηκε για 1^η φορά το 1968 (1^η έκδοση) ως στατιστικό πρόγραμμα για χρήση από κοινωνικές επιστήμες και αναπτύχθηκε από τους Norman H. Nie, Dale H. Bent και C.Hadlai Hull. Έκτοτε και ανά τακτά χρονικά διαστήματα παρουσιάζονται και νέες εκδόσεις του προγράμματος. Η έκδοση SPSS 16.0 και οι μετέπειτα εκδόσεις από αυτή «έτρεχαν» σε περιβάλλον Windows, Mac και Linux με τον κώδικα για το περιβάλλον γραφικών να είναι γραμμένος σε Java. Οι πρότερες εκδόσεις του SPSS 16.0 «έτρεχαν» σε περιβάλλον Windows, Mac OS X και Unix. Τον Ιούλιο του 2009 η εταιρεία SPSS Inc. ανακοίνωσε την εξαγορά της από την εταιρεία IBM για 1,2 δισεκατομμύρια δολάρια. Εξαιτίας μιας διαμάχης που προέκυψε για το όνομα του λογισμικού - μεταξύ των ετών 2009 και 2010 - το πρόγραμμα προσωρινά πήρε το όνομα PASW (Predictive Analytics Software). Με την ολοκλήρωση της εξαγοράς της SPSS Inc. από την IBM το 2010 το λογισμικό πήρε το όνομα IBM SPSS Statistics. Από τότε και κάθε χρόνο μέχρι σήμερα παρουσιάζονται νέες εκδόσεις από την εταιρεία.

Το SPSS παρέχει πλήθος δυνατοτήτων οι οποίες είναι προσβάσιμες είτε μέσω μενού με καταρράκτες (pull down menus) είτε γράφοντας κώδικα με εντολές (command syntax language). Η «παραθυρική» προσέγγιση είναι ιδιαίτερη εύκολη και εξαιρετικά φιλική για το χρήστη. Ωστόσο παρουσιάζει ορισμένα προβλήματα όπως η αδυναμία χρήσης ορισμένων εξειδικευμένων στατιστικών διαδικασιών με τη βοήθεια των μενού.

Αντίθετα, η χρήση κώδικα (σύνταξη εντολών) έχει τα οφέλη της επαναληψιμότητας, της διαχείρισης πολύπλοκων δεδομένων και της απλούστευσης επαναλαμβανόμενων διαδικασιών. Επιπρόσθετα, κάποιες πολύπλοκες στατιστικές διαδικασίες μπορούν να πραγματοποιηθούν μόνο με τη βοήθεια εντολών. Αξίζει να σημειωθεί ότι και η «παραθυρική» εργασία παράγει εντολές οι οποίες όμως δεν είναι άμεσα ορατές από το χρήστη. Για να γίνει αυτό θα πρέπει να πραγματοποιηθούν ορισμένες αλλαγές στις default ρυθμίσεις.

Η δημιουργία και η διαχείριση των αρχείων δεδομένων είναι απλή όπως και ο ορισμός των τύπων των δεδομένων και των μεταβλητών. Παράλληλα είναι εύκολη η προσπέλαση διαφορετικών – από το SPSS – αρχείων που προέρχονται από άλλες εφαρμογές ή διαφορετικά στατιστικά πακέτα. [26]

Το SPSS αποτελεί σήμερα ίσως το πιο γνωστό και το περισσότερο χρησιμοποιούμενο στατιστικό λογισμικό παγκοσμίως.

Κεφάλαιο 3

2X2 ή AB/BA και 3X3 Crossover δοκιμές

(Πραγματικός διάλογος κάπου στα τέλη Αυγούστου 2015 στη ραδιοφωνική εκπομπή «Της κακομοίρας» του ράδιο Μετρόπολις 95,5 fm ...

-Ακροατής: Ραπ, μήπως ξέρεις να μου πεις ποιο είναι το μυστικό της μακροζωίας;

-Κωστής Ραπτόπουλος: Η αναισθησία !!!)

3-1. Εισαγωγή

Η 2X2 ή AB/BA δοκιμή αποτελεί την πιο απλή μελέτη διασταυρούμενης μετάβασης (2X2 ή AB/BA crossover trial). Μία 2X2 crossover δοκιμή μελετά δύο παρεμβάσεις – στην πλειοψηφία τους θεραπευτικά σχήματα – Α και Β. Μία πρώτη ομάδα ασθενών (ομάδα I) λαμβάνει τα θεραπευτικά σχήματα με σειρά (αλληλουχία) AB και μια δεύτερη ομάδα ασθενών (ομάδα II) λαμβάνει τα θεραπευτικά σχήματα με σειρά (αλληλουχία) BA. Τονίζεται ότι αμφότερες οι ομάδες ασθενών είναι τυχαία επιλεγμένες. Στην πραγματικότητα η παραπάνω δοκιμή είναι μια 2X2X2 δοκιμή επειδή έχουμε δύο παρεμβάσεις (θεραπείες) σε δύο ομάδες ασθενών σε δύο χρονικές περιόδους. [5],[7]

Όπως αναφέρθηκε και στο 1^ο κεφάλαιο μεταξύ των δύο περιόδων μελέτης μεσολαβεί ένα εύλογο χρονικό διάστημα το οποίο καθορίζεται από τον ερευνητή και καλείται περίοδος έκπλυσης (washout period). Η ενδιάμεση αυτή περίοδος κατά την οποία δεν εφαρμόζεται κανενός είδους παρέμβαση σε οποιονδήποτε ασθενή αποτελεί ζωτικής σημασίας κομμάτι για την επιτυχία ενός crossover πειράματος. Η περίοδος έκπλυσης επιτρέπει στον οργανισμό του ασθενή να επιστρέψει στην προϋπάρχουσα της αρχικής θεραπείας κατάσταση και τον καθιστά έτοιμο να δεχθεί τη θεραπευτική παρέμβαση της δεύτερης περιόδου. Σε μια 2Χ2 crossover δοκιμή η περίοδος έκπλυσης είναι μία και μοναδική. [5]

Ο παρακάτω πίνακας μας δίνει μια crossover δοκιμή 2 παρεμβάσεων και 2 περιόδων.

Πίνακας 3-1

2Χ2 ή ΑΒ/ΒΑ crossover δοκιμή

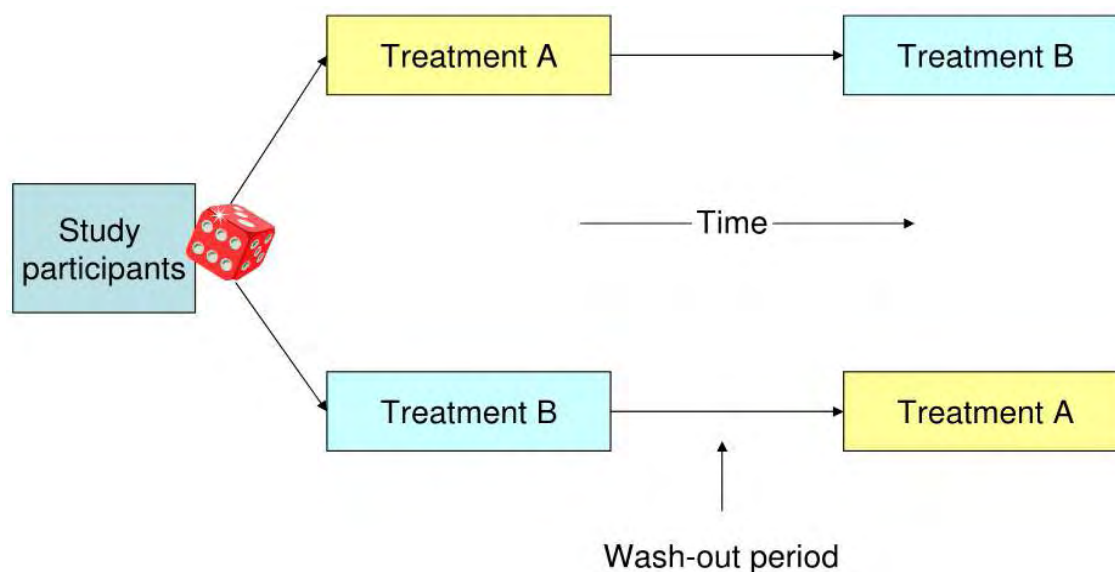
Ομάδες ασθενών	Περίοδος Προετοιμασίας	1 ^η Περίοδος	Περίοδος έκπλυσης	2 ^η Περίοδος
1 ^η Ομάδα	-	Α θεραπεία	-	Β θεραπεία
2 ^η Ομάδα	-	Β θεραπεία	-	Α θεραπεία

30

Σχηματικά μια 2Χ2 crossover δοκιμή μπορεί να αποτυπωθεί ως ακολούθως:

Σχήμα 3-1

2Χ2 ή ΑΒ/ΒΑ crossover δοκιμή



ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Η ύπαρξη του ζαριού στο παραπάνω σχήμα υποδηλώνει την τυχαιοποίηση (randomization) των συμμετεχόντων στις δύο ομάδες και αποδεικνύει ότι οι crossover δοκιμές αποτελούν σημαντική κατηγορία τυχαιοποιημένων κλινικών μελετών.

Σε μια 2X2 crossover δοκιμή υφίστανται οι έννοιες της θεραπευτικής παρέμβασης (π.χ φάρμακο Α και φάρμακο Β), της περιόδου (1^η και 2^η) και της αλληλουχίας (sequence AB και BA). Όλες οι προηγούμενες αποτελούν επεξηγηματικές μεταβλητές και είναι δίτιμες κατηγορικές. Η μεταβλητή ενδιαφέροντος ή εξαρτημένη μεταβλητή (απόκριση) μπορεί να είναι είτε συνεχής (continuous) είτε δίτιμη κατηγορική (binary). [7] **Στην παρούσα εργασία γίνεται παρουσίαση crossover δοκιμών όπου η απόκριση είναι συνεχής και οι τιμές της επιθυμούμε να προέρχονται από κανονική κατανομή.** Για παράδειγμα, μια συνεχής εξαρτημένη μεταβλητή μπορεί να είναι οι τιμές της συστολικής αρτηριακής πίεσης ή οι τιμές της ενδοφθάλμιας πίεσης κλπ.

3-2. Κέρδη και Προβλήματα

Σε μια 2X2 crossover δοκιμή αντικείμενο μελέτης αποτελεί η εξαρτημένη μεταβλητή (απόκριση) που για την παρούσα εργασία λαμβάνεται πάντα συνεχής. Συγκεκριμένα, διερευνάται αν υφίστανται στατιστικά σημαντικές διαφορές στις μετρήσεις της απόκρισης μετά την εφαρμογή δύο θεραπειών, σε δύο ομάδες πειραματικών μονάδων (ασθενών) σε δύο διαφορετικές περιόδους με τη χρήση δύο αλληλουχιών.

Η 2X2 crossover δοκιμή έχει το σημαντικό πλεονέκτημα ότι κάθε πειραματική μονάδα (ασθενής) λειτουργεί ως control του εαυτού του. Αυτό έχει ως αποτέλεσμα να επιτυγχάνεται μεγαλύτερη ακρίβεια στην εκτίμηση της επίδρασης της θεραπείας (εξαιτίας της μείωσης της – μεταξύ των ασθενών – μεταβλητότητας) καθώς και στη χρήση μικρότερων δειγμάτων ασθενών για την εφαρμογή του πειράματος. [7]

Από την άλλη, η χρήση crossover δοκιμών δημιουργεί και δυσχέρειες. Οι αποχωρήσεις ασθενών (drop outs) είναι μία από αυτές. Σ' αυτή την περίπτωση – ακόμα και αν ένας ασθενής έχει ολοκληρώσει μέρος του

πειράματος (π.χ λήψη A θεραπείας) – δεν είναι βέβαιο ότι οι μετρήσεις του θα ληφθούν υπόψιν κατά τη στατιστική επεξεργασία. [5],[7]

Το πλέον σύνηθες πρόβλημα που δημιουργείται σε τέτοιου είδους δοκιμές είναι η κατάσταση κατά την οποία η επίδραση της θεραπευτικής παρέμβασης της πρώτης περιόδου εξακολουθεί να υφίσταται και κατά τη δεύτερη περίοδο. Αυτό συνήθως προκαλείται όταν η περίοδος έκπλυσης (washout period) είναι π.χ αρκετά σύντομη. Το παραπάνω έχει ως αποτέλεσμα την εμφάνιση στατιστικά σημαντικής αλληλεπίδρασης θεραπείας και περιόδου (treatment by period interaction) η οποία συνεπάγεται την ύπαρξη επίδρασης ακολουθίας (επίδραση carry-over). [7]

Το φαινόμενο carry-over δεν αποτελεί επιθυμητή κατάσταση σ' ένα 2X2 crossover πείραμα (και γενικότερα σε crossover δοκιμές) του οποίου ο σχεδιασμός θα πρέπει να είναι τέτοιος - ώστε να αποφεύγεται η εμφάνισή του. Γενικότερα, η ύπαρξη επίδρασης carry-over προκαλεί προβλήματα και κάνει δυσχερή τη στατιστική ανάλυση των δεδομένων. [7]

Αξίζει να σημειωθεί ότι στις περιπτώσεις των 2X2 cross-over δοκιμών (και μόνο σ' αυτές) η ύπαρξη αλληλεπίδρασης θεραπείας και περιόδου ταυτίζεται με την επίδραση carry-over. Σε περισσότερο πολύπλοκα πειράματα η επίδραση carry-over και η αλληλεπίδραση θεραπείας και περιόδου πρέπει να μελετώνται χωριστά. [5],[7]

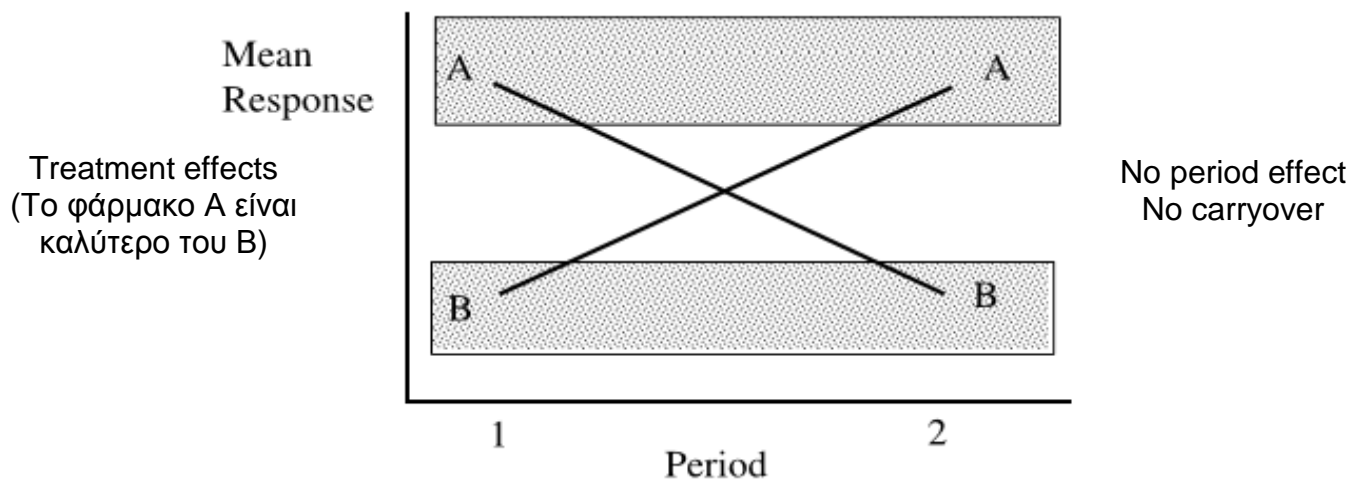
3-3. Είδη επιδράσεων

3-3-1. Επίδραση θεραπείας (treatment effects)

Η επίδραση θεραπείας (treatment effects) αναφέρεται στην επίδραση που έχει η παρέμβαση (π.χ το θεραπευτικό σκεύασμα) στις τιμές ενός χαρακτηριστικού των πειραματικών μονάδων (ασθενών) όταν αυτή εφαρμόζεται σε μια συγκεκριμένη χρονική περίοδο. Το χαρακτηριστικό αυτό αποτελεί στην ουσία την εξαρτημένη μεταβλητή (απόκριση). Ακολουθεί ένα γράφημα όπου σε μια 2X2 crossover δοκιμή υφίσταται επίδραση θεραπείας (treatment effects) χωρίς να υπάρχει οποιαδήποτε άλλου είδους επίδραση. Θεωρούμε ότι υψηλές τιμές στην εξαρτημένη μεταβλητή (απόκριση) είναι επιθυμητές (υποδεικνύουν θετικό αποτέλεσμα). [7]

Γράφημα 3-1

2X2 crossover δοκιμή με επίδραση θεραπείας

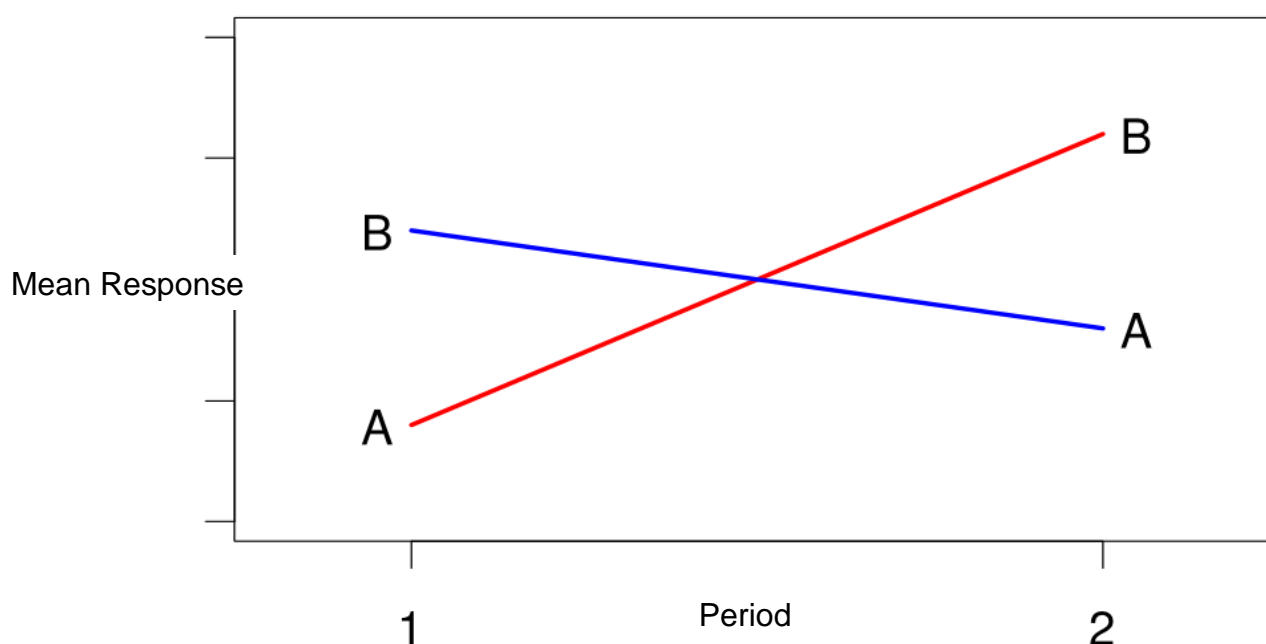
**3-3-2. Επίδραση περιόδου (period effects)**

Επίδραση περιόδου (period effects) μπορεί να έχουμε όταν οι μετρήσεις της μεταβλητής απόκρισης στην πρώτη περίοδο εμφανίζονται σημαντικά μικρότερες ή μεγαλύτερες σε σχέση με εκείνες της δεύτερης περιόδου. Ακολούθως δίνεται ένα γράφημα όπου έχουμε επίδραση περιόδου. [5]

33

Γράφημα 3-2

2X2 crossover δοκιμή με επίδραση περιόδου

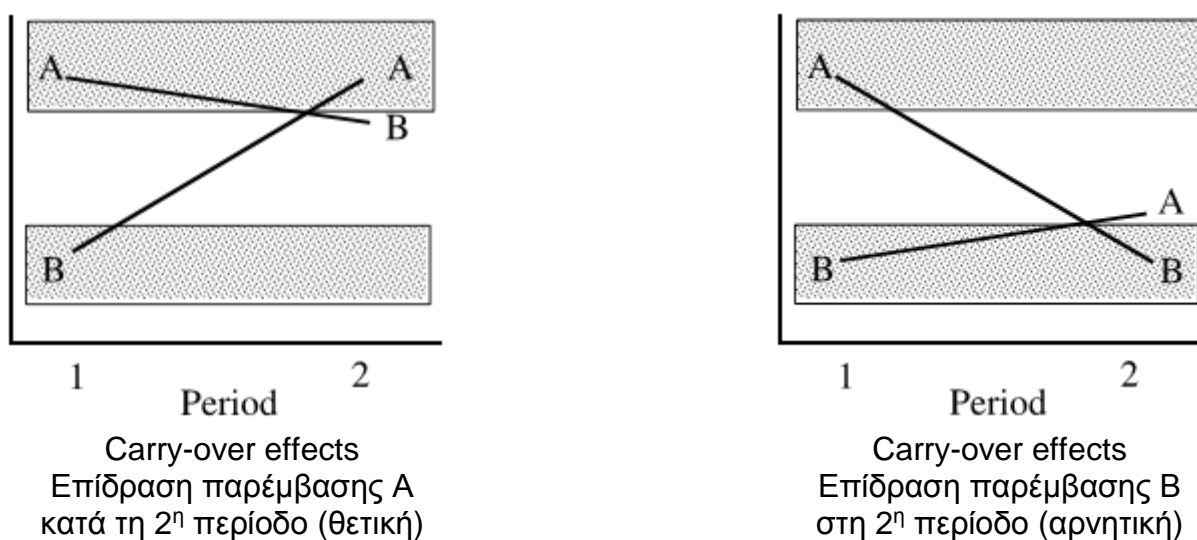


3-3-3. Επίδραση ακολουθίας (carry-over effects)

Όπως αναφέρθηκε - σε μια 2X2 crossover δοκιμή όταν η επίδραση της θεραπευτικής παρέμβασης της πρώτης περιόδου εξακολουθεί να υφίσταται και κατά τη δεύτερη περίοδο τότε έχουμε επίδραση carry-over. Ακολουθεί γράφημα με δύο περιπτώσεις στις οποίες υφίσταται επίδραση carry-over. [7]

Γράφημα 3-3

2X2 crossover δοκιμές με επίδραση carry-over



34

Σημειώνεται ότι στις 2X2 crossover δοκιμές η επίδραση carry-over ταυτίζεται με την ύπαρξη αλληλεπίδρασης θεραπείας – περιόδου (δεν ελέγχονται χωριστά).

3-4. Γραφήματα – Στατιστική ανάλυση

3-4-1. Χρησιμότητα γραφημάτων

Οι περισσότερες στατιστικές μέθοδοι στηρίζονται σε υποθέσεις που αφορούν στη συμπεριφορά των δεδομένων. Σε πολλές περιπτώσεις οι υποθέσεις αυτές μπορούν να είναι λανθασμένες και ως τούτου οι αριθμητικοί υπολογισμοί να είναι παραπλανητικοί. Σε κάθε περίπτωση είναι απαραίτητο να γίνεται διερεύνηση, του κατά πόσο οι υποθέσεις που διατυπώνονται είναι λογικά ορθές και σε περίπτωση που είναι λανθασμένες να επιδιώκεται η κατανόηση των αιτιών αυτών των λαθών. Τα γραφήματα μπορούν να αποδειχθούν εξαιρετικά πολύτιμα για αυτό το σκοπό.

Επομένως, τόσο οι στατιστικές - μαθηματικές μέθοδοι όσο και τα γραφήματα θα πρέπει να χρησιμοποιούνται από κοινού για την κατανόηση ενός συνόλου δεδομένων. Ειδικότερα, τα γραφήματα μπορούν:

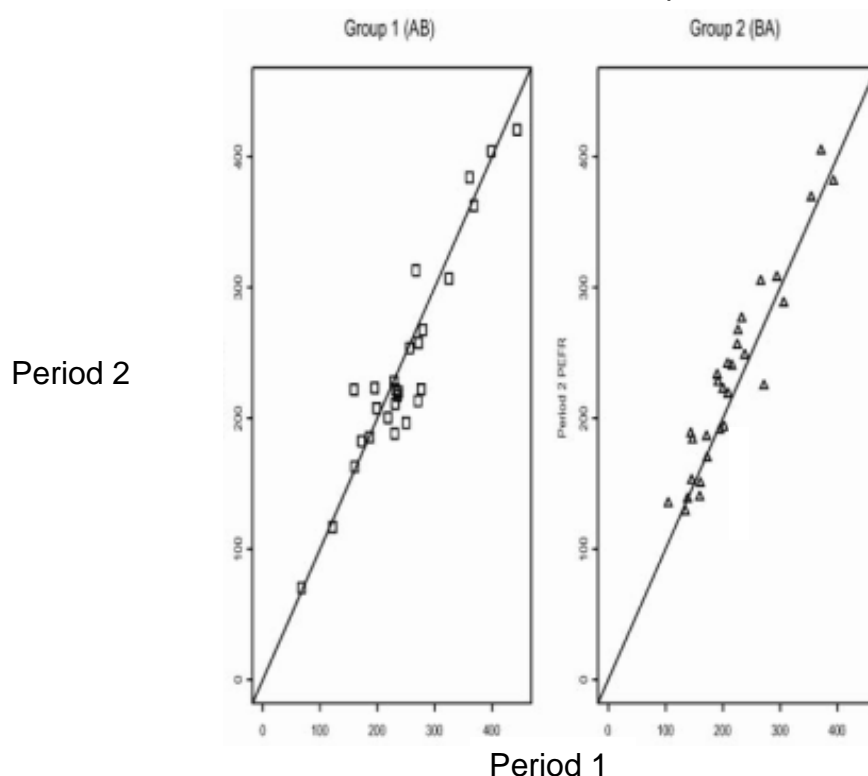
- να βοηθήσουν στην αναγνώριση και εκτίμηση αδρών χαρακτηριστικών - γνωρισμάτων των δεδομένων
- να επιτρέψουν τη διερεύνηση σ' ένα δεύτερο – πιο αναλυτικό – επίπεδο αυτών των αδρών χαρακτηριστικών

3-4-2. Γραφήματα και 2X2 crossover δοκιμές

Με το 1^ο είδος γραφημάτων έχουμε τη δυνατότητα να διερευνήσουμε την ύπαρξη επίδρασης περιόδου (period effects). Το γράφημα αποτελείται από δύο μέρη όπου στο μεν πρώτο αποτυπώνονται οι μετρήσεις (τιμές) της εξαρτημένης μεταβλητής (απόκρισης) ως προς την 1^η και 2^η περίοδο για κάθε ασθενή της 1^{ης} ομάδας και στο δε δεύτερο μέρος οι μετρήσεις (τιμές) της εξαρτημένης μεταβλητής (απόκρισης) ως προς την 1^η και 2^η περίοδο για κάθε ασθενή της 2^{ης} ομάδας. Σε καθένα από τα παραπάνω γραφήματα μπορούμε να προσθέσουμε - για καλύτερη εποπτεία - μια διαγώνιο γραμμή η οποία έχει κλίση 1 και διέρχεται από την αρχή των αξόνων. [7]

Γράφημα 3-4

Period 2 vs Period 1 plot

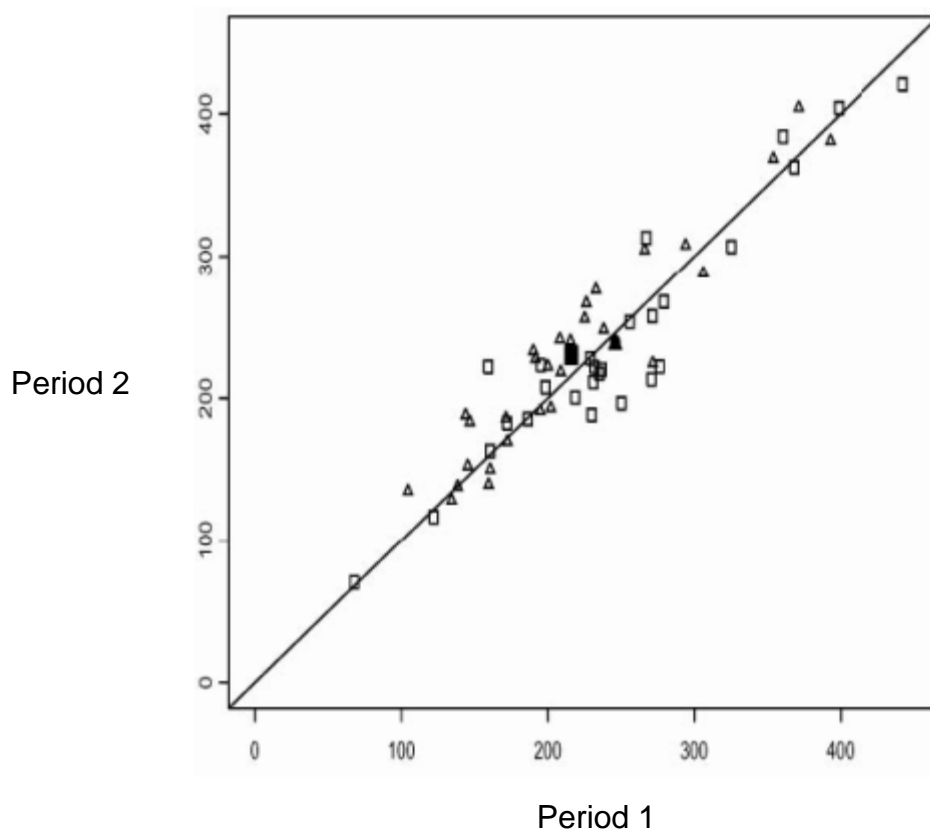


Αν τα σημεία αποτυπώνονται εκατέρωθεν της διαγωνίου, κοντά και σχεδόν συμμετρικά σ' αυτή τότε υπάρχει ένδειξη ότι δεν υφίσταται επίδραση περιόδου (period effects). Τούτο συμβαίνει επειδή η μεταξύ των ασθενών μεταβλητότητα είναι μικρή. Το αντίθετο ισχύει αν τα σημεία απομακρύνονται από τη διαγώνιο και κατανέμονται άτακτα. [7]

Με το ακόλουθο γράφημα μπορούμε να εντοπίσουμε εποπτικά την ύπαρξη επίδρασης θεραπείας (treatment effects). Κατασκευάζουμε ένα γράφημα στο οποίο αποτυπώνουμε από κοινού τις μετρήσεις (τιμές) της εξαρτημένης μεταβλητής (απόκρισης) ως προς την 1^η και 2^η περίοδο για κάθε ασθενή και για τις 2 ομάδες. Ουσιαστικά το συγκεκριμένο γράφημα αποτελεί συγχώνευση των δύο μερών του γραφήματος που παρουσιάστηκε προηγουμένως. Στο γράφημα αυτό σχεδιάζονται δύο κεντρικές τιμές (centroids) όπου η μια αντιπροσωπεύει το νέφος σημείων της 1^{ης} ομάδας και η άλλη το νέφος σημείων της 2^{ης} ομάδας. Σε περίπτωση που οι κεντρικές τιμές βρίσκονται εκατέρωθεν της διαγωνίου και σε απόσταση μεταξύ τους έχουμε ένδειξη για την ύπαρξη επίδρασης θεραπείας (treatment effects). [7]

Γράφημα 3-5

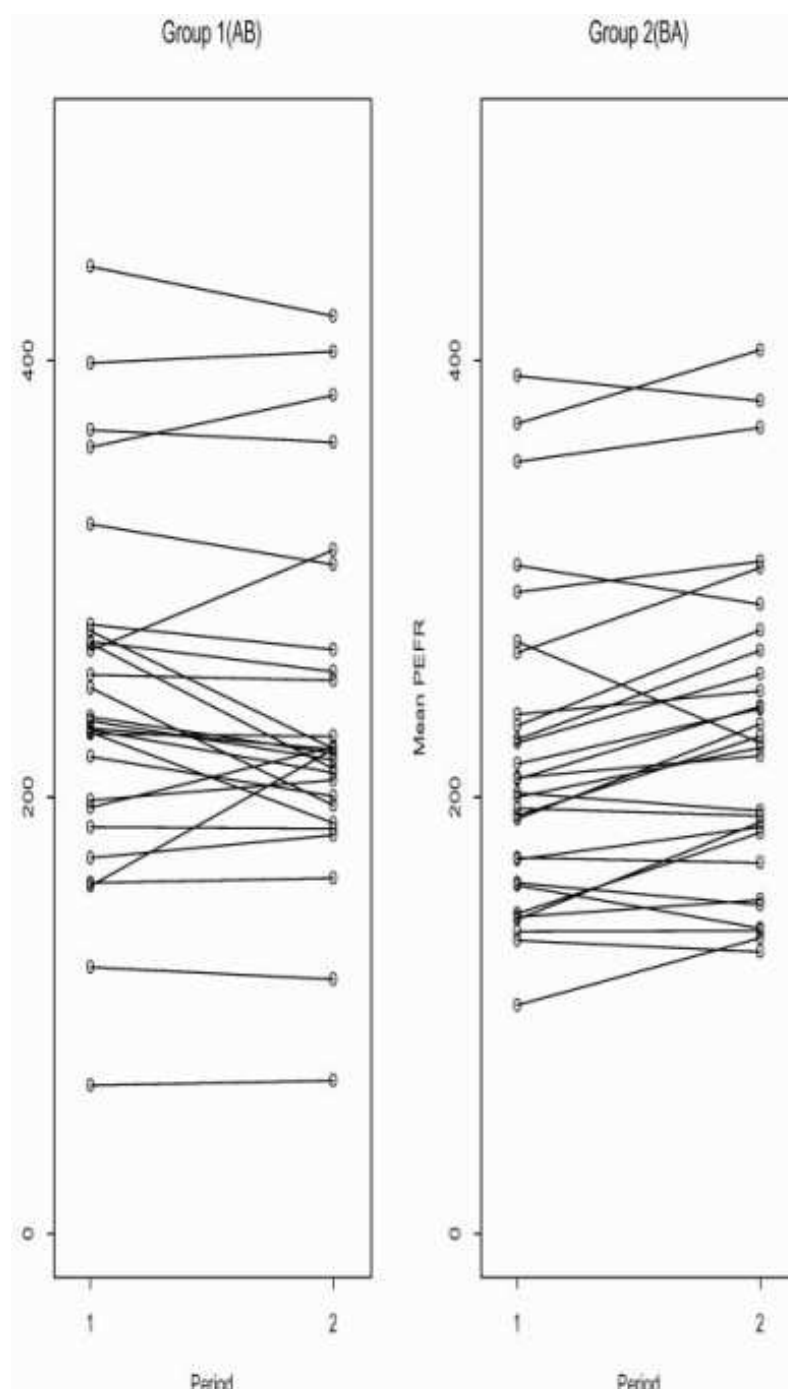
Period 2 vs Period 1 plot with centroids



Τα γραφήματα προφίλ (profiles plots) αποτελούν γραφικές απεικονίσεις των διαφορών στις τιμές της εξαρτημένης μεταβλητής (απόκρισης) για κάθε ασθενή χωριστά μεταξύ 1^{ης} και 2^{ης} περιόδου. Τα γραφήματα προφίλ δίνονται σε δύο μέρη – ένα για κάθε ομάδα. Με τον τρόπο αυτό μπορούμε να εντοπίσουμε πολύ μεγάλες διαφορές μεταξύ 1^{ης} και 2^{ης} μέτρησης για κάθε ασθενή. [7]

Γράφημα 3-6

Profiles plots for outcome data



Ένα ακόμη γράφημα που χρησιμοποιείται συχνά είναι εκείνο με το οποίο συγκρίνονται οι μέσες τιμές της εξαρτημένης μεταβλητής (απόκρισης) ως προς την κάθε ομάδα και περίοδο. Πρόκειται στην ουσία για ένα γράφημα μέσων τιμών (means plot) με τη βοήθεια του οποίου εντοπίζονται άμεσα οι διαφορές μεταξύ των μέσων τιμών στις ομάδες και ανά περίοδο. Τα γραφήματα 2-1, 2-2 και 2-3 ανήκουν σ' αυτή την κατηγορία γραφημάτων που περιγράφηκε προηγουμένως. [7]

Η επίδραση ακολουθίας (carry-over effects) δεν εντοπίζεται – σαφώς - με τη βοήθεια κάποιου γραφήματος. Το means plot μπορεί να μας δώσει μια ένδειξη για την ύπαρξή της (ή μη). [7]

3-4-3. Στατιστική ανάλυση

Έχει ήδη αναφερθεί ότι στις 2X2 crossover δοκιμές που μελετάμε - η εξαρτημένη μεταβλητή (απόκριση) θεωρείται συνεχής με τιμές που προέρχονται από κανονική κατανομή. Σε ένα 2X2 crossover πείραμα είναι κομβικής σημασίας να εξεταστεί αν υφίσταται στατιστικά σημαντική επίδραση carryover (αλλιώς αλληλεπίδραση θεραπείας – περιόδου). Η ύπαρξη (ή μη) σημαντικής επίδρασης carry-over επηρεάζει τη στατιστική προσέγγιση που θα επιλεγεί.

Αν - για παράδειγμα - η επίδραση carry-over είναι στατιστικά σημαντική τότε αναλύουμε μόνο τα αποτελέσματα της 1^{ης} περιόδου και συγκρίνουμε τις 2 ομάδες είτε χρησιμοποιώντας t-test (αν τα δεδομένα ακολουθούν κανονική κατανομή) είτε χρησιμοποιώντας το αντίστοιχο μη παραμετρικό Mann-Whitney test ή Wilcoxon rank sum test (αν τα δεδομένα δεν ακολουθούν κανονική κατανομή). Στην ουσία η 2X2 crossover δοκιμή μετατρέπεται σε μια μελέτη παράλληλων ομάδων (parallel group study). [5],[7]

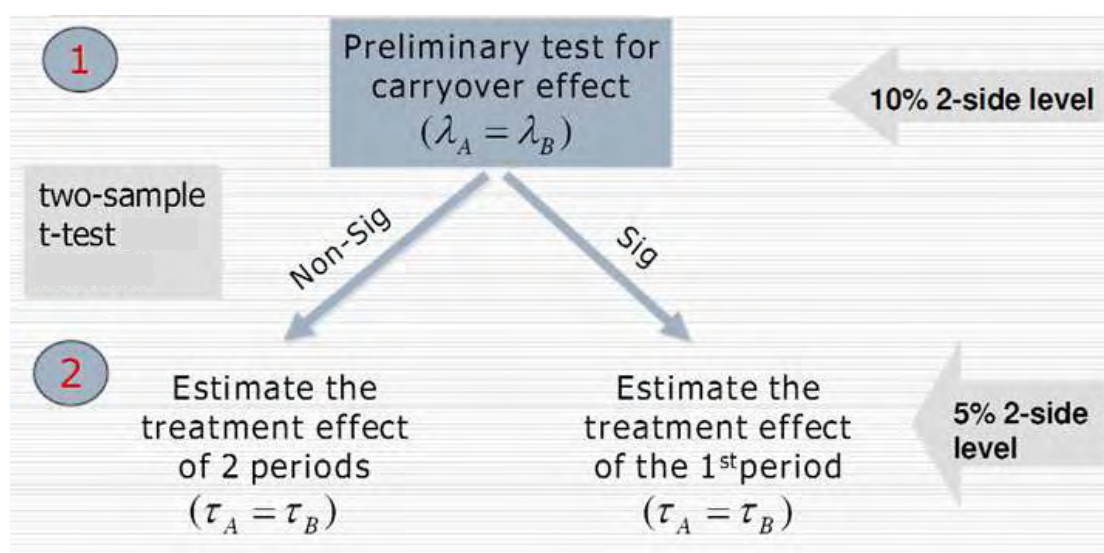
Στην περίπτωση που η επίδραση carry-over δεν είναι στατιστικά σημαντική τότε πραγματοποιούνται συγκρίσεις με τη χρήση στατιστικών test που αφορούν σε εξαρτημένα δείγματα (paired samples). Σε κάθε περίπτωση πρέπει να λαμβάνεται υπόψιν αν οι τιμές της εξαρτημένης μεταβλητής προέρχονται από κανονική κατανομή ή όχι. Στην πραγματικότητα χρησιμοποιούνται τα σταθμισμένα δεδομένα και των 2 περιόδων για να εξεταστεί αν υπάρχει στατιστικά σημαντική διαφορά στις παρεμβάσεις (θεραπείες) που εφαρμόζονται. Η περιγραφή του μαθηματικού υποδείγματος που ακολουθεί στην παράγραφο 3-6 παρουσιάζει αναλυτικά τη συγκεκριμένη προσέγγιση (μέσω t-tests). [5],[7]

Μια άλλη προσέγγιση που χρησιμοποιείται για την ανάλυση των δεδομένων είναι με τη βοήθεια γενικών γραμμικών μοντέλων και ειδικότερα των μεικτών μοντέλων (mixed models). Στην περίπτωση της χρήσης μεικτού μοντέλου υπεισέρχονται και οι έννοιες των σταθερών επιδράσεων (fixed effects) και των τυχαίων επιδράσεων (random effects). [7] Στο Κεφάλαιο 4 η εφαρμογή που παρουσιάζεται αναλύεται με τη βοήθεια μεικτού μοντέλου.

Γενικά για τις 2X2 crossover δοκιμές και τον τρόπο ανάλυσης των δεδομένων που προκύπτουν από το πείραμα - ο Grizzle - πρότεινε τον ακόλουθο αλγόριθμο:

Σχήμα 3-2

Προσέγγιση 2 σταδίων για ανάλυση δεδομένων 2x2 crossover δοκιμής



39

3-5. Μαθηματικό υπόδειγμα

Θα περιγράψουμε το υπόδειγμα που αφορά σε ένα 2X2 crossover σχεδιασμό στον οποίο υπάρχουν δύο αλληλουχίες AB και BA. Συγκεκριμένα έχουμε:

1^η αλληλουχία AB = Παρέμβαση (θεραπεία) A στην 1^η περίοδο η οποία ακολουθείται από την παρέμβαση (θεραπεία) B στην 2^η περίοδο

2^η αλληλουχία BA = Παρέμβαση (θεραπεία) B στην 1^η περίοδο η οποία ακολουθείται από την παρέμβαση (θεραπεία) A στην 2^η περίοδο

Εξαρτημένη μεταβλητή (απόκριση) Y = Συνεχής

Για να κατασκευάσουμε ένα στατιστικό μοντέλο υποθέτουμε ότι η ποσότητα y_{ijk} αποτελεί την τιμή που προέρχεται από την τυχαία μεταβλητή Y_{ijk} . Επομένως στις τιμές της μεταβλητής Y η οποία είναι συνεχής μπορούμε να

προσαρμόσουμε ένα μοντέλο που στη βασική του μορφή μπορεί να γραφεί ως εξής:

$$Y_{ijk} = \mu + \pi_j + \tau_{d[i,j]} + s_{ik} + e_{ijk} \quad \text{όπου} \quad (3-1)$$

μ = συνολικός μέσος (σταθερά)

π_j = επίδραση της j περιόδου, $j = 1, 2$

$\tau_{d[i,j]}$ = απευθείας επίδραση της θεραπείας που εφαρμόζεται την περίοδο j στην αλληλουχία i , $d_{[i,j]} = A, B$

s_{ik} = επίδραση k πειραματικής μονάδας (ασθενούς) στην αλληλουχία i , $i = 1, 2$ και $k = 1, 2, \dots, n$

e_{ijk} = τυχαία σφάλματα τα οποία ακολουθούν $N(0, \sigma_e^2)$

Ωστόσο το παραπάνω μοντέλο (σχέση 2-1) είναι ελλιπές δεδομένου ότι δεν περιλαμβάνει έναν όρο που είναι σημαντικός στους crossover σχεδιασμούς.

Αυτός ο όρος αναφέρεται στην επίδραση carryover. Αν συμβολίσουμε με $\lambda_{d[i,j-1]}$

την επίδραση carryover τότε η σχέση 2-1 γίνεται:

$$Y_{ijk} = \mu + \pi_j + \tau_{d[i,j]} + \lambda_{d[i,j-1]} + s_{ik} + e_{ijk} \quad \text{όπου} \quad (3-2)$$

$\lambda_{d[i,j-1]}$ = επίδραση carryover, $d_{[i,j-1]} = A, B$

Πριν προχωρήσουμε πρέπει να αναφερθεί ότι η επίδραση s_{ik} του k ασθενούς θεωρούμε ότι είναι μια τυχαία μεταβλητή που ακολουθεί κανονική κατανομή με μέση τιμή 0 και διακύμανση σ_s^2 . Αυτή η προσέγγιση ακολουθείται από τους ερευνητές στη συντριπτική πλειοψηφία των crossover σχεδιασμών. Άρα οι μεταβλητές s_{ik} και e_{ijk} επιδρούν κατά τρόπο τυχαίο στο μοντέλο, είναι ανεξάρτητες και έχουν διακυμάνσεις σ_s^2 και σ_e^2 αντίστοιχα.

Σ' ένα 2X2 crossover σχεδιασμό δεν συνίσταται ο έλεγχος της επίδρασης carryover. Ωστόσο – για να γίνει κατανοητή η λειτουργία του μοντέλου θα εντάξουμε σ' αυτό και τον παράγοντα της επίδρασης carryover. Ακολούθως δίνεται πίνακας που περιλαμβάνει τις σταθερές (fixed) και τυχαίες (random) επιδράσεις ως προς κάθε ομάδα, κάθε αλληλουχία και κάθε περίοδο. Σημειώνεται ότι στην 1^η περίοδο δεν υφίσταται επίδραση carry-over. [6],[7]

Πίνακας 3-2

Σταθερές και τυχαίες επιδράσεις για το πλήρες μοντέλο

Ομάδα	Αλληλουχία	Περίοδος 1	Περίοδος 2
1 ^η	AB	$\mu + \pi_1 + \tau_1 + s_{1k} + e_{11k}$	$\mu + \pi_2 + \tau_2 + \lambda_1 + s_{1k} + e_{12k}$
2 ^η	BA	$\mu + \pi_1 + \tau_2 + s_{2k} + e_{21k}$	$\mu + \pi_2 + \tau_1 + \lambda_2 + s_{2k} + e_{22k}$

Επειδή έχουμε υποθέσει ότι οι μεταβλητές s_{ik} και e_{ijk} είναι τυχαίες το μοντέλο με τις σταθερές (fixed) επιδράσεις αποτυπώνεται στον ακόλουθο πίνακα

Πίνακας 3-3

Σταθερές επιδράσεις για το πλήρες μοντέλο

Ομάδα	Αλληλουχία	Περίοδος 1	Περίοδος 2
1 ^η	AB	$\mu + \pi_1 + \tau_1$	$\mu + \pi_2 + \tau_2 + \lambda_1$
2 ^η	BA	$\mu + \pi_1 + \tau_2$	$\mu + \pi_2 + \tau_1 + \lambda_2$

3-5-1. Προσέγγιση με t-tests**α) Έλεγχος για επίδραση carryover**

Ο έλεγχος για την επίδραση carryover δεν συνίσταται για 2X2 crossover δοκιμές. Αν ωστόσο επιλεγεί - τότε υποθέτουμε ότι δεν υφίσταται αλληλεπίδραση θεραπείας και περιόδου (treatment by period interaction). Τούτο συμβαίνει επειδή η επίδραση carryover θεωρείται ότι συνδέεται εγγενώς με την αλληλεπίδραση περιόδου και θεραπείας. Εξαιτίας αυτής της σύνδεσης δεν απαιτείται ο έλεγχος για τη μη ύπαρξη επίδρασης carryover, δηλαδή του ελέγχου $\lambda_1 = \lambda_2 = 0$.

Ελέγχουμε την υπόθεση:

$$H_o : \lambda_1 = \lambda_2$$

vs

$$H_1 : \lambda_1 \neq \lambda_2$$

Τα αθροίσματα των τιμών (των αποκρίσεων) των ασθενών είναι:

$$t_{1k} = y_{11k} + y_{12k} \text{ για τον k ασθενή της 1ης ομάδας και}$$

$$t_{2k} = y_{21k} + y_{22k} \text{ για τον k ασθενή της 2ης ομάδας}$$

με αντίστοιχες αναμενόμενες τιμές

$$E[t_{1k}] = 2\mu + \pi_1 + \pi_2 + \tau_1 + \tau_2 + \lambda_1 \text{ και}$$

$$E[t_{2k}] = 2\mu + \pi_1 + \pi_2 + \tau_1 + \tau_2 + \lambda_2$$

Είναι προφανές ότι αν $\lambda_1 = \lambda_2$ τότε $E[t_{1k}] = E[t_{2k}]$

Έστω $\lambda_d = \lambda_1 - \lambda_2$ και $\hat{\lambda}_d = \bar{t}_1 - \bar{t}_2$ με $E[\hat{\lambda}_d] = \lambda_d$ και

$$\text{Var}[\hat{\lambda}_d] = 2(2\sigma_s^2 + \sigma^2)\left(\frac{1}{n_1} + \frac{1}{n_2}\right) = \sigma_T^2 * m \text{ όπου } m = \frac{n_1 + n_2}{n_1 n_2}$$

Η εκτίμηση της συνολικής διακύμανσης δίνεται από τον τύπο:

$$\hat{\sigma}_T^2 = \sum_{i=1}^2 \sum_{k=1}^{n_i} \frac{(t_{ik} - \bar{t}_i)^2}{n_1 + n_2 - 2}$$

Το στατιστικό που χρησιμοποιείται για τον έλεγχο επίδρασης carryover δίνεται από τη σχέση:

$$t_\lambda = \frac{\hat{\lambda}_d}{\sqrt{\hat{\sigma}_T^2 * m}} \sim t \text{ κατανομή με } n_1 + n_2 - 2 \text{ β.ε} \quad (3-3)$$

- Αν $|t_\lambda| > t_{\alpha/2, n_1 + n_2 - 2}$ τότε η H_o απορρίπτεται

- Αν $|t_\lambda| < t_{\alpha/2, n_1 + n_2 - 2}$ τότε η H_o δεν απορρίπτεται

Αν η μηδενική υπόθεση $H_o : \lambda_1 = \lambda_2$ απορρίπτεται τότε η ανάλυση σταματά. Αν πράγματι ισχύει $\lambda_1 = \lambda_2$ τότε η ανάλυση συνεχίζεται με τον έλεγχο για την ύπαρξη επίδρασης θεραπείας (treatments effects) και επίδρασης περιόδου (period effects). [6],[7]

β) Έλεγχος για επίδραση θεραπείας (treatments effects)

Ελέγχουμε την υπόθεση (ισχύει $\lambda_1 = \lambda_2$):

$$H_o : \tau_1 = \tau_2$$

vs

$$H_1 : \tau_1 \neq \tau_2$$

Οι διαφορές των τιμών (των αποκρίσεων) των ασθενών είναι:

$$d_{1k} = y_{11k} - y_{12k} \text{ για τον k ασθενή της 1ης ομάδας και}$$

$$d_{2k} = y_{21k} - y_{22k} \text{ για τον k ασθενή της 2ης ομάδας}$$

με αντίστοιχες αναμενόμενες τιμές

$$E[d_{1k}] = \pi_1 - \pi_2 + \tau_1 - \tau_2 \text{ και}$$

$$E[d_{2k}] = \pi_1 - \pi_2 + \tau_1 - \tau_2$$

Είναι προφανές ότι αν $\tau_1 = \tau_2$ τότε $E[d_{1k}] = E[d_{2k}]$

Έστω $\tau_d = \tau_1 - \tau_2$ και $\hat{\tau}_d = \frac{1}{2}(\bar{d}_1 - \bar{d}_2)$ με $E[\hat{\tau}_d] = \tau_d$ και

$$\text{Var}[\hat{\tau}_d] = \frac{2\sigma^2}{4} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) = \frac{\sigma_D^2}{4} * m$$

Η εκτίμηση της συνολικής διακύμανσης δίνεται από τον τύπο:

$$\hat{\sigma}_D^2 = \sum_{i=1}^2 \sum_{k=1}^{n_i} \frac{(d_{ik} - \bar{d}_i)^2}{n_1 + n_2 - 2}$$

Το στατιστικό που χρησιμοποιείται για τον έλεγχο επίδρασης θεραπείας δίνεται από τη σχέση:

$$t_\tau = \frac{\hat{\tau}_d}{\sqrt{\frac{\hat{\sigma}_D^2 * m}{4}}} \sim t \text{ κατανομή με } n_1 + n_2 - 2 \text{ β.ε} \quad (3-4)$$

- Αν $|t_\tau| > t_{\alpha/2, n_1+n_2-2}$ τότε η H_0 απορρίπτεται

- Αν $|t_\tau| < t_{\alpha/2, n_1+n_2-2}$ τότε η H_0 δεν απορρίπτεται [6],[7]

γ) Έλεγχος για επίδραση περιόδου (period effects)

Ελέγχουμε την υπόθεση (ισχύει $\lambda_1 = \lambda_2$):

$$H_0 : \pi_1 = \pi_2$$

vs

$$H_1 : \pi_1 \neq \pi_2$$

Οι διασταυρούμενες διαφορές των τιμών (των αποκρίσεων) των ασθενών είναι:

$$c_{1k} = d_{1k} = y_{11k} - y_{12k} \text{ για τον k ασθενή της 1ης ομάδας και}$$

$$c_{2k} = -d_{2k} = y_{22k} - y_{21k} \text{ για τον k ασθενή της 2ης ομάδας}$$

με αντίστοιχες αναμενόμενες τιμές

$$E[c_{1k}] = \pi_1 - \pi_2 + \tau_1 - \tau_2 \text{ και}$$

$$E[c_{2k}] = \pi_2 - \pi_1 + \tau_1 - \tau_2$$

Είναι προφανές ότι αν $\pi_1 = \pi_2$ τότε $E[c_{1k}] = E[c_{2k}]$

Έστω $\pi_d = \pi_1 - \pi_2$ και $\hat{\pi}_d = \frac{1}{2}(\bar{c}_1 - \bar{c}_2)$ με $E[\hat{\pi}_d] = \pi_d$ τότε το στατιστικό που

χρησιμοποιείται για τον έλεγχο επίδρασης περιόδου δίνεται από τη σχέση:

$$t_{\pi} = \frac{\hat{\pi}_d}{\sqrt{\frac{\hat{\sigma}_D^2 * m}{4}}} \sim t \text{ κατανομή με } n_1 + n_2 - 2 \text{ β.ε} \quad (3-5)$$

- Αν $|t_{\pi}| > t_{\alpha/2, n_1+n_2-2}$ τότε η H_o απορρίπτεται

- Αν $|t_{\pi}| < t_{\alpha/2, n_1+n_2-2}$ τότε η H_o δεν απορρίπτεται

Στην πράξη όταν έχουμε να αναλύσουμε ένα 2X2 crossover πείραμα ελέγχουμε την ύπαρξη (ή μη) στατιστικά σημαντικής αλληλεπίδρασης θεραπείας και περιόδου. [6],[7]

3-5-2. Προσέγγιση μέσω ανάλυσης διακύμανσης

Εκτός από τη χρήση t-tests - ένας άλλος τρόπος για να προσεγγιστούν οι υποθέσεις μας είναι με τη βοήθεια γραμμικών μοντέλων και της ανάλυσης διακύμανσης. Η ανάλυση διακύμανσης χρησιμοποιείται κυρίως σε περιπτώσεις όπου έχουμε περισσότερο πολύπλοκους σχεδιασμούς. Χωρίς να υπεισέλθουμε σε παρουσίαση του πίνακα ANOVA για ένα 2X2 crossover σχεδιασμό δίνουμε εν συντομία τα στατιστικά που χρησιμοποιούνται για τον έλεγχο των υποθέσεων που μας ενδιαφέρουν. Στα προβλήματα ανάλυσης διακύμανσης το στατιστικό που χρησιμοποιείται είναι το $F = t^2$. [6],[7]

44

α) Έλεγχος για επίδραση carryover

Ελέγχουμε την υπόθεση:

$$H_o : \lambda_1 = \lambda_2$$

vs

$$H_1 : \lambda_1 \neq \lambda_2$$

Το στατιστικό που χρησιμοποιείται για τον έλεγχο επίδρασης carryover δίνεται από τη σχέση:

$$F_{\lambda} = \frac{MS_{carryover}}{MS_{B-S Residual}} \text{ όπου MS το μέσο τετράγωνο} \quad (3-6)$$

- Αν $|F_{\lambda}| > F_{1, n_1+n_2-2}$ τότε η H_o απορρίπτεται

- Αν $|F_{\lambda}| < F_{1, n_1+n_2-2}$ τότε η H_o δεν απορρίπτεται [6],[7]

β) Έλεγχος για επίδραση θεραπείας (treatments effects)

Ελέγχουμε την υπόθεση (ισχύει $\lambda_1 = \lambda_2$):

$$H_o : \tau_1 = \tau_2$$

vs

$$H_1 : \tau_1 \neq \tau_2$$

Το στατιστικό που χρησιμοποιείται για τον έλεγχο επίδρασης θεραπείας δίνεται από τη σχέση:

$$F_\tau = \frac{MS_{Treatments}}{MS_{W-S Residual}} \quad \text{όπου MS το μέσο τετράγωνο} \quad (3-7)$$

- Αν $|F_\tau| > F_{1, n_1+n_2-2}$ τότε η H_o απορρίπτεται

- Αν $|F_\tau| < F_{1, n_1+n_2-2}$ τότε η H_o δεν απορρίπτεται [6],[7]

γ) Έλεγχος για επίδραση περιόδου (period effects)

Ελέγχουμε την υπόθεση (ισχύει $\lambda_1 = \lambda_2$):

$$H_o : \pi_1 = \pi_2$$

vs

$$H_1 : \pi_1 \neq \pi_2$$

Το στατιστικό που χρησιμοποιείται για τον έλεγχο επίδρασης περιόδου δίνεται από τη σχέση:

$$F_\pi = \frac{MS_{Periods}}{MS_{W-S Residual}} \quad \text{όπου MS το μέσο τετράγωνο} \quad (3-8)$$

- Αν $|F_\pi| > F_{1, n_1+n_2-2}$ τότε η H_o απορρίπτεται

- Αν $|F_\pi| < F_{1, n_1+n_2-2}$ τότε η H_o δεν απορρίπτεται [6],[7]

Το μεγάλο πλεονέκτημα της προσέγγισης με τη βοήθεια γραμμικών μοντέλων είναι ότι μπορεί να ελεγχθεί η καταλληλότητα και η επάρκεια του μοντέλου με τη βοήθεια της ανάλυσης υπολοίπων (Residual analysis).

Το υπόλοιπο (residual) δίνεται από τη σχέση:

$$r_{ijk} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \bar{y}_{i.k} - \bar{y}_{ij.} + \bar{y}_{i..} \quad (3-9)$$

όπου \hat{y}_{ijk} είναι η εκτιμώμενη τιμή της εξαρτημένης μεταβλητής (απόκρισης) όπως προκύπτει από το μοντέλο που έχουμε εκτιμήσει.

Τα υπόλοιπα r_{ijk} αποτελούν εκτιμήσεις των τυχαίων σφαλμάτων e_{ijk} και αν το μοντέλο που έχουμε εκτιμήσει είναι κατάλληλο, τότε τα r_{ijk} έχουν παρόμοιες ιδιότητες με τα e_{ijk} . Τα υπόλοιπα r_{ijk} ακολουθούν κανονική κατανομή με μέσο 0, ωστόσο δεν είναι ανεξάρτητα και δεν έχουν ίσες διακυμάνσεις. Σε

περίπτωση που έχουμε $n_1 = n_2$ και τα n_1, n_2 είναι αρκετά μεγάλα τότε τα υπόλοιπα r_{ijk} είναι περίπου ανεξάρτητα με ίσες διακυμάνσεις. Κατά τον έλεγχο υπολοίπων είναι προτιμότερο να χρησιμοποιούνται τα studentized υπόλοιπα. [7]

3-5-3. Προσέγγιση με Bayesian analysis

Ένας άλλος τρόπος προκειμένου να περιγραφεί και να αναλυθεί ένας 2X2 crossover σχεδιασμός είναι με τη χρήση της ανάλυσης κατά Bayes. [7] Η συγκεκριμένη προσέγγιση δεν θα περιγραφεί διότι ξεφεύγει από τους σκοπούς της συγκεκριμένης εργασίας.

3-6. Παράδειγμα 2X2 crossover δοκιμής

Δεκατέσσερα παιδιά ηλικίας από 7 έως 14 ετών που αντιμετώπιζαν πρόβλημα άσθματος κατανεμήθηκαν κατά τρόπο τυχαίο σε δύο ομάδες A και B (7 στην ομάδα A και 7 στην ομάδα B) με σκοπό τη συμμετοχή τους σε κλινική έρευνα. Για τη διεξαγωγή της έρευνας επιλέχθηκε crossover σχεδιασμός δεδομένου ότι θα εξετάζονταν δύο θεραπευτικές παρεμβάσεις (φάρμακα). Τα υπό έρευνα φάρμακα ήταν το Formoterol (A) και το Salbutamol (B). Πριν την έναρξη της έρευνας υπήρξε αποχώρηση ενός παιδιού που είχε κατανεμηθεί στην ομάδα B με αποτέλεσμα σ' αυτή να απομείνουν 6 παιδιά.

Στα 7 παιδιά της 1^{ης} ομάδας χορηγήθηκαν τα φάρμακα με τη σειρά AB ή For/Sal και στα 6 παιδιά της 2^{ης} ομάδας χορηγήθηκαν τα φάρμακα με τη σειρά BA ή Sal/For. Η δομή του 2X2 crossover πειράματος αποτυπώνεται στον ακόλουθο πίνακα:

Πίνακας 3-4

For-Sal/Sal-For crossover δοκιμή

Ομάδες ασθενών	Αλληλουχία	1 ^η Περίοδος	Περίοδος έκπλυσης	2 ^η Περίοδος
1 ^η Ομάδα	For/Sal	Formoterol	Χωρίς θεραπεία	Salbutamol
2 ^η Ομάδα	Sal/For	Salbutamol	Χωρίς θεραπεία	Formoterol

Πραγματοποιήθηκαν μετρήσεις του δείκτη PEF (Peak expiratory flow) για το σύνολο των παιδιών σύμφωνα με τις προδιαγραφές που έθετε το

πρωτόκολλο και μετά το πέρας της έρευνας οι μετρήσεις καταγράφηκαν σε ένα αρχείο.

Οι μετρήσεις του δείκτη PEF αποτέλεσαν την εξαρτημένη μεταβλητή (απόκριση) του πειράματος η οποία ήταν συνεχής. Μετά από έλεγχο με το test Shapiro – Wilk (μικρό μέγεθος δείγματος) διαπιστώθηκε ότι οι τιμές PEF ακολουθούν την κανονική κατανομή τόσο για το φάρμακο For όσο και για το φάρμακο Sal (For: $p=0.256$ και Sal: $p=0.116$). Το ερώτημα που ενδιέφερε να απαντήσουν οι ερευνητές ήταν αν υπήρχε στατιστικά σημαντική διαφοροποίηση στις τιμές PEF μετά τη χορήγηση των φαρμάκων A και B. Σημειώνεται ότι οι υψηλές τιμές του δείκτη PEF ήταν επιθυμητές.

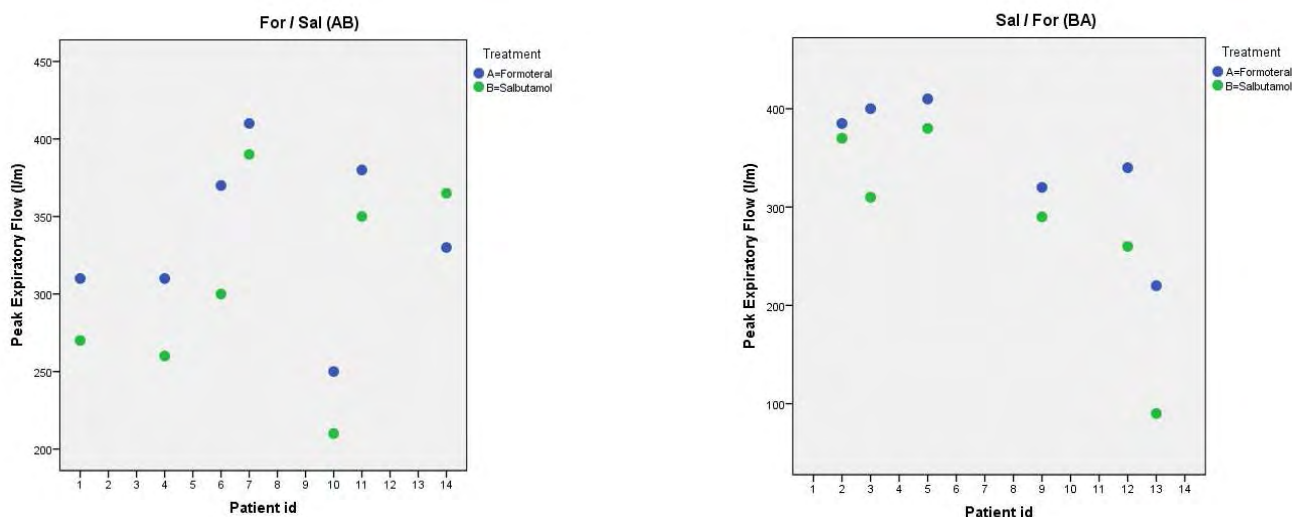
Για την ανάλυση της παραπάνω 2X2 crossover δοκιμής χρησιμοποιήθηκε το στατιστικό λογισμικό SPSS και επιλέχθηκε η προσέγγιση με τη χρήση μεικτού μοντέλου (mixed model).

Αρχικά δίνεται μια γραφική απεικόνιση των δεδομένων ανά ασθενή ως προς τις 2 αλληλουχίες For/Sal και Sal/For. Στον κάθετο άξονα κάθε γραφήματος δίνεται η τιμή PEF για κάθε ασθενή.

Γράφημα 3-7

Τιμές PEF για ομάδες For/Sal και Sal/For

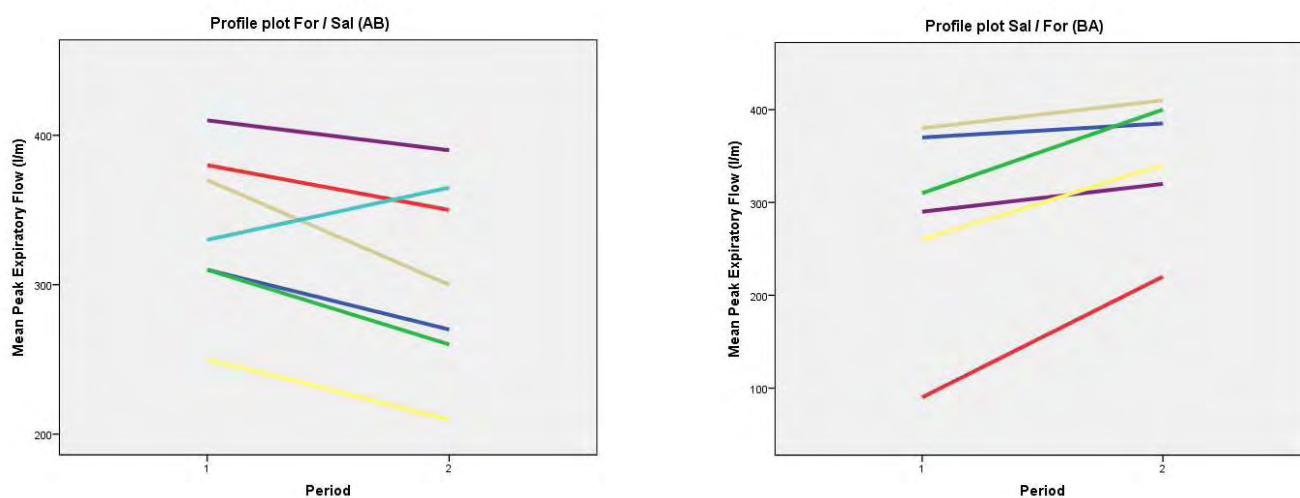
47



Στη συνέχεια δίνονται τα profiles plots των διαφορών στις τιμές PEF για κάθε ασθενή χωριστά μεταξύ 1^{ης} και 2^{ης} περιόδου και ως τις ομάδες For/Sal και Sal/For.

Γράφημα 3-8

Profiles Plots για ομάδες For/Sal και Sal/For



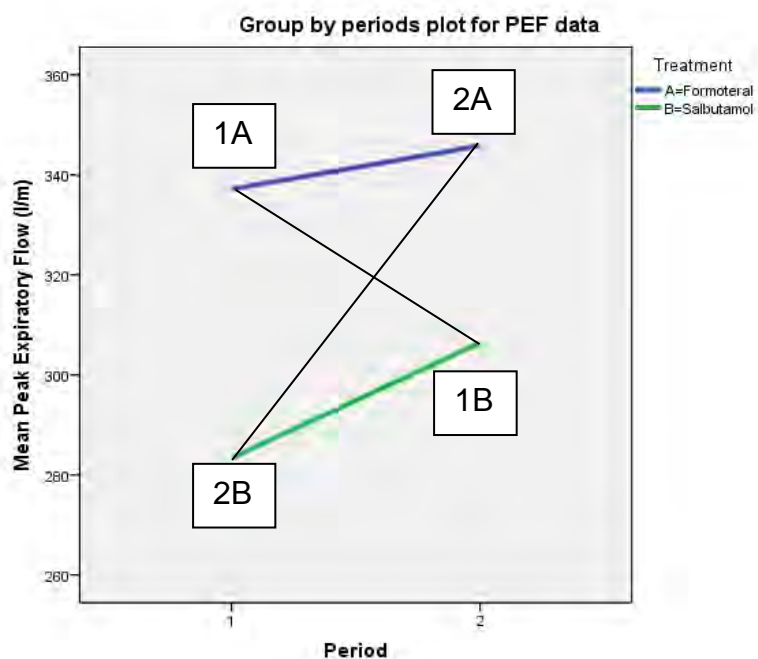
Παρατηρούμε ότι στην ομάδα For/Sal οι μετρήσεις του δείκτη PEF για το σύνολο των παιδιών (εκτός από μια περίπτωση) μειώνονται από την 1^η στη 2^η περίοδο. Αντίθετα στην ομάδα Sal/For παρατηρούμε ότι για το σύνολο των παιδιών οι τιμές αυξάνονται από την 1^η στη 2^η περίοδο.

Ακολουθεί ένα means plot όπου δίνονται οι μέσες τιμές της PEF ανά περίοδο και ομάδα.

48

Γράφημα 3-9

Means plot ως προς τις ομάδες και την περίοδο



Στο παραπάνω γράφημα φαίνεται ότι - πιθανόν - δεν υφίσταται επίδραση carryover κάτι που αποτελεί ευχάριστη είδηση. Είναι φανερό ότι η μέση τιμή της PEF μετά τη χορήγηση του φαρμάκου For είναι υψηλότερη σε σχέση με τη μέση τιμή της PEF μετά τη χορήγηση του φαρμάκου Sal τόσο για την 1^η όσο και για τη 2^η ομάδα. Συνεπώς το φάρμακο For φαίνεται ότι είναι περισσότερο αποτελεσματικό από το Sal. Υπάρχει επομένως σαφής ένδειξη ότι υφίσταται στατιστικά σημαντική επίδραση θεραπείας (treatment effects). Ακόμη ενδεχομένως να υφίσταται μια μικρή επίδραση περιόδου (period effects) θετική από την 1^η στη 2^η περίοδο. Για το αν είναι σημαντική θα αποφανθούμε μετά τη χρήση αριθμητικών υπολογισμών.

Με τη βοήθεια μεικτού μοντέλου (mixed model) θα προσεγγίσουμε μαθηματικά το πρόβλημα. Αρχικά πρέπει να ελέγξουμε αν υφίσταται επίδραση carryover κάτι που δεν φαίνεται να ισχύει παρατηρώντας το γράφημα means plot. Όπως έχει ήδη αναφερθεί – για ένα 2X2 crossover σχεδιασμό – ελέγχουμε την ύπαρξη στατιστικά σημαντικής αλληλεπίδρασης θεραπείας και περιόδου (treatment by period interaction).

Θεωρώντας ως εξαρτημένη μεταβλητή (απόκριση) τις τιμές του PEF ελέγχουμε τον παράγοντα της αλληλεπίδρασης θεραπείας και περιόδου. Επειδή γίνεται χρήση μεικτού μοντέλου (mixed model) η αλληλεπίδραση προκαλεί σταθερές επιδράσεις (fixed effects) και οι πειραματικές μονάδες (παιδιά) προκαλούν τυχαίες επιδράσεις (random effects).

Πίνακας 3-5

Type III Tests of Fixed Effects^a

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	7,350	195,341	,000
period * treat	3	21,004	1,003	,411

a. Dependent Variable: Peak Expiratory Flow (l/m).

Ο πίνακας Type III tests of fixed effects μας πληροφορεί ότι η αλληλεπίδραση θεραπείας - περιόδου δεν είναι στατιστικά σημαντική ($p=0.411$). Τούτο σημαίνει ότι – όπως αρχικά είχαμε υποθέσει – δεν υφίσταται επίδραση carryover κάτι που μας επιτρέπει την διερεύνηση ύπαρξης διαφορετικών είδους επιδράσεων. Ο ακόλουθος πίνακας που δίνει τις εκτιμήσεις των

σταθερών επιδράσεων της αλληλεπίδρασης επιβεβαιώνει το προηγούμενο συμπέρασμα.

Πίνακας 3-6

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	335,789349	34,127160	15,146	9,839	,000	263,110193	408,468504
[period=1] * [treat=1]	30,714286	38,288903	21	,802	,431	-48,911847	110,340419
[period=1] * [treat=2]	-24,018533	39,857709	21,006	-,603	,553	-106,905820	58,868754
[period=2] * [treat=1]	38,481467	39,857709	21,006	,965	,345	-44,405820	121,368754
[period=2] * [treat=2]	0 ^b	0

a. Dependent Variable: Peak Expiratory Flow (l/m).

b. This parameter is set to zero because it is redundant.

Στη συνέχεια προσαρμόζουμε ένα μεικτό μοντέλο στις τιμές της PEF λαμβάνοντας υπόψιν τους παράγοντες θεραπεία (treatment), περίοδος (period) και σειρά-αλληλουχία χορήγησης των φαρμάκων (sequence).

Τρέχοντας έχουμε:

Πίνακας 3-7

Type III Tests of Fixed Effects^a

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	11	250,558	,000
treat	1	11,000	18,704	,001
period	1	11,000	2,175	,168
seq	1	11	,032	,861

a. Dependent Variable: Peak Expiratory Flow (l/m).

Ο προηγούμενος πίνακας μας πληροφορεί ότι η επίδραση θεραπείας είναι στατιστικά σημαντική ($p=0.001$). Τούτο σημαίνει ότι υπάρχει διαφοροποίηση στην επίδραση που προκαλούν τα φάρμακα For και Sal στις τιμές της μεταβλητής PEF. Το φάρμακο For είναι καλύτερο του Sal δεδομένου ότι είναι περισσότερο αποτελεσματικό. Αντίθετα δεν υπάρχουν στατιστικά σημαντικές επιδράσεις των παραγόντων περίοδος ($p=0.168$) και σειρά χορήγησης ($p=0.861$).

Στα παραπάνω συμπεράσματα καταλήγουμε και από τη μελέτη του ακόλουθου πίνακα:

Πίνακας 3-8

Estimates of Fixed Effects^a

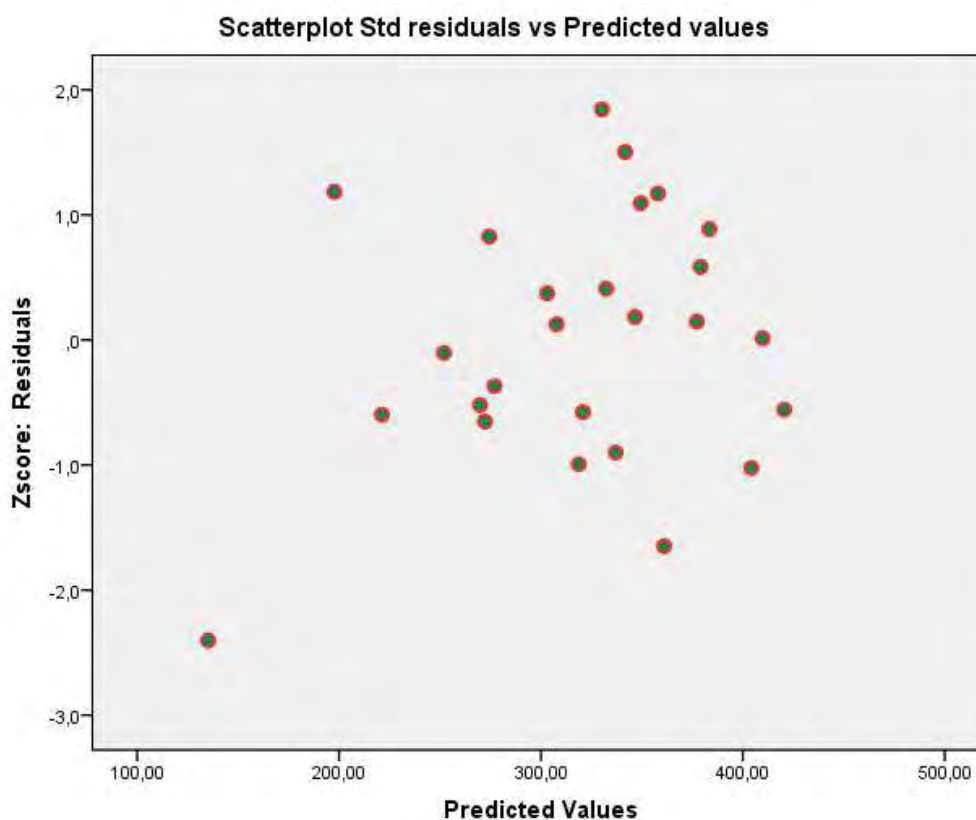
Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	299,226190	30,395565	12,350	9,844	,000	233,207415	365,244966
[treat=1]	46,607143	10,776560	11,000	4,325	,001	22,888095	70,326191
[treat=2]	0 ^b	0
[period=1]	-15,892857	10,776560	11,000	-1,475	,168	-39,611905	7,826191
[period=2]	0 ^b	0
[seq=1]	7,202381	40,202646	11	,179	,861	-81,283047	95,687809
[seq=2]	0 ^b	0

a. Dependent Variable: Peak Expiratory Flow (l/m).

b. This parameter is set to zero because it is redundant.

Το μοντέλο που έχουμε προσαρμόσει φαίνεται κατάλληλο δεδομένου ότι τα υπόλοιπα (residuals) ακολουθούν κανονική κατανομή ($p=0.927$).

Γράφημα 3-10



Ταυτόχρονα από τη μελέτη του παραπάνω διαγράμματος των τυποποιημένων υπολοίπων ως προς τις προβλεπόμενες τιμές (predicted values) παρατηρούμε ότι υπάρχει μια τυχαία κατανομή των residuals πάνω και κάτω από τη νοητή γραμμή που διέρχεται από το σημείο μηδέν. Δεν υπάρχουν συσσωρεύσεις σημείων ούτε και πρότυπα (patterns). Άρα υφίσταται η παραδοχή της γραμμικότητας και της ομοσκεδαστικότητας.

3-7. 3X3 crossover δοκιμή

3-7-1. Γενικά

Μια 3X3 crossover δοκιμή αποτελεί μια τυχαιοποιημένη μελέτη στην οποία οι πειραματικές μονάδες εκτίθενται σε μια ακολουθία (αλληλουχία) 3 διαφορετικών παρεμβάσεων σε 3 χρονικές περιόδους. Αν η crossover δοκιμή αφορά κλινική μελέτη τότε N ασθενείς κατανέμονται κατά τρόπο τυχαίο σε 3 ομάδες στις οποίες χορηγούνται 3 θεραπείες σε 3 χρονικές περιόδους. Σκοπός είναι ο εντοπισμός στατιστικά σημαντικών διαφορών ανάμεσα στις 3 θεραπείες [9], [16]

Μια crossover δοκιμή 3 περιόδων και 3 παρεμβάσεων έχει $3! = 6$ πιθανές ακολουθίες (αλληλουχίες), δηλαδή:

Πίνακας 3-9

Crossover σχεδιασμός 3 περιόδων και 3 παρεμβάσεων

Ακολουθία (Sequence)	Περίοδος (Period)		
	1η	2η	3η
ABC	A	B	C
BCA	B	C	A
CAB	C	A	B
ACB	A	C	B
BAC	B	A	C
CBA	C	B	A

Πρέπει να τονιστεί ότι η εφαρμογή λατινικού τετραγώνου απαιτεί μόνο 3 ακολουθίες (αλληλουχίες). Σ' αυτή την περίπτωση το λατινικό τετράγωνο είναι ομοιόμορφο (uniform) και ισορροπημένο (balanced).

Επειδή το παραπάνω λατινικό τετράγωνο είναι μικρό και έχει δύο βαθμούς ελευθερίας καταφεύγουμε στη μέθοδο των επαναλήψεων (του λατινικού τετραγώνου) όπως περιγράφηκε στην παράγραφο 2-2-7. Αυτή ακριβώς η μεθοδολογία χρησιμοποιείται και κατά τη στατιστική προσέγγιση μιας 3X3 crossover δοκιμής η οποία πραγματοποιείται με τη βοήθεια ενός μεικτού γραμμικού μοντέλου (mixed model).

Η στατιστική ανάλυση περισσότερο πολύπλοκων από 2X2 crossover πειραμάτων είναι περίπλοκη και δεδομένου ότι πρέπει να εξεταστεί η επίδραση της παρέμβασης (θεραπείας), η επίδραση της περιόδου, η επίδραση της αλληλουχίας (sequence) και η επίδραση της αλληλεπίδρασης παρέμβασης (θεραπείας) και περιόδου στην εξαρτημένη μεταβλητή (απόκριση) επιλέγεται η προσέγγιση μέσω μεικτού γραμμικού μοντέλου. Θυμίζουμε ότι η εξαρτημένη μεταβλητή (απόκριση) θεωρείται ότι είναι συνεχής τυχαία μεταβλητή. [16]

Ακόμη πρέπει να τονιστεί ότι σε περισσότερο πολύπλοκα crossover πειράματα (από 3X3 και πάνω) μελετάται χωριστά η επίδραση carry-over και η αλληλεπίδραση θεραπείας και περιόδου (treatment by period interaction). Ένα σημαντικό και δύσκολο ζήτημα είναι η κωδικοποίηση της επίδρασης carry-over από τους ερευνητές. Η δημιουργία δηλαδή μιας νέας μεταβλητής στο αρχείο δεδομένων που θα αποτυπώνει την επίδραση carry-over. Για παράδειγμα, σε μια 3X3 crossover κλινική δοκιμή όπου δοκιμάζονται τρία φάρμακα A, B και C στην αλληλουχία ABC (χορήγηση A φαρμάκου στην 1^η περίοδο, B φαρμάκου στη 2^η περίοδο κοκ) είναι προφανές ότι στην 1^η περίοδο δεν υπάρχει επίδραση carry-over. Αυτό κωδικοποιείται από τους ερευνητές με μηδέν. Στη 2^η περίοδο (χορήγηση B φαρμάκου) μπορεί να υφίσταται επίδραση carry-over η οποία στην προκειμένη περίπτωση κωδικοποιείται με A. Θυμίζουμε ότι η αλληλουχία που εξετάζουμε είναι η ABC πράγμα που σημαίνει ότι στην 1^η περίοδο χορηγήθηκε το φάρμακο A. Ομοίως στην 3^η περίοδο μπορεί να υφίσταται επίδραση carry-over η οποία κωδικοποιείται με B κοκ. Σε κάθε περίπτωση κατά τη στατιστική ανάλυση η μεταβλητή carry-over λαμβάνεται ως κατηγορική που για ένα 3X3 crossover πείραμα θα έχει 4 κατηγορίες (Null, A, B, C). Όλα τα παραπάνω

μπορούν να αποτυπωθούν εποπτικά με τη βοήθεια του παρακάτω πίνακα (για την αλληλουχία ABC):

Πίνακας 3-10

Κωδικοποίηση επίδρασης carry-over σε 3X3 crossover δοκιμή

	Περίοδος		
Αλληλουχία	1η	2η	3η
ABC	A	B	C
	↓	↓	↓
Κωδικογράφηση επίδρασης carryover	Null (ή 0)	A (ή 1)	B (ή 2)

Ένας άλλος τρόπος κωδικοποίησης της επίδρασης carry-over είναι με τη βοήθεια ψευδομεταβλητών. Σε ένα 3X3 crossover πείραμα με τρεις παρεμβάσεις (θεραπείες) A, B και C δημιουργούμε δύο νέες μεταβλητές X_1 και X_2 και κωδικοποιούμε ως εξής:

54

Πίνακας 3-11

Κωδικοποίηση επίδρασης carry-over με τη χρήση ψευδομεταβλητών

Θεραπεία	X_1	X_2
A	1	0
B	0	1
C	-1	-1

Τούτο σημαίνει ότι στο αρχείο δεδομένων δημιουργούνται 2 νέες μεταβλητές η X_1 και η X_2 οι οποίες κατά τη στατιστική ανάλυση εντάσσονται στο μοντέλο ως συμμεταβλητές (covariates). Για να αποφανθούμε ωστόσο για τη σημαντικότητα της επίδρασης carry-over θα πρέπει να τρέξουμε το μοντέλο και χωρίς τις μεταβλητές X_1 και X_2 και εν συνεχεία να συγκρίνουμε το

αποτελέσματά του με το αρχικό. Αυτό πραγματοποιείται με τη βοήθεια του στατιστικού $\Delta G^2 = (-\log L_{Reduced}) - (-\log L_{Full})$. [24]

3-7-2. Μεικτό μοντέλο

Τα μεικτά μοντέλα (mixed models) αναφέρονται σε μια ευρεία γκάμα πολύπλοκων γραμμικών μοντέλων που έχουν ως κύριο χαρακτηριστικό την ύπαρξη επιδράσεων (effects) οι οποίες είναι είτε σταθερές (fixed) είτε τυχαίες (random). Αυτές οι επιδράσεις προκαλούνται από τις επεξηγηματικές (ανεξάρτητες) μεταβλητές του μοντέλου οι οποίες στα συγκεκριμένα μοντέλα είναι κατηγορικές.

Ο διαχωρισμός των μεταβλητών που προκαλούν είτε σταθερές (fixed) είτε τυχαίες (random) επιδράσεις δεν είναι εύκολος. Η ύπαρξη σταθερών (fixed) επιδράσεων συνδέεται με την υπόθεση ότι οι τιμές της μεταβλητής είναι μεταξύ τους ανεξάρτητες. Για παράδειγμα, η μεταβλητή φύλο θεωρείται ότι προκαλεί σταθερές (fixed) επιδράσεις διότι παίρνει δύο τιμές (άνδρας – γυναίκα) οι οποίες μεταξύ τους είναι αμοιβαία αποκλειόμενες (ανεξάρτητες) και αμετάβλητες εις το διηνεκές. Από την άλλη η ύπαρξη τυχαίων (random) επιδράσεων υπονοεί ότι υφίσταται κάποιου είδους συσχέτιση μεταξύ των τιμών της μεταβλητής. Για παράδειγμα, η μεταβλητή τάξη δημοτικού θεωρείται ότι προκαλεί τυχαίες (random) επιδράσεις δεδομένου ότι συνήθως λαμβάνεται δείγμα παρατηρήσεων (όχι από όλες τις τάξεις) και κανένας μαθητής δεν παραμένει εσαεί στην ίδια τάξη. [17] **Πρέπει να τονιστεί ότι οι μεταβλητές που προκαλούν σταθερές (fixed) επιδράσεις επηρεάζουν μόνο τη μέση τιμή της εξαρτημένης μεταβλητής (απόκρισης). Από την άλλη, οι μεταβλητές που προκαλούν τυχαίες (random) επιδράσεις επηρεάζουν μόνο τη διακύμανση της εξαρτημένης μεταβλητής (απόκρισης).**

Στην πραγματικότητα ένα μεικτό μοντέλο χρησιμοποιείται για να επιλύσει ένα σύνθετο πρόβλημα ανάλυσης διακύμανσης. Ένα μεικτό μοντέλο αποτελεί συνδυασμό επαναλαμβανόμενων μετρήσεων μιας συνεχούς μεταβλητής ως προς τις ομάδες (κατηγορίες) περισσότερο των δύο ανεξάρτητων κατηγορικών μεταβλητών. Για παράδειγμα, αν έχουμε δύο κατηγορικές μεταβλητές X_1 και X_2 και επιθυμούμε να μελετήσουμε τις από κοινού επιδράσεις αυτών των μεταβλητών στην εξαρτημένη μεταβλητή (απόκριση), η μεταβλητή X_1

λαμβάνεται ως τυχαία (random) και η μεταβλητή X_2 λαμβάνεται ως σταθερή (fixed). [17]

Για να γίνει περισσότερο κατανοητό ένα μεικτό μοντέλο παραθέτουμε ένα μοντέλο με δύο ανεξάρτητες κατηγορικές μεταβλητές (one fixed and one random) στο οποίο λαμβάνουμε υπόψιν και την αλληλεπίδρασή τους:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad \begin{array}{l} i = 1, 2, \dots, r \\ j = 1, 2, \dots, c \\ k = 1, 2, \dots, n \end{array}$$

όπου Y_{ijk} = η k παρατήρηση της i γραμμής και της j στήλης

μ = ο συνολικός μέσος

α_i = το αποτέλεσμα της i γραμμής για τον παράγοντα α - Ο παράγοντας α προκαλεί τυχαίες επιδράσεις (random effects) και ακολουθεί $N(0, \sigma_\alpha^2)$

β_j = το αποτέλεσμα της j στήλης για τον παράγοντα που προκαλεί σταθερές επιδράσεις (fixed effects) και

$(\alpha\beta)_{ij}$ = αλληλεπίδραση παράγοντα α με τον παράγοντα β

e_{ijk} = τα τυχαία σφάλματα που ακολουθούν $N(0, \sigma^2)$ [16]

Ο πίνακας ανάλυσης διακύμανσης για το προηγούμενο μοντέλο προσομοιάζει με εκείνο των λατινικών τετραγώνων που δόθηκε στο κεφάλαιο 2.

Κεφάλαιο 4

Εφαρμογή: Test γευσιγνωσίας σοκολάτας

(Χαϊδέψτε έναν κύκλο, και θα γίνει φαύλος – Ευγένιος Ιονέσκο)

4-1. Το πείραμα

Τα δεδομένα του test γευσιγνωσίας σοκολάτας είναι πραγματικά και προέρχονται από μελέτη που διενεργήθηκε από το Εργαστήριο Τεχνολογίας, Ελέγχου Ποιότητας και Ασφάλειας Τροφίμων του Τμήματος Γεωπονίας, Φυτικής Παραγωγής και Αγροτικού Περιβάλλοντος του Πανεπιστημίου Θεσσαλίας. Εξήντα (60) γευσιγνώστες κλήθηκαν να δοκιμάσουν – με διαφορετική σειρά - τρεις ίδιες σοκολάτες A, B και C όπου καθεμία περιείχε διαφορετική γλυκαντική ουσία. Συγκεκριμένα, η σοκολάτα A περιείχε ως γλυκαντική ουσία λευκή ζάχαρη, η σοκολάτα B καστανή ζάχαρη και η σοκολάτα C στέβια. Σημειώνεται ότι κανένας γευσιγνώστης δεν γνώριζε το είδος της γλυκαντικής ουσίας και τί περιλάμβανε κάθε σοκολάτα.

Στη συνέχεια, οι 60 γευσιγνώστες κατανεμήθηκαν με τυχαίο τρόπο σε μία από τις παρακάτω έξι (6) ομάδες:

- α) 10 γευσιγνώστες στην ομάδα ABC (1^η δοκιμή με σοκολάτα A, 2^η δοκιμή με σοκολάτα B και 3^η δοκιμή με σοκολάτα C)
- β) 10 γευσιγνώστες στην ομάδα CAB (1^η δοκιμή με σοκολάτα C, 2^η δοκιμή με σοκολάτα A και 3^η δοκιμή με σοκολάτα B)
- γ) 10 γευσιγνώστες στην ομάδα BCA (1^η δοκιμή με σοκολάτα B, 2^η δοκιμή με σοκολάτα C και 3^η δοκιμή με σοκολάτα A)
- δ) 10 γευσιγνώστες στην ομάδα CBA (1^η δοκιμή με σοκολάτα C, 2^η δοκιμή με σοκολάτα B και 3^η δοκιμή με σοκολάτα A)
- ε) 10 γευσιγνώστες στην ομάδα ACB (1^η δοκιμή με σοκολάτα A, 2^η δοκιμή με σοκολάτα C και 3^η δοκιμή με σοκολάτα B)
- στ) 10 γευσιγνώστες στην ομάδα BAC (1^η δοκιμή με σοκολάτα B, 2^η δοκιμή με σοκολάτα A και 3^η δοκιμή με σοκολάτα C).

Κάθε γευσιγνώστης - μετά από κάθε δοκιμή - βαθμολόγησε τις τρεις σοκολάτες ως προς τα ακόλουθα πέντε (5) χαρακτηριστικά:

- α) χρώμα (ανοικτό – σκούρο)
- β) σκληρότητα (μαλακή – σκληρή)
- γ) γλυκύτητα (πικρή – γλυκιά)
- δ) επίγευση (απαλή – έντονη) και
- ε) προτίμηση (μικρότερη – μεγαλύτερη)

Η κλίμακα της βαθμολογίας (σκορ) που χρησιμοποίησε κάθε γευσιγνώστης ήταν από 1 έως και 5 (από το αρνητικό προς το θετικό). Στην ουσία, κάθε γευσιγνώστης μπορούσε να βαθμολογήσει με 1, 2, 3, 4 και 5 το καθένα από τα παραπάνω χαρακτηριστικά. Από αυτά τα χαρακτηριστικά επιλέχθηκε το πέμπτο – η βαθμολογία ως προς την προτίμηση (στο εξής προτίμηση) – για περαιτέρω μελέτη. Επομένως, η μεταβλητή ενδιαφέροντος είναι η «προτίμηση» και αποτελεί για το πείραμά μας την εξαρτημένη μεταβλητή (απόκριση). Η κλίμακα μέτρησης της εξαρτημένης μεταβλητής αποτελεί μια 5-βάθμια κλίμακα Likert η οποία λαμβάνει διακριτές τιμές.

Ως περίοδος έκπλυσης (washout period) για τα test γευσιγνωσίας γενικά, αποτελεί η ακόλουθη διαδικασία:

Ο γευσιγνώστης δοκιμάζει την 1^η μπουκιά/γουλιά. Χρειάζεται κάποια λεπτά για να γευτεί και να κατεβάσει τη μπουκιά/γουλιά και πίνει νερό. Πριν προχωρήσει στη δοκιμή της 2^{ης} μπουκιάς/γουλιάς ξεπλένει καλά το στόμα του με νερό. Κάτι ανάλογο συνέβη και στο test γευσιγνωσίας σοκολάτας.

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Από τα παραπάνω είναι φανερό ότι το test γευσιγνωσίας σοκολάτας αποτελεί μια 3X3 μελέτη διασταυρούμενης μετάβασης (3X3 crossover δοκιμή) όπου 60 γευσιγνώστες (πειραματικές μονάδες) δοκιμάζουν τρεις σοκολάτες A, B και C (παρεμβάσεις ή θεραπείες) σε 3 διαφορετικές χρονικές περιόδους 1^η, 2^η και 3^η. Οι γευσιγνώστες έχουν τυχαιοποιηθεί σε έξι (6) ομάδες ABC, CAB, BCA, CBA, ACB και BAC που αποτελούν τις αλληλουχίες (σειρές) με την οποία δοκιμάζονται οι τρεις σοκολάτες.

4-2. Στόχος

Μεταβλητή ενδιαφέροντος ή μελέτης αποτελεί η βαθμολογία ως προς την προτίμηση η οποία για τις ανάγκες του πειράματος αλλά κυρίως για τη στατιστική ανάλυση λαμβάνεται ως συνεχής μεταβλητή (scale).

Κύριος στόχος της crossover δοκιμής είναι να διερευνηθεί αν διαφέρουν σημαντικά οι τρεις σοκολάτες (καθεμία με άλλη γλυκαντική ουσία) ως προς την προτίμηση. Ακόμη, αντικείμενα μελέτης είναι αν είναι σημαντική η σειρά δοκιμής (η σειρά με την οποία δόθηκαν οι σοκολάτες στους γευσιγνώστες) καθώς η περίοδος (χρονική στιγμή) δοκιμής κάθε σοκολάτας.

Οι παραπάνω έλεγχοι πραγματοποιούνται με τη βοήθεια τεσσάρων (4) γνωστών στατιστικών λογισμικών, ήτοι JMP, Minitab, R και SPSS. Τα συγκεκριμένα λογισμικά είναι αρκετά γνωστά και χρησιμοποιούνται ευρέως από στατιστικούς και άλλους επιστήμονες. Ειδικότερα, το JMP και το Minitab αποτελούν κλασσικά λογισμικά ανάλυσης πειραματικών σχεδίων (όπως το συγκεκριμένο crossover πείραμα), το R είναι το πιο διαδεδομένο freeware λογισμικό και το SPSS το πιο εμπορικό.

4-3. Αρχείο δεδομένων (Data file)

Οι εξήντα (60) γευσιγνώστες αποτελούν τις πειραματικές μονάδες και θεωρητικά αναμένουμε από τον καθένα τρεις (3) βαθμολογίες ως προς την προτίμηση. Οι πειραματικές μονάδες (cases) αποτυπώνονται σε ένα αρχείο δεδομένων ως σειρές στο φύλλο δεδομένων. Αντίθετα, οι μεταβλητές (variables) αποτυπώνονται σε στήλες.

Το αρχείο δεδομένων (data file) του πειράματος αποτελείται από τις ακόλουθες μεταβλητές:

-Εξαρτημένη μεταβλητή (απόκριση) = Προτίμηση (preference) – Διακριτή μεταβλητή ιεράρχησης (ordinal) με κλίμακα μέτρησης Likert από 1 έως 5

-Ανεξάρτητες μεταβλητές (επεξηγηματικές)=

α) Αλληλουχία (σειρά) δοκιμής σοκολατών (sequence) – Κατηγορική μεταβλητή (nominal) με έξι (6) κατηγορίες ABC, CAB, BCA, CBA, ACB και BAC

β) Είδος Σοκολάτας (treatment) – Κατηγορική μεταβλητή (nominal) με τρεις (3) κατηγορίες A, B και C και

γ) Περίοδος δοκιμής σοκολάτας (period) – Κατηγορική μεταβλητή (nominal) με τρεις (3) κατηγορίες 1^η, 2^η και 3^η.

Έχει ήδη αναφερθεί ότι η εξαρτημένη μεταβλητή «προτίμηση» λαμβάνεται ως συνεχής τυχαία μεταβλητή, παρά το γεγονός ότι οι τιμές της είναι 5-βάθμιας κλίμακας Likert. **Σε προβλήματα τέτοιου τύπου η κλίμακα Likert θεωρείται ότι αντιστοιχεί σε λανθάνουσα (latent) κανονική κατανομή.** Επιπλέον, εξαιτίας της ύπαρξης ικανοποιητικού μεγέθους δείγματος (n=180) έχει εφαρμογή το Κεντρικό Οριακό Θεώρημα (Central Limit Theorem). [24]

4-4. Επίδραση ακολουθίας (carry-over effects)

Στα προηγούμενα κεφάλαια περιγράφηκε διεξοδικά το φαινόμενο της ύπαρξης επίδρασης ακολουθίας (carry-over effects) και το ρόλο που διαδραματίζει σε ένα crossover πείραμα. Για να διερευνηθεί η ύπαρξη σημαντικής (ή μη) επίδρασης carry-over θα πρέπει πρώτα να αποτυπωθεί. Τούτο σημαίνει ότι στο αρχείο δεδομένων (data file) θα πρέπει να δημιουργηθεί μια νέα μεταβλητή που θα εκφράζει την επίδραση carry-over. Η δημιουργία αυτής της μεταβλητής είναι υπόθεση και απόφαση του ερευνητή.

A priori, θεωρούμε ότι υφίσταται **επίδραση carry-over 1^{ης} τάξης. Δηλαδή, η επίδραση carry-over διαρκεί ακριβώς μια περίοδο.** Άρα, εφόσον υπάρχει επίδραση carry-over αυτή εμφανίζεται μόνο στην αμέσως επόμενη περίοδο και δεν συνεχίζει να υπάρχει στη μεθεπόμενη (περίοδο).

Η λογική της επίδρασης carry-over 1^{ης} τάξης σε ένα 3X3 crossover πείραμα παρουσιάζεται στον ακόλουθο πίνακα:

Πίνακας 4-1

Επίδραση carry-over σε 3X3 crossover πείραμα

Αλληλουχία	Περίοδος		
	1 ^η	2 ^η	3 ^η
ABC	A	B-a	C-b
ACB	A	C-a	B-c
BAC	B	A-b	C-a
BCA	B	C-b	A-c
CBA	C	B-c	A-b
CAB	C	A-c	B-a

Η επίδραση carry-over εμφανίζεται με ένα μικρό γράμμα δίπλα από το κεφαλαίο γράμμα. Παρατηρούμε ότι στην 1^η περίοδο δεν υπάρχει επίδραση carry-over κάτι που είναι απολύτως αναμενόμενο και λογικό. Στην αλληλουχία ABC έχουμε επίδραση carry-over a (από τη σοκολάτα A) στη 2^η περίοδο κατά τη δοκιμή της σοκολάτας B. Και πάλι στην αλληλουχία ABC, έχουμε επίδραση carry-over b (από τη σοκολάτα B στη 2^η περίοδο) στην 3^η περίοδο κατά τη δοκιμή της σοκολάτας C.

Η κωδικοποίηση της επίδρασης carry-over δίνεται στον ακόλουθο πίνακα:

Πίνακας 4-2

Κωδικοποίηση carry-over σε 3X3 crossover πείραμα

Αλληλουχία	1η Περίοδος		2η Περίοδος		3η Περίοδος	
	Δοκιμή	Επίδραση carry-over	Δοκιμή	Επίδραση carry-over	Δοκιμή	Επίδραση carry-over
ABC	A	Null ή 0	B	A ή 1	C	B ή 2
ACB	A	Null ή 0	C	A ή 1	B	C ή 3
BAC	B	Null ή 0	A	B ή 2	C	A ή 1
BCA	B	Null ή 0	C	B ή 2	A	C ή 3
CBA	C	Null ή 0	B	C ή 3	A	B ή 2
CAB	C	Null ή 0	A	C ή 3	B	A ή 1

Η νέα μεταβλητή carry-over ή residual effects του αρχείου δεδομένων είναι κατηγορική μεταβλητή με τέσσερις (4) κατηγορίες Null, A, B και C (ή 0, 1, 2 και 3).

4-5. Έλεγχος για ελλείπουσες και ακραίες τιμές

Πριν ξεκινήσουμε με την περιγραφική παρουσίαση των μεταβλητών θα πρέπει να ελέγξουμε το αρχείο δεδομένων για την ύπαρξη ελλειπουσών τιμών (missing values). Με τη βοήθεια του λογισμικού R και με την εντολή `apply(is.na(cho),2,sum)` διαπιστώνουμε ότι δεν υπάρχει ελλείπουσα τιμή σε καμία μεταβλητή στο αρχείο δεδομένων. Πράγματι:

Πίνακας 4-3

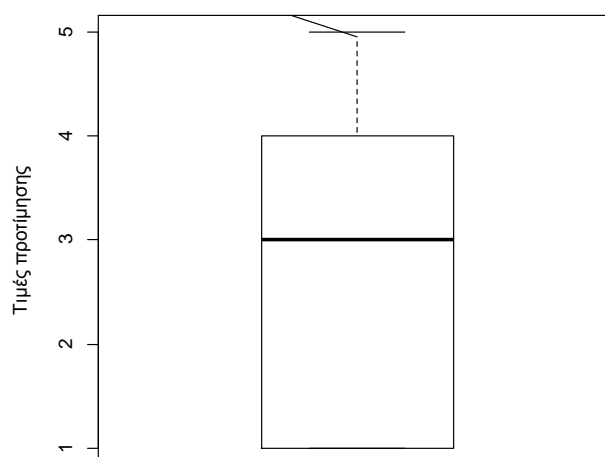
Έλεγχος για ελλείπουσες τιμές

Variables	id	Sequence	Preference	Period	Carry-over
Missing values (N)	0	0	0	0	0

Στη συνέχεια ελέγχουμε την ύπαρξη ακραίων τιμών στις τιμές της εξαρτημένης μεταβλητής «προτίμηση» με τη βοήθεια ενός θηκογράμματος. Η εντολή που μας δίνει το θηκόγραμμα είναι `boxplot(cho$prefer,main="Θηκόγραμμα προτίμησης",ylab="Τιμές προτίμησης")`.

Γράφημα 4-1

Θηκόγραμμα προτίμησης



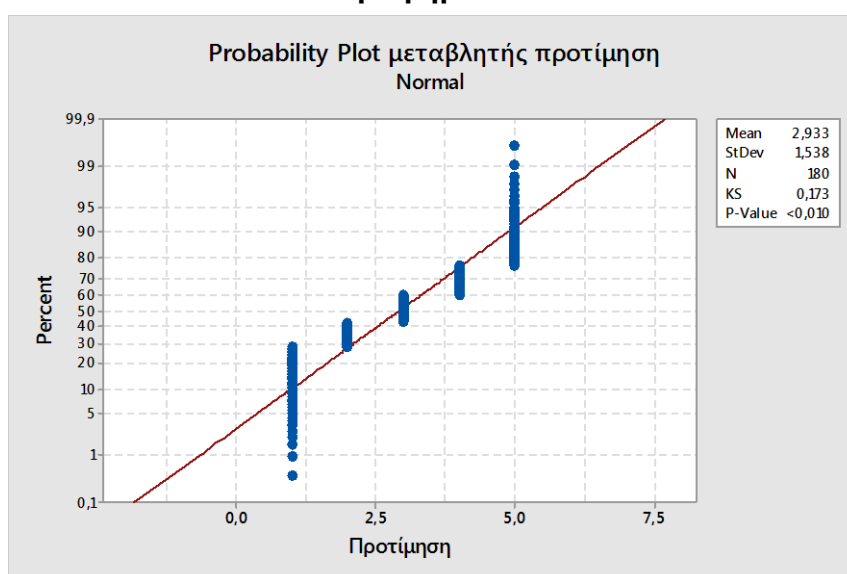
Παρατηρούμε ότι δεν υφίστανται ακραίες τιμές (outliers) σε αυτές της εξαρτημένης μεταβλητής «προτίμηση».

4-6. Περιγραφική στατιστική

4-6-1. Εξαρτημένη μεταβλητή (Περιγραφικά μέτρα – Κατανομές – Κανονικότητα)

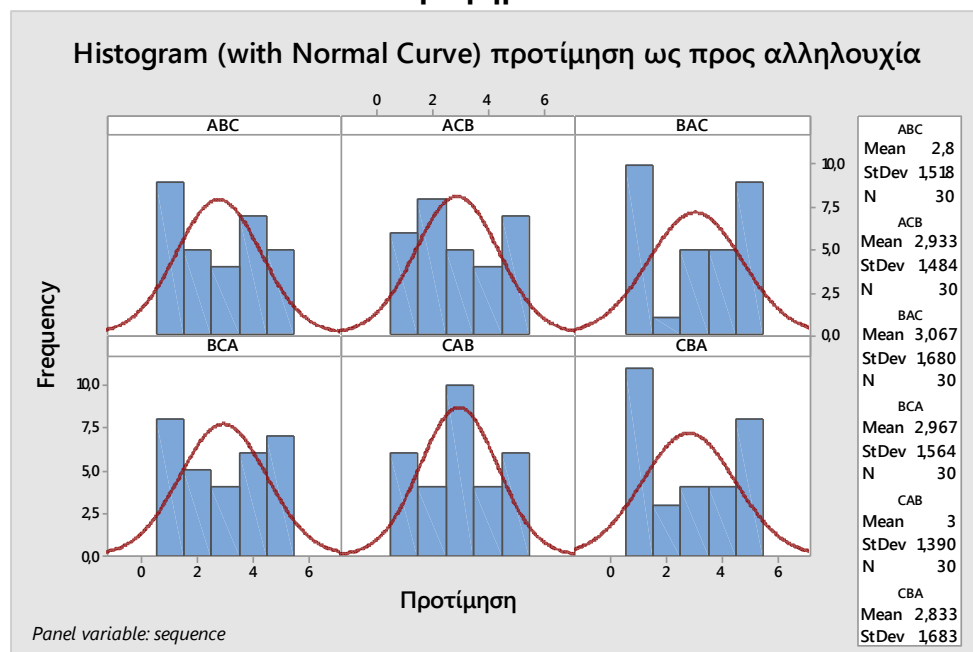
Με τη βοήθεια των λογισμικών Minitab και SPSS πραγματοποιούμε έλεγχο κανονικότητας της εξαρτημένης μεταβλητής «προτίμηση» και υπολογίζουμε κάποια βασικά περιγραφικά μέτρα. Αυτό που επιθυμούμε είναι οι τιμές της μεταβλητής «προτίμηση» να ακολουθούν κανονική κατανομή (από μόνες τους) αλλά και στις κατηγορίες των ανεξάρτητων μεταβλητών αλληλουχία (sequence), είδος σοκολάτας (treatment), περίοδος (period) και επίδραση ακολουθίας (carry-over). Από το γράφημα 4-2 παρατηρούμε ότι οι τιμές της μεταβλητής «προτίμηση» δεν ακολουθούν κανονική κατανομή ($p = 0.01$).

Γράφημα 4-2



Στο επόμενο γράφημα 4-3 δίνονται έξι (6) ιστογράμματα της μεταβλητής «προτίμηση» ως προς τις κατηγορίες της μεταβλητής αλληλουχία (σειρά) δοκιμής των σοκολατών. Παρατηρούμε ότι οι τιμές της μεταβλητής «προτίμηση» δεν ακολουθούν σε καμιά από τις 6 κατηγορίες κανονική κατανομή. Αυτό επιβεβαιώνεται και από τον πίνακα 4-4. Ταυτόχρονα, στο γράφημα 4-3 δίνονται ορισμένα βασικά περιγραφικά μέτρα της μεταβλητής «προτίμηση» ως προς τις 6 κατηγορίες. Για παράδειγμα, στην αλληλουχία BAC έχουμε τη μεγαλύτερη μέση τιμή της μεταβλητής «προτίμηση».

Γράφημα 4-3



Πίνακας 4-4

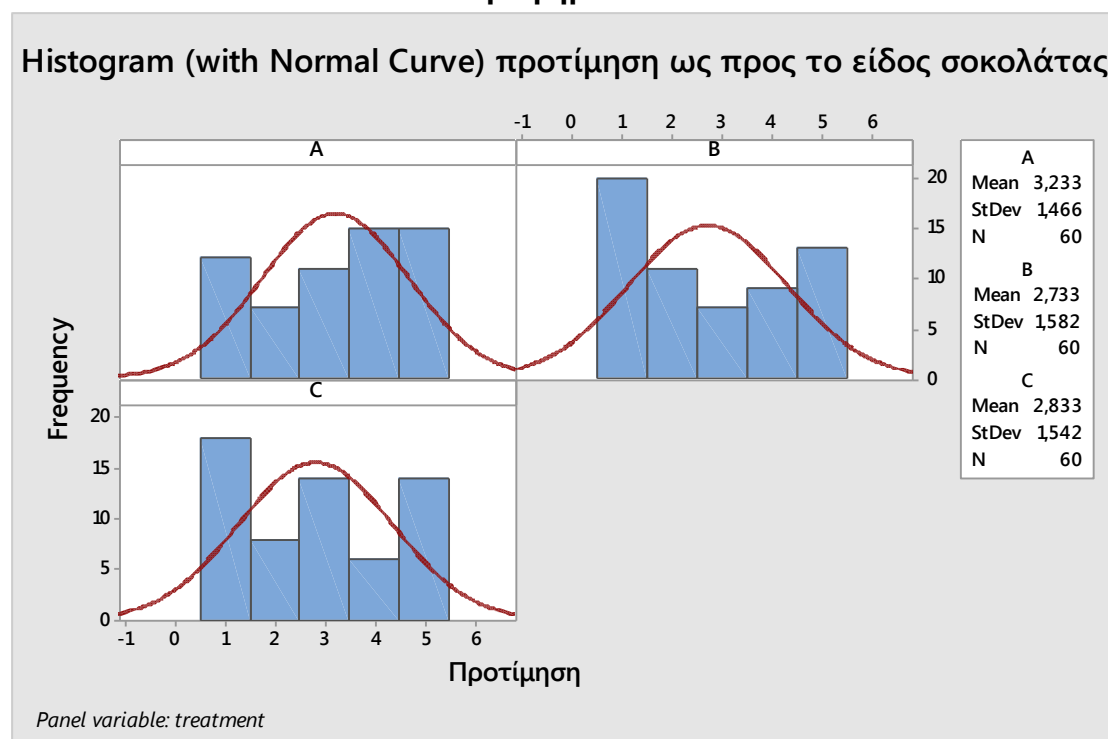
Tests of Normality

Sequence	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Choco Preference ABC	,185	30	,010	,860	30	,001
CAB	,167	30	,033	,891	30	,005
BCA	,179	30	,015	,857	30	,001
CBA	,229	30	,000	,812	30	,000
ACB	,202	30	,003	,872	30	,002
BAC	,224	30	,001	,811	30	,000

a. Lilliefors Significance Correction

Στο επόμενο γράφημα 4-4 παραθέτουμε τρία (3) ιστογράμματα της μεταβλητής «προτίμηση» ως προς τις κατηγορίες της μεταβλητής είδος σοκολάτας. Παρατηρούμε ότι οι τιμές της μεταβλητής «προτίμηση» δεν ακολουθούν σε καμιά από τις 3 κατηγορίες κανονική κατανομή. Αυτό επιβεβαιώνεται και από τον πίνακα 4-5. Επιπρόσθετα, στο γράφημα 4-4 δίνονται ορισμένα βασικά περιγραφικά μέτρα της μεταβλητής «προτίμηση» ως προς τις 3 κατηγορίες. Για παράδειγμα, για τη σοκολάτα A έχουμε τη μεγαλύτερη μέση τιμή της μεταβλητής «προτίμηση».

Γράφημα 4-4



Πίνακας 4-5

Tests of Normality

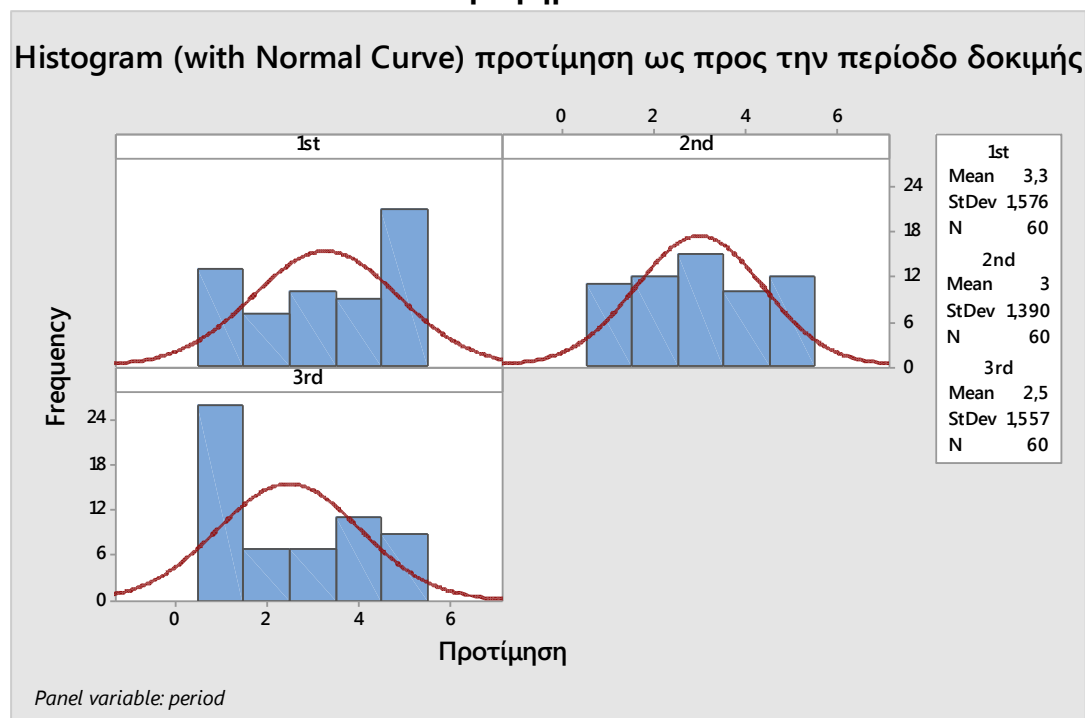
Chocolate	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Choco Preference A	,200	60	,000	,866	60	,000
B	,197	60	,000	,833	60	,000
C	,183	60	,000	,851	60	,000

a. Lilliefors Significance Correction

Στο επόμενο γράφημα 4-5 παραθέτουμε τρία (3) ιστογράμματα της μεταβλητής «προτίμηση» ως προς τις κατηγορίες της μεταβλητής περιόδου δοκιμής. Παρατηρούμε ότι οι τιμές της μεταβλητής «προτίμηση» δεν ακολουθούν σε καμιά από τις 3 κατηγορίες κανονική κατανομή. Αυτό επιβεβαιώνεται και από τον πίνακα 4-6. Επιπλέον, στο γράφημα 4-5 δίνονται ορισμένα βασικά περιγραφικά μέτρα της μεταβλητής «προτίμηση» ως προς τις 3 κατηγορίες. Για παράδειγμα, στην 1^η περίοδο έχουμε τη μεγαλύτερη μέση τιμή της μεταβλητής «προτίμηση». Το αποτέλεσμα αυτό φαντάζει απολύτως λογικό. Είναι γνωστό πως όταν τρώμε σοκολάτα επηρεάζεται το δίκτυο των νευροδιαβιβαστών, δηλαδή των χημικών αγγελιοφόρων του εγκεφάλου, που λειτουργούν μεταφέροντας ηλεκτρικά μηνύματα ανάμεσα στους νευρώνες του. Αυτά τα μηνύματα προκαλούν τις αλλαγές στις αισθήσεις

και στα συναισθήματα που νιώθουμε. Επομένως, η κατανάλωση σοκολάτας στην 1^η περίοδο προκαλεί μεγαλύτερη αίσθηση σε σχέση με τις επόμενες.

Γράφημα 4-5



Πίνακας 4-6

Tests of Normality

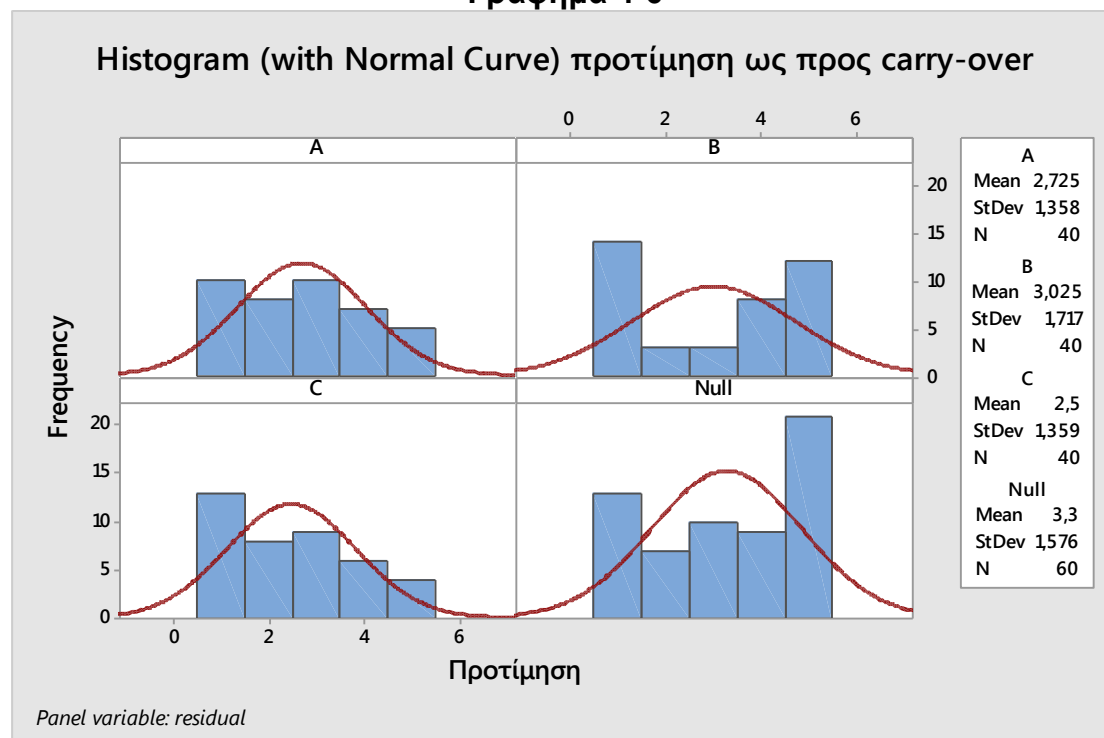
Period	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Choco Preference 1	,210	60	,000	,833	60	,000
2	,147	60	,002	,895	60	,000
3	,266	60	,000	,806	60	,000

a. Lilliefors Significance Correction

Στο επόμενο γράφημα 4-6 δίνονται τέσσερα (4) ιστογράμματα της μεταβλητής «προτίμηση» ως προς τις κατηγορίες της μεταβλητής επίδραση carry-over. Παρατηρούμε ότι οι τιμές της μεταβλητής «προτίμηση» δεν ακολουθούν σε καμιά από τις 4 κατηγορίες κανονική κατανομή. Αυτό επιβεβαιώνεται και από τον πίνακα 4-7. Παράλληλα, στο γράφημα 4-6 δίνονται ορισμένα βασικά περιγραφικά μέτρα της μεταβλητής «προτίμηση» ως προς τις 3 κατηγορίες. Για παράδειγμα, στην κατηγορία Null (έλλειψη carry-over λόγω 1^{ης} περιόδου) έχουμε τη μεγαλύτερη μέση τιμή της μεταβλητής «προτίμηση».

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Γράφημα 4-6



Πίνακας 4-7

Tests of Normality

Carryover effect	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Choco Preference A	,153	40	,019	,894	40	,001
B	,231	40	,000	,793	40	,000
C	,190	40	,001	,872	40	,000
Null	,210	60	,000	,833	60	,000

a. Lilliefors Significance Correction

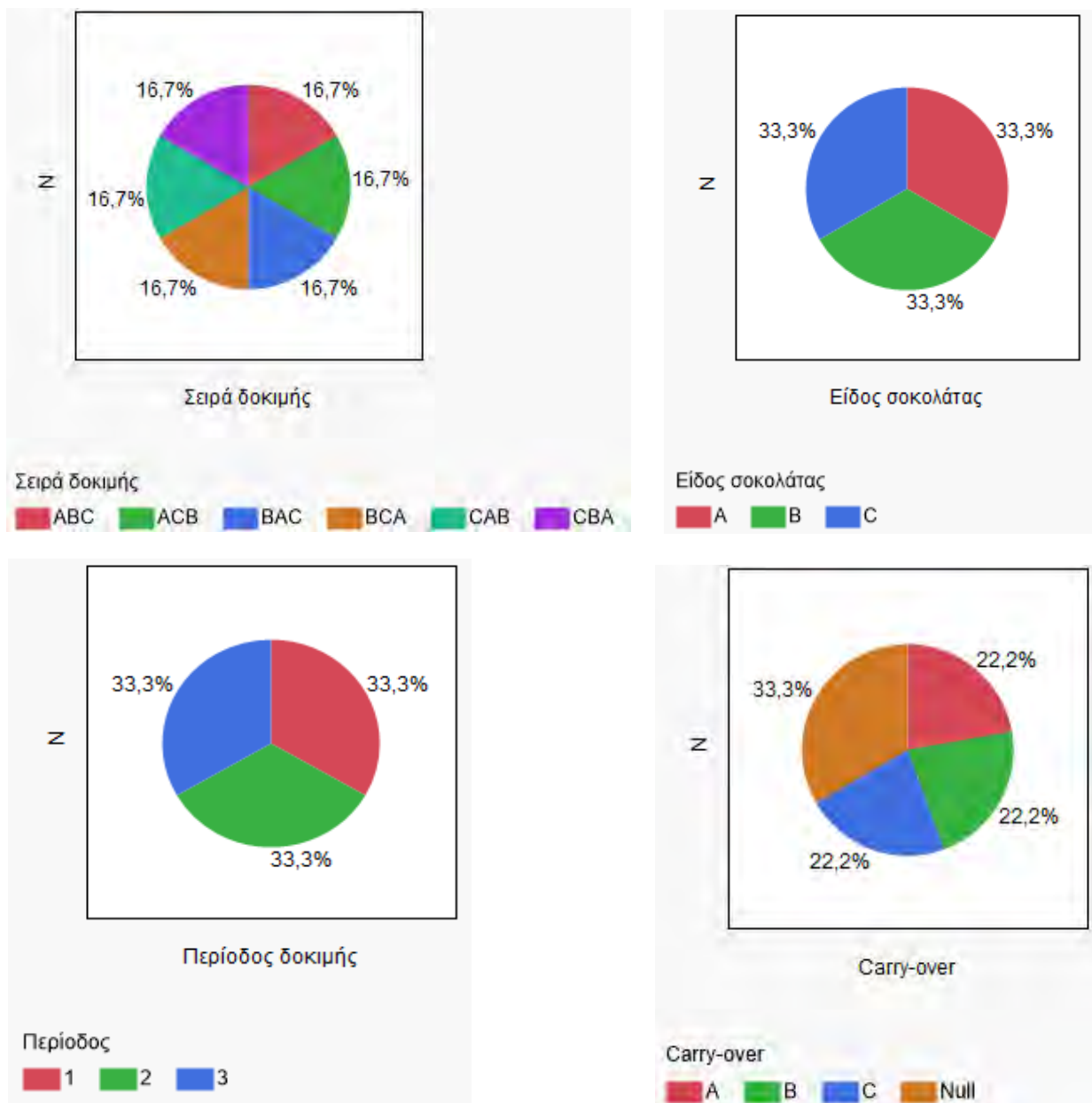
Όπως έχει ήδη αναφερθεί το μέγεθος του δείγματος είναι ικανοποιητικό, συνεπώς η έλλειψη κανονικότητας της μεταβλητής «προτίμηση» δεν αποτελεί αξεπέραστο εμπόδιο για περαιτέρω στατιστική ανάλυση.

4-6-2. Ανεξάρτητες μεταβλητές (Κατανομές)

Με τη βοήθεια του λογισμικού JMP ακολουθούν τέσσερα (4) γραφήματα των ανεξαρτήτων μεταβλητών του πειράματος. Υπενθυμίζεται ότι και οι 4 ανεξάρτητες μεταβλητές είναι κατηγορικές. Τα γραφήματα δίνονται σε κυκλικά διαγράμματα για να υπάρχει μια εποπτική εικόνα της κατανομής των τιμών ως προς τις κατηγορίες κάθε μεταβλητής.

Γράφημα 4-7

Κυκλικά διαγράμματα 4 ανεξάρτητων μεταβλητών



4-7. Testing τα στατιστικά λογισμικά

Πριν προχωρήσουμε στην ανάλυση που αφορά στο 3X3 crossover πείραμα γευσίγνωσίας, θα πρέπει να ελέγξουμε αν τα τέσσερα (4) στατιστικά λογισμικά που θα χρησιμοποιήσουμε δίνουν ίδια αποτελέσματα στη χρήση περισσότερο πολύπλοκων στατιστικών διαδικασιών. Με αυτό τον τρόπο γίνεται και έλεγχος των αρχείων δεδομένων που χρησιμοποιεί κάθε λογισμικό (μήπως κάποιο από τα αρχεία έχει υποστεί κάποια αλλοίωση).

Επιλέξαμε να «τρέξουμε» και με τα 4 λογισμικά ένα γενικό γραμμικό μοντέλο (general linear model) λαμβάνοντας ως εξαρτημένη μεταβλητή την προτίμηση (preference) και ως ανεξάρτητες - τις κατηγορικές μεταβλητές είδος σοκολάτας (treatment) και την περίοδο δοκιμής (period). Στον παρακάτω πίνακα δίνονται ενδεικτικά τα αποτελέσματα της ανάλυσης διακύμανσης (anova table) και για τα 4 λογισμικά. Το σύνολο της ανάλυσης για κάθε λογισμικό δίνεται στο παράρτημα. Σημειώνεται ότι για το λογισμικό R δίνεται μόνο ο κώδικας (για οικονομία χώρου) ο οποίος μπορεί πολύ εύκολα να αναπαραχθεί.

Πίνακας 4-8

F-ratios και p-values από Anova Table

Variable	Statistic	JMP	Minitab	R	SPSS
Treatment	F-Ratio	1.859	1.86	1.859	1.86
	p-value	0.158	0.159	0.158	0.159
Period	F-Ratio	4.339	4.34	4.339	4.34
	p-value	0.014	0.014	0.014	0.014

Παρατηρούμε ότι και τα 4 λογισμικά δίνουν ίδια αποτελέσματα. Επομένως μπορούμε να προχωρήσουμε στην ανάλυση των δεδομένων του 3X3 crossover πειράματος γευσίγνωσίας.

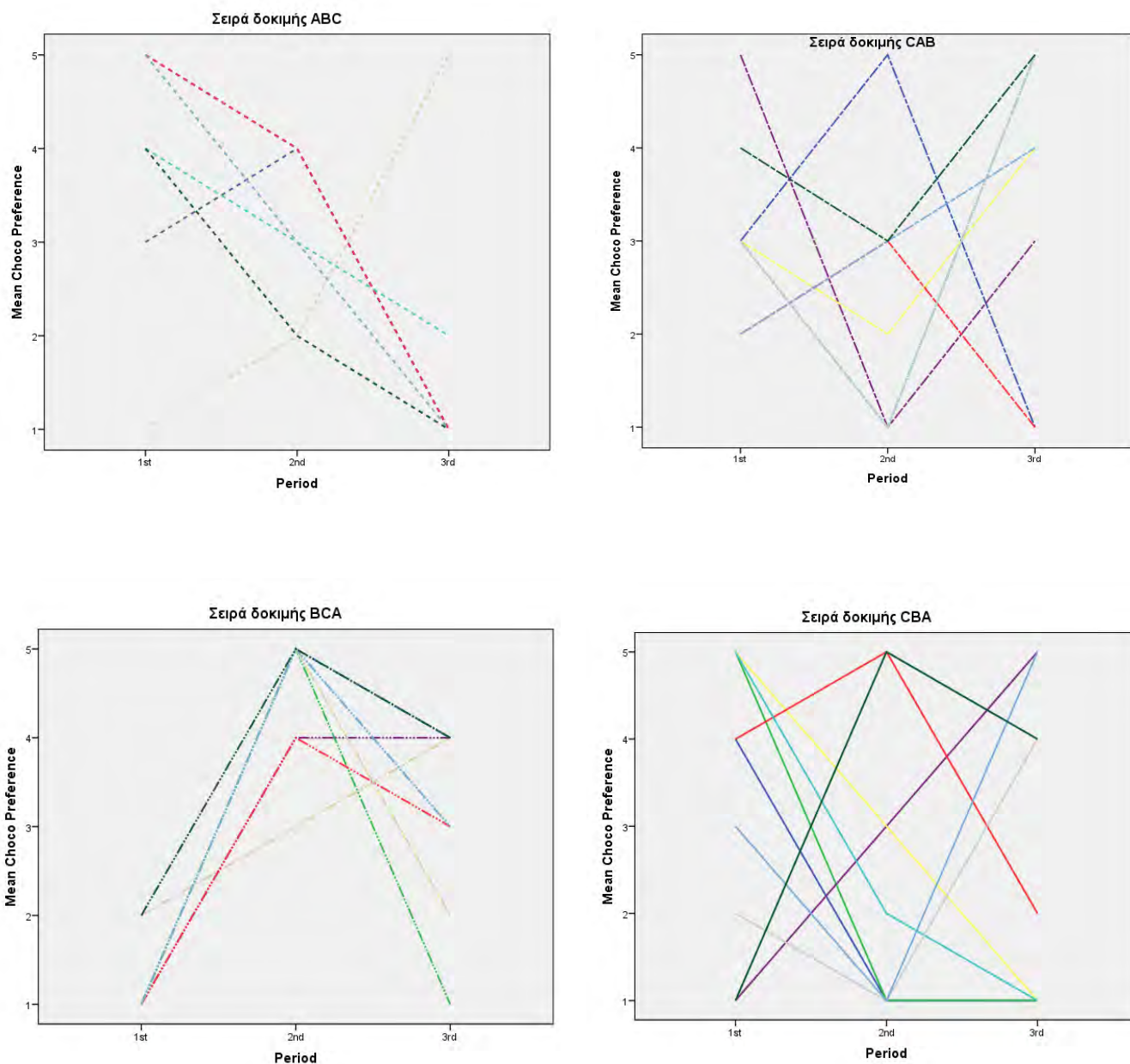
4-8. Χρήσιμα γραφήματα για crossover δοκιμές

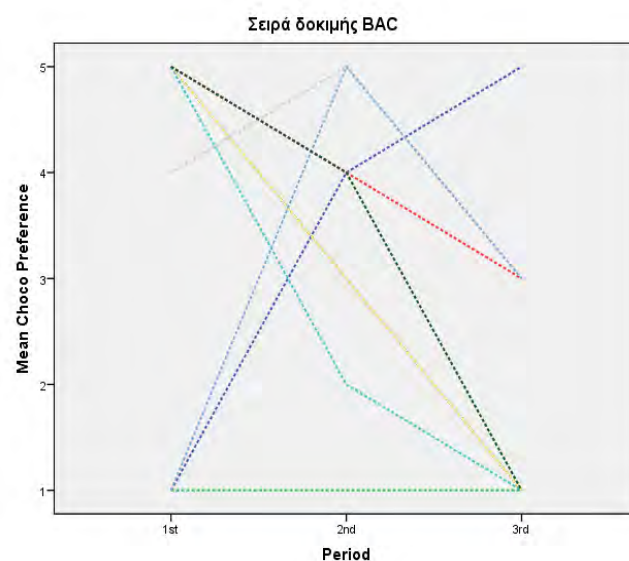
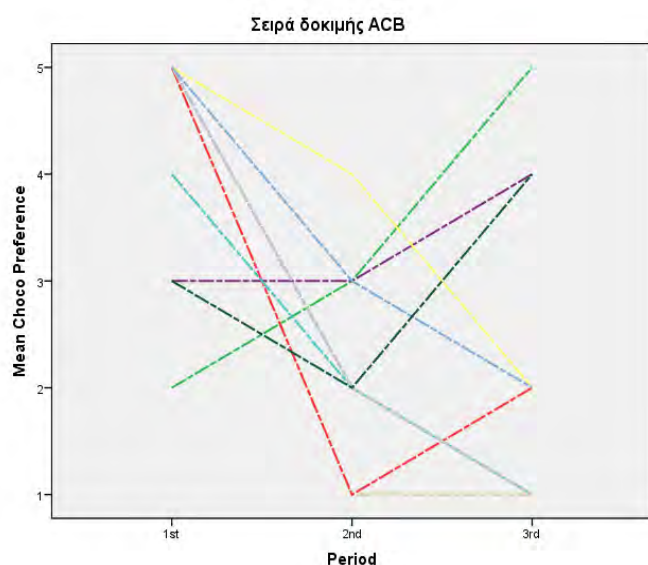
Είδαμε σε προηγούμενο κεφάλαιο ότι σε ένα crossover σχεδιασμό μπορούν να χρησιμοποιηθούν ορισμένα είδη γραφημάτων προκειμένου να γίνει μια πρώτη εκτίμηση των αδρών γνωρισμάτων των δεδομένων και της συμπεριφοράς των μεταβλητών. Σε 3X3 crossover πειράματα

χρησιμοποιούνται κυρίως δύο (2) είδη γραφημάτων, ήτοι τα profiles plots και τα means plots. Με τη βοήθεια του SPSS δίνονται τα profiles plots των βαθμολογιών ως προς την προτίμηση των γευσιγνωστών στις 3 περιόδους δοκιμής. Δίνονται συνολικά 6 profiles plots όσες και οι αλληλουχίες (σειρές) δοκιμής.

Γράφημα 4-8

Profiles plots γευσιγνωστών





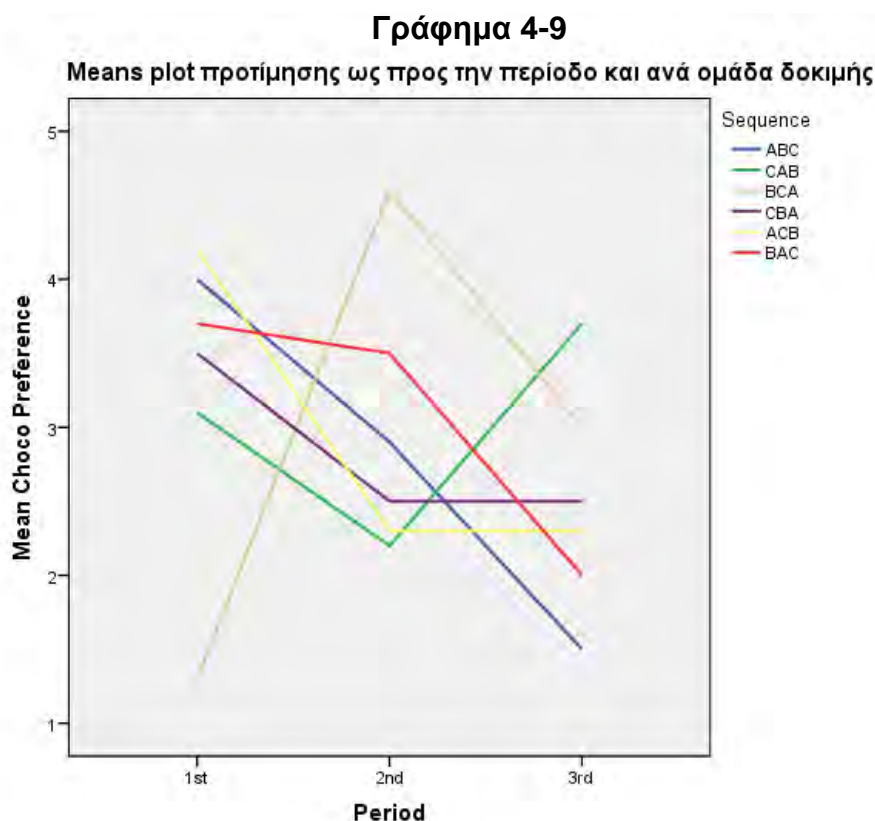
Σε καθένα από τα παραπάνω γραφήματα παριστάνονται οι βαθμολογίες ως προς την προτίμηση για τις τρεις χρονικές περιόδους για 10 γευσιγνώστες. Μια πρώτη εικόνα είναι ότι οι περισσότεροι γευσιγνώστες – ανεξάρτητα από την ομάδα δοκιμής στην οποία ανήκουν – φαίνεται να δίνουν την υψηλότερη βαθμολογία κατά την 1^η περίοδο. Εξαιρεση αποτελούν οι γευσιγνώστες που κατανεμήθηκαν στην σειρά δοκιμής BCA όπου σχεδόν στο σύνολό τους – κατά την 1^η περίοδο – δίνουν χαμηλές βαθμολογίες 1 ή 2. Στις σειρές δοκιμών CBA, ACB και BAC υπάρχει μεγάλη ανομοιομορφία στις βαθμολογίες των γευσιγνωστών.

Ακολουθεί πίνακας και γράφημα όπου αποτυπώνονται οι μέσες βαθμολογίες προτίμησης ως προς την περίοδο και ανά ομάδα δοκιμής.

Πίνακας 4-9

Μέσες βαθμολογίες προτίμησης ως προς την περίοδο και την ομάδα δοκιμής

			Period		
			1st	2nd	3rd
			Mean	Mean	Mean
Sequence	ABC	Choco Pref	4,00	2,90	1,50
	CAB	Choco Pref	3,10	2,20	3,70
	BCA	Choco Pref	1,30	4,60	3,00
	CBA	Choco Pref	3,50	2,50	2,50
	ACB	Choco Pref	4,20	2,30	2,30
	BAC	Choco Pref	3,70	3,50	2,00



Στην 1^η περίοδο οι μεγαλύτερες μέσες τιμές παρατηρούνται στις σειρές ACB ($m_{ACB} = 4.20$) και ABC ($m_{ABC} = 4.00$). Στη 2^η περίοδο η μεγαλύτερη μέση τιμή παρατηρείται στη σειρά BCA ($m_{BCA} = 4.60$). Στην 3^η περίοδο η μεγαλύτερη μέση τιμή παρατηρείται στη σειρά CAB ($m_{CAB} = 3.70$). Οι μέσες τιμές στις σειρές ABC, CBA, ACB και BAC ακολουθούν πτωτική πορεία από την 1^η στην 3^η περίοδο.

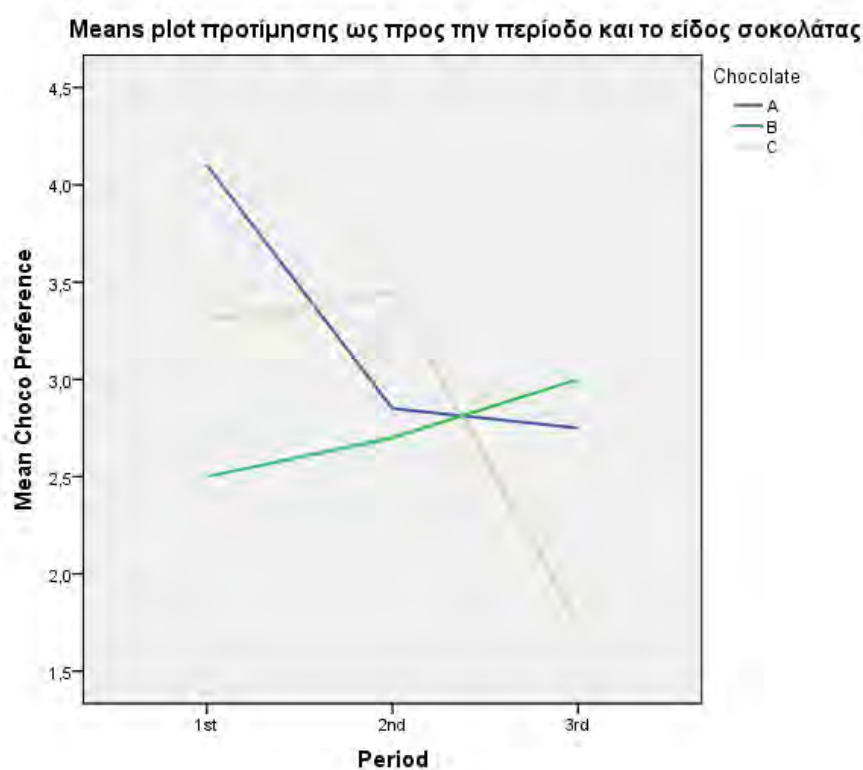
Ακολουθεί πίνακας και γράφημα όπου αποτυπώνονται οι μέσες βαθμολογίες προτίμησης ως προς την περίοδο και ανά σοκολάτα.

Πίνακας 4-10

Μέσες βαθμολογίες προτίμησης ως προς την περίοδο και την ομάδα δοκιμής

			Period		
			1st	2nd	3rd
			Mean	Mean	Mean
Chocolate	A	Choco Pref	4,10	2,85	2,75
	B	Choco Pref	2,50	2,70	3,00
	C	Choco Pref	3,30	3,45	1,75

Γράφημα 4-10



Στην 1^η περίοδο η σοκολάτα A έχει τη μεγαλύτερη μέση βαθμολογία ($m_A = 4.10$). Στη 2^η περίοδο η σοκολάτα C έχει τη μεγαλύτερη μέση βαθμολογία ($m_C = 3.45$). Στην 3^η περίοδο η σοκολάτα B έχει την υψηλότερη μέση τιμή ($m_B = 3.00$). Για τη σοκολάτα B παρατηρούμε ότι η μέση βαθμολογία της αυξάνεται από την 1^η περίοδο έως την 3^η.

73

4-9. Έλεγχος επίδρασης carry-over

Σε ένα crossover πείραμα αποτελεί πρώτη προτεραιότητα να εξεταστεί αν υφίσταται στατιστικά σημαντική επίδραση carry-over. Η ύπαρξη (ή μη) σημαντικής επίδρασης carry-over επηρεάζει αποφασιστικά τη μεθοδολογία που θα επιλεγεί για τη στατιστική ανάλυση των δεδομένων ενός crossover πειράματος.

Για τον έλεγχο της ύπαρξης επίδρασης carry-over στο test γευσίγνωσας σοκολάτας χρησιμοποιούμε μεικτό γραμμικό μοντέλο (mixed linear model). Η εξαρτημένη μεταβλητή «προτίμηση» που αποτελεί τη μεταβλητή ενδιαφέροντος λαμβάνεται ως συνεχής και η ύπαρξη αρκετά μεγάλου δείγματος μας καθυσιάζει σχετικά με την μη ύπαρξη κανονικότητας.

Στο συγκεκριμένο μεικτό μοντέλο ως μεταβλητή που προκαλεί σταθερές (fixed) επιδράσεις θεωρείται η επίδραση carry-over (ή residual) και ως μεταβλητή που προκαλεί – ενδεχομένως – τυχαίες (random) επιδράσεις ο A/A των γευσιγνωστών (id ή subject).

Άρα το μοντέλο έχει τη μορφή:

Προτίμηση ~ Επίδραση carry-over + id γευσιγνωστών + error

Η προσαρμογή του μεικτού μοντέλου στα δεδομένα και η στατιστική επεξεργασία πραγματοποιείται με 4 στατιστικά λογισμικά - το JMP, το Minitab, το R και το SPSS. Με τη βοήθεια του ακόλουθου πίνακα θα δούμε τον τρόπο με τον οποίο μπορούμε να προσεγγίσουμε τη χρήση μεικτού μοντέλου για καθένα από τα παραπάνω προγράμματα.

Πίνακας 4-11

Προσέγγιση μεικτού μοντέλου με 4 στατιστικά λογισμικά

Software command	Statistical softwares			
	JMP	Minitab	R	SPSS
Mixed Model	Menu Analyze>Fit Model>Model Specification>Pick Role Variables: Προτίμηση Construct Model Effects: Carry-over id (Random) EM: REML SS: Type III	Menu Stat>ANOVA>GLM>Fit GLM>Responses: Προτίμηση Factors: Carry-over id Random/Nest: Carry-over-Fixed id-Random EM: - SS: Type III	Κώδικας Packages: nlme, lme4 Functions: lme, lmer, lsmeans EM: REML SS: Type III	Menu ⁷⁴ Analyze>Mixed Models>Linear> Dependent: Προτίμηση Factor(s): Carry-over (Fixed) id (Random) EM: REML SS: Type III

Το JMP και το SPSS δίνουν τη δυνατότητα στον χρήστη για άμεση χρήση μεικτού μοντέλου - σε αντίθεση με το Minitab που επιτρέπει έμμεσα την προσαρμογή μεικτού μοντέλου στα δεδομένα. Το Minitab προσεγγίζει τη διαδικασία χρήσης μεικτού μοντέλου μέσω του γενικού γραμμικού μοντέλου

(general linear model) και δεν δίνει τη δυνατότητα επιλογής ακόμα και της μεθόδου εκτίμησης (REML ή ML) που προσφέρουν τα υπόλοιπα λογισμικά. Στο R - τα πακέτα nlme και lme4 - επιτρέπουν στο χρήστη να προσαρμόσει μεικτό μοντέλο στα δεδομένα με τη βοήθεια των συναρτήσεων lmer και lme.

Ως μέθοδος εκτίμησης από τα λογισμικά JMP, R και SPSS χρησιμοποιείται η Restricted Maximum Likelihood (REML). Το R χρησιμοποιεί by default τη μέθοδο Maximum Likelihood (ML) για αυτό χρειάζεται ιδιαίτερη προσοχή ώστε να γίνει η κατάλληλη αλλαγή. Η προσέγγιση που χρησιμοποιείται για τον υπολογισμό του αθροίσματος τετραγώνων SS είναι με Type III και για τα 4 λογισμικά.

Γενικά για το σύνολο των λογισμικών θα πρέπει να γίνεται έλεγχος των default ρυθμίσεων επειδή αυτές μπορούν να επηρεάσουν σημαντικά τα αποτελέσματα.

Πίνακας 4-12

Έλεγχος καλής προσαρμογής και επίδρασης carry-over

Category	Criterion	Statistical Softwares			
		JMP	Minitab	R	SPSS
Summary of fit	R^2	0.041	0.041	0.041	X
	$Adj.R^2$	0.025	0.025	X	X
	Root Mean Square Error	1.517	1.517	X	X
Goodness of fit tests	-2LogLik	664.31	X	X	661.54
	LogLik	X	X	-330.77	X
	AIC	678.80	X	675.54	665.54
	AICC	X	X	X	665.61
	CAIC	X	X	X	673.88
	BIC	695.47	X	697.73	671.88
Fixed effects tests for Carry-over	Df	3	3	3	3
	Df Den.	176	176	X	176
	F-Ratio	2.55	2.55	2.55	2.55
	p-value	0.057	0.057	0.058	0.057

Στον πίνακα 4-12 δίνονται από το JMP και το Minitab ο συντελεστής προσδιορισμού R^2 και ο προσαρμοσμένος συντελεστής προσδιορισμού $Adj.R^2$ (εξαιρετικά χαμηλές τιμές). Στο R το R^2 δίνεται με τη βοήθεια του πακέτου MuMIn και της συνάρτησης r.squaredGLMM. Τα συνήθη μέτρα καλής προσαρμογής του μοντέλου -2Log Likelihood, AIC και BIC δίνονται από το JMP, το R (έμμεσα το -2LogLik) και το SPSS με μικρές διαφοροποιήσεις

στις τιμές. Το Minitab δεν υπολογίζει κανένα μέτρο καλής προσαρμογής. Το SPSS δίνει δύο επιπλέον μέτρα για την καλή προσαρμογή του μοντέλου – το AICC και το CAIC. Χαμηλές τιμές για τα παραπάνω μέτρα είναι επιθυμητές.

Το σημαντικότερο αποτέλεσμα του παραπάνω πίνακα είναι ο έλεγχος των σταθερών (fixed) επιδράσεων για την επίδραση carry-over συνολικά. Όλα τα λογισμικά δίνουν ίδιο αποτέλεσμα ($p\text{-value} = 0.057$) πράγμα που σημαίνει ότι η επίδραση carry-over δεν είναι στατιστικά σημαντική. Επομένως, μπορούμε να προχωρήσουμε σε περαιτέρω στατιστική ανάλυση χρησιμοποιώντας τα δεδομένα και των 3 περιόδων.

Πίνακας 4-13

Εκτιμήσεις σταθερών (fixed) επιδράσεων

Category	Term	Criterion	Statistical Softwares			
			JMP	Minitab	R	SPSS
Estimates of fixed effects	Intercept	Estimate	2.88	3.3	3.3	3.3
		Std Error	0.114	0.196	0.195	0.195
		Df Den.	176	X	X	X
		t-Ratio	25.13	16.84	16.83	16.83
		p-value	<0.0001	0.000	0.000	0.000
		95% LL	2.66	2.91	2.91	2.91
		95% UL	3.11	3.68	3.68	3.68
	Carry-over (A)	Estimate	-0.16	-0.575	-0.575	-0.575
		Std Error	0.204	0.310	0.309	0.309
		Df Den.	176	X	X	X
		t-Ratio	-0.79	-1.86	-1.85	-1.85
		p-value	0.42	0.065	0.066	0.066
		95% LL	-0.56	-1.17	-1.17	-1.18
		95% UL	0.24	0.02	0.02	0.03
	Carry-over (B)	Estimate	0.13	-0.275	-0.275	-0.275
		Std Error	0.204	0.310	0.309	0.309
		Df Den.	176	X	X	X
		t-Ratio	0.67	-0.89	-0.88	-0.88
		p-value	0.50	0.376	0.376	0.376
		95% LL	-0.26	-0.87	-0.87	-0.88
		95% UL	0.54	0.32	0.32	0.33
	Carry-over (C)	Estimate	-0.38	-0.8	-0.8	-0.8
		Std Error	0.204	0.310	0.309	0.309
		Df Den.	176	X	X	X
t-Ratio		-1.89	-2.58	-2.58	-2.58	
p-value		0.06	0.011	0.011	0.011	
95% LL		-0.79	-1.40	-1.40	-1.41	
95% UL		0.01	-0.9	-0.19	-0.18	

Ως κατηγορία αναφοράς για τα 4 λογισμικά χρησιμοποιήθηκε η κατηγορία Null της μεταβλητής carry-over. Η επιλογή της κατηγορίας αναφοράς στο SPSS αποτελεί δύσκολη διαδικασία δεδομένου ότι απαιτήσε επανακωδικοποίηση της μεταβλητής carry-over. Στο R έγινε χρήση της εντολής `cho$residual<-relevel(cho$residual, ref="Null")`. Τα τρία λογισμικά δίνουν ίδια αποτελέσματα για τις κατηγορίες A, B και C της μεταβλητής carry-over. Μικρές διαφοροποιήσεις προκύπτουν στο λογισμικό JMP χωρίς όμως να μεταβάλλεται η ουσία των συμπερασμάτων. Η διαφοροποίηση αυτή φαίνεται ότι είναι καθαρά αλγοριθμική.

Πίνακας 4-14

Εκτιμήσεις διακύμανσης από μεταβλητές μοντέλου

Category	Variable	Criterion	Statistical Softwares			
			JMP	Minitab	R	SPSS
Variance component estimates	Γευσιγνώστες (id)	Var Comp.	0.00	-0.744	0.00	0.00
		Std Dev	X	0.00	0.00	X
		Std Error	0.00	0.00	X	X
		95% LL	0.00	X	X	X
		95% UL	0.00	X	X	X
		% of Total	0.00	0.00	X	X
	Residual	Var Comp.	2.304	3.04	2.304	2.304
		Std Dev	X	1.74	1.158	X
		Std Error	0.245	X	X	X
		95% LL	1.889	X	X	X
		95% UL	2.873	X	X	X
		% of Total	100.00	100.00	X	X

77

Έχει αναφερθεί ότι η μεταβλητή id γευσιγνωστών – ενδεχομένως – να προκαλεί τυχαίες (random) επιδράσεις στο μοντέλο. Ο πίνακας 4-14 δείχνει ότι η **μεταβλητή id γευσιγνωστών δεν έχει καμιά απολύτως επίδραση στο μοντέλο**. Το ίδιο αποτέλεσμα εξάγεται και από τα 4 λογισμικά. Ο δείκτης μάλιστα % of Total που μας δίνουν τα λογισμικά JMP και Minitab είναι ιδιαίτερα αποκαλυπτικός και μας πληροφορεί ότι η μεταβλητή id γευσιγνωστών ερμηνεύει το 0% της διακύμανσης της εξαρτημένης μεταβλητής «προτίμηση». Πολύ εύκολα θα μπορούσε να παραληφθεί και να ακολουθήσουμε διαφορετικό τρόπο στατιστικής προσέγγισης στα δεδομένα του πειράματος.

Ο πίνακας 4-15 που ακολουθεί δίνει τις μέσες τιμές στις 4 κατηγορίες της επίδρασης carry-over. Το λογισμικό Minitab υπολογίζει τις προσαρμοσμένες

μέσες τιμές (fitted values) για αυτό υπάρχει και μια μικρή διαφοροποίηση σε σχέση με τις τιμές που δίνουν τα υπόλοιπα 3 λογισμικά. Στα λογισμικά JMP, R και SPSS οι τιμές (ακόμα και στο 95% διάστημα εμπιστοσύνης) ταυτίζονται. Πρέπει να σημειωθεί ότι στο λογισμικό R έγινε χρήση του πακέτου lsmeans και της συνάρτησης lsmeans.

Πίνακας 4-15

Μέσες τιμές επίδρασης carry-over (κατηγορίες)

Category	Term	Criterion	Statistical Softwares			
			JMP	Minitab	R	SPSS
Least Squares Means	Carry-over (Null)	Mean	3.30	3.30	3.30	3.30
		Std Error	0.19	X	0.19	0.19
		95% LL	2.91	X	2.90	2.91
		95% UL	3.68	X	3.69	3.68
	Carry-over (A)	Mean	2.72	2.70	2.72	2.72
		Std Error	0.24	X	0.24	0.24
		95% LL	2.25	X	2.24	2.25
		95% UL	3.19	X	3.20	3.19
	Carry-over (B)	Mean	3.02	3.10	3.02	3.02
		Std Error	0.24	X	0.24	0.24
		95% LL	2.55	X	2.54	2.55
		95% UL	3.49	X	3.50	3.49
	Carry-over (C)	Mean	2.50	2.45	2.50	2.50
		Std Error	0.24	X	0.24	0.24
		95% LL	2.02	X	2.01	2.02
		95% UL	2.97	X	2.98	2.97

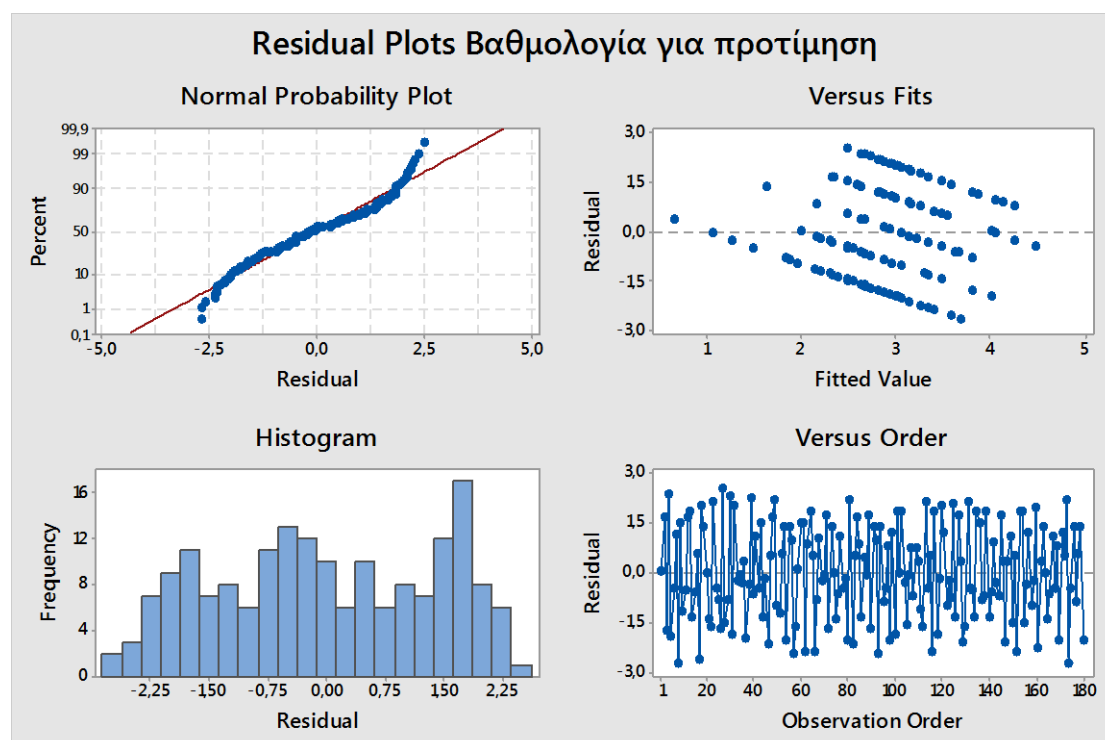
78

Ο έλεγχος της καταλληλότητας του μοντέλου πραγματοποιείται με τη χρήση γραφημάτων τα οποία δίνονται και με τα 4 λογισμικά. Το JMP με τη βοήθεια της επιλογής Effect Leverage δίνει διάφορα γραφήματα όπως αυτό των τιμών της «προτίμησης» ως προς τις εκτιμώμενες τιμές, των υπολοίπων ως προς τις εκτιμώμενες τιμές, των υπολοίπων ως προς τη σειρά A/A κλπ. Το SPSS δίνει τη δυνατότητα αποθήκευσης των υπολοίπων και των εκτιμωμένων τιμών και είναι εύκολη η κατασκευή διαγραμμάτων διασποράς. Στο R ο υπολογισμός των υπολοίπων και των εκτιμωμένων τιμών είναι σχετικά απλή διαδικασία όπως και η δημιουργία γραφημάτων διασποράς. Όλα τα παραπάνω καθώς και ο κώδικας για το R δίνονται στο παράρτημα της παρούσας εργασίας.

Ακολουθως παραθέτουμε τα γραφήματα που δίνονται by default από το λογισμικό Minitab. Είναι φανερό ότι δεν πληρούνται οι προϋποθέσεις για να

χαρακτηρίσουμε το μοντέλο που έχουμε προσαρμόσει στα δεδομένα μας κατάλληλο.

Γράφημα 4-11



4-10. Πλήρες μοντέλο

Στην προηγούμενη παράγραφο είδαμε – με τη βοήθεια 4 στατιστικών λογισμικών - ότι η επίδραση carry-over δεν είναι στατιστικά σημαντική. Το αποτέλεσμα αυτό μας επιτρέπει να προχωρήσουμε περαιτέρω στη στατιστική ανάλυση των δεδομένων του 3X3 crossover πειράματος προσαρμόζοντας ένα περισσότερο πολύπλοκο μοντέλο. Το μεικτό γραμμικό μοντέλο (mixed linear model) που χρησιμοποιούμε περιλαμβάνει ως εξαρτημένη μεταβλητή την βαθμολογία της προτίμησης (preference) και ως ανεξάρτητες την σειρά δοκιμής σοκολάτας (sequence), το είδος σοκολάτας (treatment), την περίοδο δοκιμής (period), την αλληλεπίδραση είδους σοκολάτας με την περίοδο δοκιμής (treatment by period interaction) και το id γευσιγνωστών.

Στο μοντέλο **δεν περιλαμβάνεται η επίδραση carry-over** αφενός επειδή **δεν είναι στατιστικά σημαντική** για τις τιμές της «προτίμησης» και αφετέρου επειδή έχει **μικρό αριθμό βαθμών ελευθερίας (df)**.

Επομένως το μεικτό μοντέλο θα έχει τη μορφή:

Προτίμηση ~ Σειρά δοκιμής + Είδος σοκολάτας + Περίοδος δοκιμής + Αλληλεπίδραση σοκολάτας*περιόδου + id γευσιγνωστών + error

Στο παραπάνω μοντέλο οι μεταβλητές προκαλούν τις ακόλουθες επιδράσεις:

- 1) Σειρά δοκιμής – Σταθερές επιδράσεις (Fixed)
- 2) Είδος σοκολάτας – Σταθερές επιδράσεις (Fixed)
- 3) Περίοδος δοκιμής – Σταθερές επιδράσεις (Fixed)
- 4) id γευσιγνωστών – Τυχαίες επιδράσεις (Random)

Για την προσαρμογή του μεικτού μοντέλου χρησιμοποιήθηκαν 4 στατιστικά λογισμικά – το JMP, το Minitab, το R και το SPSS όπως φαίνεται και στον πίνακα 4-16.

Πίνακας 4-16

Προσέγγιση πλήρους μεικτού μοντέλου με 4 στατιστικά λογισμικά

Software command	Statistical softwares			
	JMP	Minitab	R	SPSS
Mixed Model	Menu Analyze>Fit Model>Model Specification>Pick Role Variables: Προτίμηση Construct Model Effects: Σειρά δοκιμής Σοκολάτα Περίοδος Αλληλεπίδραση Σοκολάτας*Περίοδος id (Random) EM: REML SS: Type III	Menu Stat>ANOVA>GLM>Fit GLM>Responses: Προτίμηση Factors: Carry-over id Random/Nest: Σειρά δοκιμής Σοκολάτα Περίοδος Αλληλεπίδραση Σοκολάτας*Περίοδος id-Random EM: - SS: Type III	Κώδικας Packages: nlme, lme4 Functions: lme, lmer, lsmeans EM: REML SS: Type III	Menu Analyze>Mixed Models>Linear Dependent: Προτίμηση Factor(s): Σειρά δοκιμής Σοκολάτα Περίοδος Αλληλεπίδραση Σοκολάτας*Περίοδος id (Random) EM: REML SS: Type III

Επειδή για το λογισμικό R χρησιμοποιείται η ML ως μέθοδος εκτίμησης by default θα πρέπει να γραφεί η κατάλληλη εντολή για να γίνει χρήση της

μεθόδου REML. Το ίδιο ισχύει και για τον υπολογισμό των αθροισμάτων τετραγώνων (SS) όπου το R χρησιμοποιεί by default τον τύπο I (Type I). Για να γίνει χρήση του τύπου III (Type III) στο R χρησιμοποιούμε την εντολή *anova.lme(mixF,type="marginal",adjustSigma=F)*. Το JMP χρησιμοποιεί για τον υπολογισμό των αθροισμάτων τετραγώνων (SS) by default μόνο τον τύπο III (Type III). Τα λογισμικά Minitab και SPSS δίνουν τη δυνατότητα προσέγγισης τόσο με τον τύπο I (Type I) όσο και με τον τύπο III (Type III).

Η επιλογή της προσέγγισης μέσω του τύπου III (Type III) για τον υπολογισμό των αθροισμάτων τετραγώνων δεν είναι τυχαία. Το συγκεκριμένο 3X3 πείραμα γευσιγνωσίας σοκολάτας που μελετάμε με τη χρήση του παραπάνω μεικτού μοντέλου αποτελεί ένα πολύπλοκο σχεδιασμό. Ο σχεδιασμός αυτός είναι παραγοντικός (Factorial) και δεν είναι ισορροπημένος (unbalanced). Είναι γνωστό ότι ο υπολογισμός των αθροισμάτων τετραγώνων (SS) μέσω του τύπου I (Type I) δεν είναι ενδεδειγμένος. Για αυτό το λόγο γίνεται – αυστηρά - χρήση της προσέγγισης μέσω τύπου III (Type III).

Πίνακας 4-17

Έλεγχος καλής προσαρμογής μεικτού μοντέλου

Category	Criterion	Statistical Softwares			
		JMP	Minitab	R	SPSS
Summary of fit	R^2	0.283	0.283	0.267	X
	$Adj.R^2$	0.226	0.226	X	X
	Root Mean Square Error	1.351	1.351	X	X
Goodness of fit tests	-2LogLik	627.37	X	X	610,61
	LogLik	X	X	-305.30	X
	AIC	662.71	X	644.61	614,61
	AICC	X	X	X	614,68
	CAIC	X	X	X	622,83
	BIC	710.46	X	697.51	620,83

Στον πίνακα 4-17 παρατηρούμε ότι το R και το SPSS δίνουν σχεδόν ίδια τιμή για το μέτρο ελέγχου καλής προσαρμογής του μοντέλου $-2\text{LogLik}=610.6$. Τα λογισμικά JMP και Minitab υπολογίζουν και τον προσαρμοσμένο συντελεστή προσδιορισμού $Adj.R^2=0.22$ που έχει αρκετά χαμηλή τιμή. Τούτο πρακτικά σημαίνει ότι μόλις 22% της μεταβλητότητας της εξαρτημένης μεταβλητής «προτίμηση» ερμηνεύεται από το μοντέλο που έχουμε προσαρμόσει.

Πίνακας 4-18

Σταθερές επιδράσεις επεξηγηματικών μεταβλητών

Category	Variable	Criterion	Statistical Softwares			
			JMP	Minitab	R	SPSS
Fixed effects tests	Sequence	df	5	5	5	5
		F-Ratio	5.55	5.56	5.55	5.55
		p-value	<0.0001	0.000	<0.0003	0.000
	Treatment	df	2	2	2	2
		F-Ratio	2.29	10.88	10.88	2.29
		p-value	0.1037	0.000	<0.0001	0.104
	Period	df	2	2	2	2
		F-Ratio	5.36	15.26	15.26	5.36
		p-value	0.0055	0.000	<0.0001	0.006
	Treat*Period	df	4	4	4	4
		F-Ratio	12.34	12.34	12.34	12.34
		p-value	<0.0001	0.000	<0.0001	0.000

Από τον έλεγχο των σταθερών (fixed) επιδράσεων των ανεξάρτητων μεταβλητών του μοντέλου και της αλληλεπίδρασης παρατηρούμε ότι τα λογισμικά JMP και SPSS δίνουν ίδια αποτελέσματα. Συγκεκριμένα η μεταβλητή είδος σοκολάτας (treatment) δεν είναι στατιστικά σημαντική για το μοντέλο ($p\text{-value} = 0.104$). Αντίθετα οι μεταβλητές σειρά δοκιμής (sequence) με $p\text{-value} < 0.001$, περίοδος (period) με $p\text{-value} = 0.006$ και η αλληλεπίδραση σοκολάτας και περιόδου (treatment by period interaction) με $p\text{-value} < 0.001$ είναι στατιστικά σημαντικές για το μοντέλο.

Ομοίως τα λογισμικά Minitab και R δίνουν τα ίδια αποτελέσματα σε ότι αφορά στις σταθερές (Fixed) επιδράσεις των μεταβλητών του μοντέλου. Όλες οι ανεξάρτητες μεταβλητές (Sequence, Treatment και Period) είναι στατιστικά σημαντικές. Το ίδιο ισχύει και για την αλληλεπίδραση σοκολάτας και περιόδου ($p\text{-value} < 0.001$).

Επομένως, υπάρχει μια σοβαρή διαφοροποίηση μεταξύ των 4 λογισμικών σε σχέση με τη μεταβλητή «Είδος σοκολάτας». Τα λογισμικά JMP και SPSS δίνουν ότι το είδος σοκολάτας δεν είναι στατιστικά σημαντικό αποτέλεσμα για το μοντέλο. Από την άλλη μεριά τα λογισμικά Minitab και R δίνουν στατιστικά σημαντικό αποτέλεσμα για τη συγκεκριμένη μεταβλητή.

Πρέπει να τονιστεί - για ακόμη μια φορά - ότι η προσέγγιση των δεδομένων του πειράματος και η προσαρμογή μεικτού μοντέλου σε αυτά

πραγματοποιήθηκε για τα 3 λογισμικά με τη χρήση της μεθόδου REML (ως μεθόδου εκτίμησης). Για το Minitab δεν μπορούμε να γνωρίζουμε με ασφάλεια. Επιπρόσθετα, για τον υπολογισμό των αθροισμάτων των τετραγώνων (SS) χρησιμοποιήθηκε η προσέγγιση τύπου III (Type III).

Η προσέγγιση με τη χρήση τύπου I (Type I) δίνει ακριβώς ίδια αποτελέσματα – σε ότι αφορά τις μεταβλητές του μεικτού μοντέλου – για τα λογισμικά Minitab, R και SPSS. Σε αυτή την περίπτωση οι μεταβλητές σειρά δοκιμής (sequence) και είδος σοκολάτας (treatment) δεν είναι στατιστικά σημαντικές ενώ η μεταβλητή περίοδος (period) και η αλληλεπίδραση είναι. Το λογισμικό JMP δεν δίνει τη δυνατότητα επιλογής τύπου I προκειμένου να «τρέξουμε» το μεικτό μοντέλο. **Όπως έχει αναφερθεί η συγκεκριμένη προσέγγιση είναι λανθασμένη.**

Η διαφοροποίηση του αποτελέσματος ανάμεσα στην 1^η ομάδα λογισμικών (JMP και SPSS) με τη 2^η ομάδα (Minitab και το R) - σε ότι αφορά στη μεταβλητή «είδος σοκολάτας» - πιθανότατα να οφείλεται στο διαφορετικό τρόπο με τον οποίο διαχειρίζονται τα παραπάνω λογισμικά μη ισορροπημένα (unbalanced) δεδομένα. Επομένως, αυτή η διαφορά είναι μάλλον αλγοριθμική.

Πίνακας 4-19

Εκτιμήσεις διακύμανσης από μεταβλητές μοντέλου

Category	Variable	Criterion	Statistical Softwares			
			JMP	Minitab	R	SPSS
Variance component estimates	Γευσιγνώστες (id)	Var Comp.	0.00	0.00	0.00	0.00
		Std Dev	X	0.00	0.00	X
		Std Error	0.00	0.00	X	X
		95% LL	0.00	X	X	X
		95% UL	0.00	X	X	Xe
		% of Total	0.00	0.00	X	X
	Residual	Var Comp.	1.82	1.82	1.82	1,82
		Std Dev	X	1.35	1.35	X
		Std Error	0.2	X	X	X
		95% LL	1.49	X	X	X
		95% UL	2.29	X	X	X
		% of Total	100.00	100.00	X	X

Η μεταβλητή id γευσιγνωστών δεν έχει καμιά επίδραση στο μοντέλο. Η συγκεκριμένη μεταβλητή δεν προκαλεί τυχαίες (random) επιδράσεις στο μοντέλο δεδομένου ότι ερμηνεύει απειροελάχιστο ποσοστό (πρακτικά 0%) της

μεταβλητότητας της εξαρτημένης μεταβλητής «προτίμηση». Θα μπορούσε να μην συμπεριληφθεί στο μοντέλο και να γίνει προσέγγιση με τη βοήθεια γενικού γραμμικού μοντέλου (general linear model). **Τα αποτελέσματα που προκύπτουν με τη εφαρμογή GLM με τη χρήση των 4 λογισμικών είναι ακριβώς ίδια με αυτά του πίνακα 4-17!**

Πριν προχωρήσουμε στην εκτίμηση των παραμέτρων του μοντέλου πρέπει να σημειωθεί ότι για τη μεταβλητή «σειρά δοκιμής» ως κατηγορία αναφοράς λαμβάνεται η αλληλουχία BAC, για τη μεταβλητή «είδος σοκολάτας» η C και για τη μεταβλητή «περίοδος» η 3^η.

Αρχικά, θα σχολιάσουμε τις εκτιμήσεις παραμέτρων για τις μεταβλητές που είναι στατιστικά σημαντικές για το μοντέλο όπως αυτές δίνονται από τα λογισμικά JMP και SPSS.

Σχήμα 4-1

Εκτιμήσεις παραμέτρων (JMP και SPSS)

Parameter Estimates					
Term	Estimate	Std Error	DFDen	t Ratio	Prob> t
Intercept	2,93	0,10	166	29,11	0,000
sequence[ABC]	-0,09	0,30	166	-0,29	0,769
sequence[ACB]	-1,18	0,30	166	-3,90	0,000
sequence[BCA]	0,14	0,30	166	0,48	0,633
sequence[CAB]	-0,09	0,30	166	-0,29	0,769
sequence[CBA]	-0,18	0,30	166	-0,59	0,557
treatment[A]	0,30	0,14	166	2,11	0,037
treatment[B]	-0,20	0,14	166	-1,40	0,162
period[1]	0,37	0,14	166	2,57	0,011
period[2]	0,07	0,14	166	0,47	0,641
treatment[A]*period[1]	1,13	0,29	166	3,98	0,000
treatment[A]*period[2]	-1,10	0,29	166	-3,86	0,000
treatment[B]*period[1]	-1,37	0,29	166	-4,79	0,000
treatment[B]*period[2]	0,03	0,29	166	0,12	0,907

Παρατηρούμε ότι η αλληλουχία ACB είναι σημαντική και προκαλεί στην «προτίμηση» μείωση 1.18. Ομοίως η 1^η περίοδος είναι σημαντική και προκαλεί μικρή αύξηση στην «προτίμηση» 0.37. Η αλληλεπίδραση σοκολάτας A με 1^η περίοδο προκαλούν αύξηση στην «προτίμηση» 1.13. Οι αλληλεπιδράσεις σοκολάτας A και 2^{ης} περιόδου καθώς και σοκολάτας B και 1^{ης} περιόδου προκαλούν μείωση στην «προτίμηση» 1.1 και 1.37 αντίστοιχα

84

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	2,49	,38	166,000	6,601	,000	1,74	3,23
[sequence=ABC]	-1,48	,45	166	-3,279	,001	-2,37	-,59
[sequence=CAB]	-1,48	,45	166	-3,279	,001	-2,37	-,59
[sequence=BCA]	-1,24	,45	166	-2,761	,006	-2,13	-,35
[sequence=CBA]	-1,57	,49	166	-3,173	,002	-2,54	-,59
[sequence=ACB]	-2,57	,49	166	-5,199	,000	-3,54	-1,59
[period=1st]	2,33	,55	166,000	4,227	,000	1,24	3,42
[period=2nd]	2,87	,55	166,000	5,194	,000	1,78	3,96
[treatment=A] * [period=1]	-,37	,86	166	-,429	,669	-2,05	1,32
[treatment=A] * [period=2]	-3,43	,86	166	-4,015	,000	-5,12	-1,75
[treatment=B] * [period=1]	-4,23	,86	166	-4,951	,000	-5,92	-2,55
[treatment=B] * [period=2]	-3,67	,86	166	-4,288	,000	-5,35	-1,98

a. Dependent Variable: Choco Preference.

Στο SPSS παρατηρούμε ότι το σύνολο των κατηγοριών (εκτός από μία) επηρεάζουν σημαντικά την «προτίμηση». Ενδεικτικά αναφέρουμε ότι η αλληλουχία ACB προκαλεί στην «προτίμηση» μείωση 2.57. Τόσο η 1^η περίοδος όσο και η 2^η προκαλούν αύξηση στην «προτίμηση» 2.33 και 2.87 αντίστοιχα.

Σχήμα 4-2

Εκτιμήσεις παραμέτρων (R και Minitab)

	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.488889	0.3770487	112	6.600975	0.0000
as.factor(sequence)ABC	-1.477778	0.4506594	54	-3.279146	0.0018
as.factor(sequence)ACB	-2.566667	0.4936726	54	-5.199127	0.0000
as.factor(sequence)BCA	-1.244444	0.4506594	54	-2.761386	0.0078
as.factor(sequence)CAB	-1.477778	0.4506594	54	-3.279146	0.0018
as.factor(sequence)CBA	-1.566667	0.4936726	54	-3.173493	0.0025
treatmentA	1.666667	0.5519427	112	3.019637	0.0031
treatmentB	2.533333	0.5519427	112	4.589848	0.0000
period1st	2.333333	0.5519427	112	4.227492	0.0000
period2nd	2.866667	0.5519427	112	5.193776	0.0000
treatmentA:period1st	-0.366667	0.8550660	112	-0.428817	0.6689
treatmentB:period1st	-4.233333	0.8550660	112	-4.950885	0.0000
treatmentA:period2nd	-3.433333	0.8550660	112	-4.015285	0.0001
treatmentB:period2nd	-3.666667	0.8550660	112	-4.288168	0.0000

Στο R παρατηρούμε ότι μόνο μια κατηγορία (αλληλεπίδραση σοκολάτας A με 1^η περίοδο) δεν είναι σημαντική για το μοντέλο. Αύξηση στην «προτίμηση» κατά 1.6 και 2.5 αντίστοιχα προκαλούν οι σοκολάτες A και B.

Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant	2,489	0,377	6,60	0,000
sequence				
ABC	-1,478	0,451	-3,28	0,001
ACB	-2,567	0,494	-5,20	0,000
BCA	-1,244	0,451	-2,76	0,006
CAB	-1,478	0,451	-3,28	0,001
CBA	-1,567	0,494	-3,17	0,002
treatment				
A	1,667	0,552	3,02	0,003
B	2,533	0,552	4,59	0,000
period				
1st	2,333	0,552	4,23	0,000
2nd	2,867	0,552	5,19	0,000
treatment*period				
A 1st	-0,367	0,855	-0,43	0,669
A 2nd	-3,433	0,855	-4,02	0,000
B 1st	-4,233	0,855	-4,95	0,000
B 2nd	-3,667	0,855	-4,29	0,000

Τα αποτελέσματα των εκτιμήσεων των παραμέτρων που δίνει το Minitab είναι ακριβώς ίδια με αυτά που δίνει και το R. Το συγκεκριμένο λογισμικό (όπως και το R) δίνει στατιστικά σημαντικές τις σοκολάτες A και B.

Ένα παράδοξο που παρατηρούμε είναι ότι οι εκτιμήσεις των παραμέτρων που δίνει το λογισμικό SPSS ταυτίζονται με αυτές που δίνουν τα λογισμικά JMP και Minitab. Από όλα τα παραπάνω είναι προφανές ότι χρειαζόμαστε – ενδεχομένως – περισσότερες επαναλήψεις του πειράματος και να ελέγξουμε εκ νέου τα αποτελέσματα για να διαπιστώσουμε τί ακριβώς ισχύει. Όταν λέμε περισσότερες επαναλήψεις του πειράματος εννοούμε πρακτικά ότι περισσότεροι γευσιγνώστες πρέπει να λάβουν μέρος στο test γευσιγνωσίας, δηλαδή χρειαζόμαστε μεγαλύτερο δείγμα.

Σε ότι αφορά τις μέσες βαθμολογίες ως προς την προτίμηση για τις κατηγορίες κάθε μιας από τις μεταβλητές που μετέχουν στο μοντέλο υπάρχει ταύτιση στα αποτελέσματα και για τα 4 λογισμικά. Για κάθε μεταβλητή δίνονται

οι μέσες βαθμολογίες της προτίμησης με τη χρήση διαφορετικού – κάθε φορά – λογισμικού.

Σχήμα 4-3

Least Square Means μεταβλητών

```

sequence  lsmean      SE df lower.CL upper.CL
BAC       4.322222  0.3186643 59 3.684576 4.959868
ABC       2.844444  0.3186643 54 2.205561 3.483328
ACB       1.755556  0.3186643 54 1.116672 2.394439
BCA       3.077778  0.3186643 54 2.438894 3.716661
CAB       2.844444  0.3186643 54 2.205561 3.483328
CBA       2.755556  0.3186643 54 2.116672 3.394439
    
```

```

treatment
A         3,233    0,175
B         2,733    0,175
C         2,833    0,175
    
```

Least Squares Means Table

Level	Sq Mean	Std Error
A,1	4,73	0,36
A,2	2,20	0,36
A,3	2,77	0,36
B,1	1,73	0,36
B,2	2,83	0,36
B,3	3,63	0,36
C,1	3,43	0,36
C,2	3,97	0,36
C,3	1,10	0,36

Estimates^a

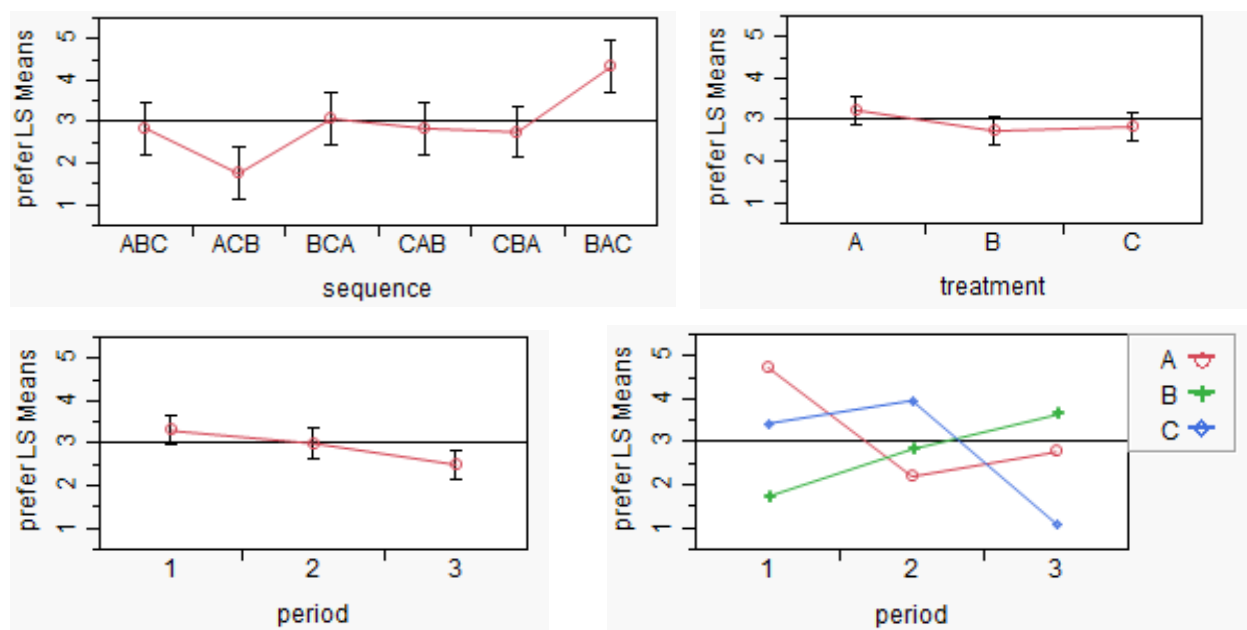
Period	Mean	Std. Error	df	95% Confidence Interval	
				Lower Bound	Upper Bound
1st	3,300	,175	166	2,955	3,645
2nd	3,000	,175	166	2,655	3,345
3rd	2,500	,175	166	2,155	2,845

a. Dependent Variable: Choco Preference.

Στη συνέχεια ακολουθούν τα 4 γραφήματα μέσω των τιμών για τις παραπάνω μεταβλητές όπως δίνονται από το JMP.

Γράφημα 4-12

Least Squares Means Plot



4-11. Συμπέρασμα

Η εφαρμογή των 4 στατιστικών λογισμικών – JMP, Minitab, R και SPSS – στο 3X3 crossover πείραμα γευσιγνωσίας σοκολάτας έδωσε στις περισσότερες περιπτώσεις ίδια αποτελέσματα και σε κάποιες διαφορετικά. Ο πίνακας που ακολουθεί αποτυπώνει συγκεντρωτικά τις ομοιότητες και τις διαφορές μεταξύ των 4 στατιστικών πακέτων σε ότι αφορά στην προσέγγιση που χρησιμοποιεί το καθένα για ένα crossover πείραμα. Τα παρακάτω συμπεράσματα αναφέρονται στο πλήρες μοντέλο.

Πίνακας 4-20

Σύγκριση JMP, Minitab, R και SPSS

Statistical softwares		JMP	Minitab	R	SPSS
Μοντελοποίηση					
Μοντέλο		Mixed	Mixed (Μέσω GLM)	Mixed	Mixed
Προσέγγιση		Άμεση	Έμμεση	Άμεση	Άμεση
Μέθοδος Εκτίμησης		REML (Default)	Άγνωστη	REML (Default είναι η ML)	REML (Default)
Υπολογισμός SS		Type III (Default-Δεν δίνει τη δυνατότητα για χρήση Type I)	Type III (Default)	Type III (Χρήση συνάρτησης anova.lme)	Type III
Αποτελέσματα					
Goodness of fit	Tests	-2LogLik, AIC, BIC	X	LogLik, AIC, BIC	-2LogLik, AIC, AICC, CAIC, BIC
	Results	Παρόμοια	X	Παρόμοια	Παρόμοια
Fixed effects tests		Ίδια με SPSS	Ίδια με R	Ίδια με Minitab	Ίδια με JMP
Variance components estimates		Ίδια	Ίδια	Ίδια	Ίδια
Estimates of parameters		Αρκετές διαφορές	Ίδια με R και SPSS	Ίδια με Minitab και SPSS	Ίδια με Minitab και R
Least squares means		Ίδια	Ίδια	Ίδια	Ίδια

Statistical softwares	JMP	Minitab	R	SPSS
Γραφικές παραστάσεις για mixed models				
Ποικιλία	*****	***	*****	***
Ευκολία κατασκευής	*****	****	*	****
Ποιότητα	****	****	***	**
Γενική Αξιολόγηση				
Αποτελέσματα (output)	*****	**	***	***
Αρχείο δεδομένων	*****	***	****	*****
Δυνατότητες	*****	***	*****	***
Κόστος	*	**	*****	**
Στατιστικές μέθοδοι	*****	***	*****	***
Φιλικό για χρήστη	*****	****	*	*****

Τα αποτελέσματα για το πλήρες μεικτό μοντέλο παρουσιάστηκαν και αναλύθηκαν διεξοδικά στις προηγούμενες παραγράφους.

Στο τελευταίο μέρος του πίνακα 4-19 με τον τίτλο «Γενική αξιολόγηση» γίνεται μια προσπάθεια εκτίμησης – η οποία δεν μπορεί να είναι εντελώς αντικειμενική – των 4 λογισμικών ως προς διάφορα χαρακτηριστικά.

Το JMP και το R αποτελούν – σε κάθε περίπτωση – πανίσχυρα στατιστικά λογισμικά με πολύ μεγάλες δυνατότητες. Το βασικό μειονέκτημα του JMP είναι το κόστος του.

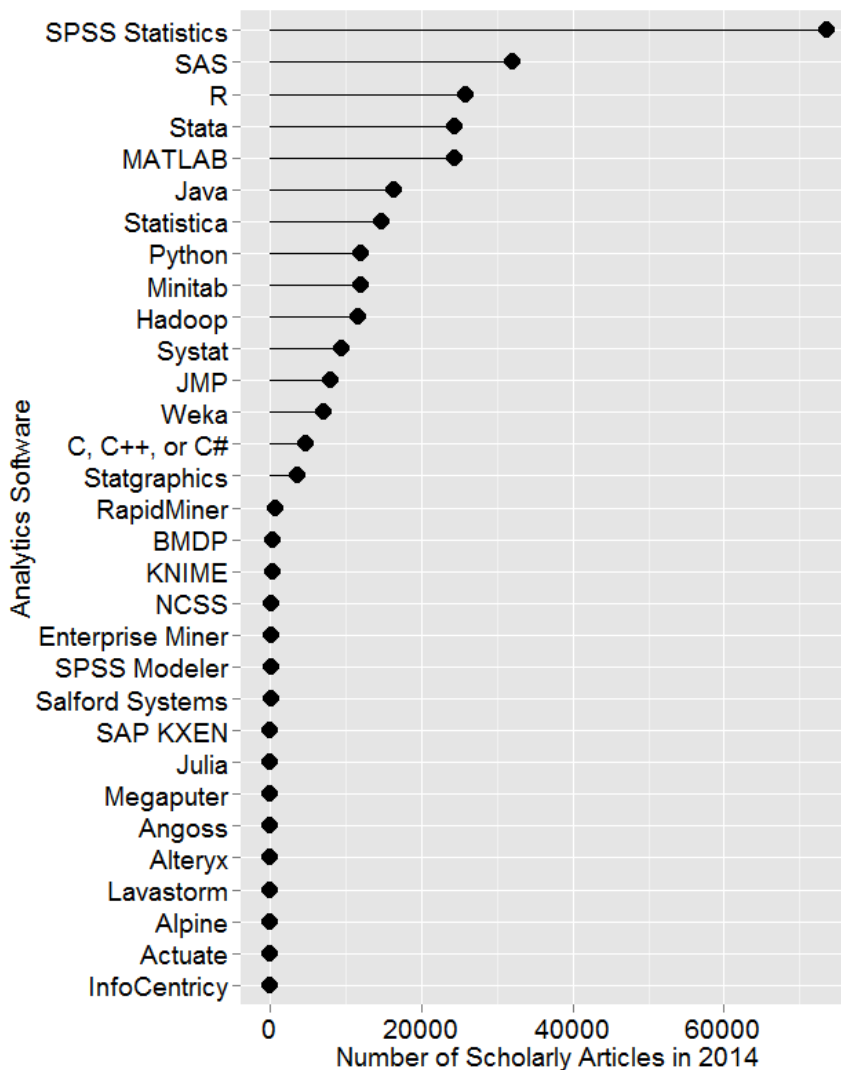
Από την άλλη το R διατίθεται δωρεάν, ωστόσο επειδή είναι open source λογισμικό και απαιτεί χρήση κώδικα δεν προτιμάται από τους χρήστες. Επομένως, το R δεν θεωρείται user friendly λογισμικό.

Όπως φαίνεται και από το ακόλουθο γράφημα 4-13, το SPSS είναι το περισσότερο διαδεδομένο πρόγραμμα και χρησιμοποιείται από πλήθος ερευνητών. Παρόλα αυτά παρουσιάζει κάποιες αδυναμίες στην εφαρμογή πολύπλοκων πειραματικών σχεδιασμών.

Τέλος, το Minitab χρησιμοποιείται κυρίως σε προβλήματα που έχουν να κάνουν με τον στατιστικό ποιοτικό έλεγχο. Είδαμε ότι ακόμη και στην περίπτωση προσαρμογής μεικτού μοντέλου εμφανίζει ορισμένες αδυναμίες σε σχέση με τα υπόλοιπα 3 λογισμικά.

Γράφημα 4-13

Χρήση λογισμικών σε δημοσιεύσεις το 2014



Κεφάλαιο 5

Ανάλυση ευαισθησίας (Sensitivity analysis)

(Πατήρ, Υιός και Άγιο Πνεύμα και όλα τ' άλλα klein mein! – Σάκης Παλαμάς)

5-1. Εισαγωγή

Η ανάλυση ευαισθησίας (sensitivity analysis) παίζει σημαντικό ρόλο προκειμένου να αξιολογηθεί η ανθεκτικότητα των αποτελεσμάτων και των συμπερασμάτων που προέκυψαν από την αρχική ανάλυση των δεδομένων του πειράματος [19]. Με απλά λόγια, αποτελεί μια μέθοδο με την οποία δίνεται η δυνατότητα στον εκάστοτε ερευνητή να ελέγξει την επίδραση διαφόρων παραμέτρων του μοντέλου που έχει εκτιμήσει μετά την αρχική του ανάλυση. Ειδικότερα, μπορεί να ελέγξει εκ νέου τη συμπεριφορά της εξαρτημένης μεταβλητής (απόκρισης), να προσεγγίσει τα δεδομένα με διαφορετική μέθοδο, να διαχειριστεί διαφορετικά τις ακραίες τιμές (outliers) και τις ελλείπουσες τιμές (missing data).

Επομένως, η ανάλυση ευαισθησίας πραγματοποιείται για να διερευνηθεί η ισχύς των συμπερασμάτων της αρχικής ανάλυσης υπό διαφορετικές όμως

συνθήκες. Η σταθερότητα των αποτελεσμάτων μετά την ανάλυση ή τις αναλύσεις ευαισθησίας δίνει σε αυτά μεγαλύτερη αξιοπιστία [20].

5-2. Ανάλυση ευαισθησίας και crossover δοκιμές

Έχουμε ήδη αναφέρει ότι οι μελέτες διασταυρούμενης μετάβασης (crossover studies) αποτελούν τυχαίοποιημένες έρευνες που κυρίως εφαρμόζονται στον τομέα της ανάπτυξης θεραπευτικών σκευασμάτων (φαρμάκων) και στην βιομηχανία τροφίμων. Θυμίζουμε ότι στον τομέα των κλινικών δοκιμών οι crossover δοκιμές έχουν πολύ μεγάλη εφαρμογή κατά τη διενέργεια των φάσεων I και II.

Κατά το σχεδιασμό των περισσότερων κλινικών δοκιμών προτείνεται από τους περισσότερους ερευνητές να προβλεφθεί στο πρωτόκολλο και η εφαρμογή ανάλυσης ευαισθησίας. Ως εκ τούτου, και στις crossover δοκιμές είναι συνετό να εφαρμόζεται ανάλυση ευαισθησίας προκειμένου να αξιολογούνται τα ευρήματα της αρχικής ανάλυσης. [19]

5-3. Ορισμός και χρησιμότητα ανάλυσης ευαισθησίας

Ένας γενικός ορισμός για την ανάλυση ευαισθησίας αναφέρει ότι πρόκειται για τη μελέτη του τρόπου με τον οποίο η αβεβαιότητα του αποτελέσματος (output) ενός μαθηματικού μοντέλου ή συστήματος μπορεί να καταμεριστεί στις επιμέρους εισροές του (inputs). [20] Πρόκειται ασφαλώς για μια προσέγγιση που προέρχεται από τον τομέα της πληροφορικής.

Στο σημείο αυτό πρέπει να διευκρινιστεί ότι **η ανάλυση ευαισθησίας είναι μια μέθοδος που - σε καμία περίπτωση - δεν μελετά μια νέα υπόθεση** που να σχετίζεται με το αποτέλεσμα. Αντίθετα, με την ανάλυση ευαισθησίας αντιμετωπίζεται η ίδια αρχική ερώτηση με διαφορετικό τρόπο ή προσέγγιση.

Συνεπώς, η ανάλυση ευαισθησίας είναι μια μέθοδος με την οποία καθορίζεται η ανθεκτικότητα μιας αξιολόγησης και διερευνά ποια αποτελέσματα επηρεάζονται από αλλαγές στις μεθόδους, τα μοντέλα και τις τιμές των μεταβλητών με σκοπό τον εντοπισμό αυτών (των αποτελεσμάτων) που είναι περισσότερο εξαρτημένα από αμφισβητήσιμες υποθέσεις. [19]

Ο σχεδιασμός και η ανάλυση crossover μελετών βασίζεται σε υποθέσεις οι οποίες αν δεν ικανοποιούνται τότε μπορεί να υπάρχει σημαντική επίδραση

στα αποτελέσματα και ως εκ τούτου στα συμπεράσματα. Αυτές οι επιδράσεις εκτιμώνται με τη βοήθεια της ανάλυσης ευαισθησίας. Η συνάφεια μεταξύ των αποτελεσμάτων της αρχικής ανάλυσης και της ανάλυσης ευαισθησίας ενισχύουν σημαντικά την αξιοπιστία των ευρημάτων και τα συμπεράσματα της έρευνας. Από τα παραπάνω είναι προφανής η χρησιμότητα εφαρμογής της ανάλυσης ευαισθησίας στις crossover δοκιμές αλλά και σε κάθε είδους πειραματικό σχεδιασμό. [19]

Στην πράξη και παρά τη σπουδαιότητά της, η ανάλυση ευαισθησίας δεν χρησιμοποιείται αρκετά συχνά. Στον τομέα της υγείας η χρήση της είναι μικρή και κυρίως σε μελέτες που αφορούν τα οικονομικά της υγείας παρά σε έρευνες που αφορούν στο κλινικό κομμάτι.

5-4. Βασικές προσεγγίσεις ανάλυσης ευαισθησίας

Οι κυριότερες επιδράσεις που μελετώνται με τη βοήθεια της ανάλυσης ευαισθησίας παρουσιάζονται σε αδρές γραμμές παρακάτω:

α) Επίδραση στην εξαρτημένη μεταβλητή (απόκριση). Στην περίπτωση όπου η απόκριση είναι δίτιμη κατηγορική μεταβλητή μπορεί να πραγματοποιηθεί ανάλυση ευαισθησίας με την επιλογή διαφορετικού διαχωριστικού ορίου (cut off point) σε σχέση με την αρχική ανάλυση. Από την άλλη, αν η απόκριση είναι συνεχής μεταβλητή τότε η ανάλυση ευαισθησίας μπορεί να πραγματοποιηθεί προσθέτοντας «θόρυβο» στις τιμές (της απόκρισης).

β) Επίδραση χρήσης διαφορετικών μεθόδων ανάλυσης. Σε αυτή την περίπτωση μπορεί να πραγματοποιηθεί ανάλυση ευαισθησίας με τη χρήση μιας διαφορετικής στατιστικής μεθόδου σε σχέση με εκείνη που επιλέχθηκε στην αρχική ανάλυση. Για παράδειγμα, στην αρχική προσέγγιση χρησιμοποιήθηκε από τον ερευνητή ένα γενικό γραμμικό μοντέλο (two way anova) και στην ανάλυση ευαισθησίας έγινε χρήση μεικτού μοντέλου.

γ) Επίδραση ακραίων τιμών (outliers). Μια τιμή θεωρείται ακραία αν διαφέρει σημαντικά από τις υπόλοιπες τιμές ενός σετ δεδομένων. Ο εντοπισμός του γίνεται συνήθως γραφικά με τη βοήθεια ενός γραφήματος διασποράς (scatterplot) ή ενός θηκογράμματος (boxplot) ή με τη βοήθεια των z-τιμών. Σε αυτή την περίπτωση, η ανάλυση ευαισθησίας μπορεί να γίνει λαμβάνοντας

υπόψιν (ή μη) την ή τις ακραίες τιμές. Η απόφαση αυτή εξαρτάται από το αν συμπεριλήφθηκαν στην αρχική ανάλυση οι ακραίες τιμές ή όχι.

δ) Επίδραση ελλειπουσών τιμών (missing data). Μια τιμή χαρακτηρίζεται ελλείπουσα όταν δεν καταγράφεται ως παρατήρηση για κάποια μεταβλητή. Οι λόγοι της μη καταγραφής μπορεί να είναι ποικίλοι. Έχει αναπτυχθεί ολόκληρη θεωρία σχετικά με τη διαχείριση των ελλειπουσών τιμών η οποία δεν αφορά τους στόχους της παρούσας εργασίας. [21] Με την ανάλυση ευαισθησίας, υπάρχει η δυνατότητα διαφορετικής διαχείρισης των ελλειπουσών τιμών σε μεταβλητή ή μεταβλητές σε σχέση με εκείνη της αρχικής ανάλυσης.

5-5. Ανάλυση ευαισθησίας στο test γευσιγνωσίας σοκολάτας

Στην εφαρμογή του κεφαλαίου 4 η εξαρτημένη μεταβλητή (απόκριση) αναφερόταν στη βαθμολογία ως προς την προτίμηση που έδωσαν εξήντα (60) γευσιγνώστες ως προς τρεις ίδιες σοκολάτες όπου κάθε μία περιείχε διαφορετική γλυκαντική ουσία. Η βαθμολογία που μπορούσαν να επιλέξουν οι γευσιγνώστες ήταν από 1 έως και 5, όπου η τιμή 1 υποδήλωνε ελάχιστη προτίμηση και η τιμή 5 μέγιστη προτίμηση.

Επομένως, η εξαρτημένη μεταβλητή (απόκριση) ήταν μια διακριτή μεταβλητή που έπαιρνε συγκεκριμένες διακριτές τιμές 1, 2, 3, 4 και 5. Με απλά λόγια, η μεταβλητή βαθμολογία προτίμησης αποτελούσε κατηγορική μεταβλητή ιεράρχησης. Η κλίμακα μέτρησης της προσομοιάζε με 5-βαθμια κλίμακα Likert. Θυμίζουμε ότι για τις ανάγκες της εφαρμογής στο κεφάλαιο 4 η βαθμολογία ως προς την προτίμηση θεωρήθηκε ότι είναι συνεχής τυχαία μεταβλητή.

Στο προηγούμενο κεφάλαιο μελετήσαμε αρχικά - και με τη βοήθεια 4 στατιστικών λογισμικών - την ύπαρξη στατιστικά σημαντικής επίδρασης ακολουθίας (carry-over effects). Η προσέγγιση η οποία επιλέχθηκε ήταν με τη χρήση μεικτού γραμμικού μοντέλου (mixed model) το οποίο περιλάμβανε τους παρακάτω όρους:

-Εξαρτημένη μεταβλητή (απόκριση) = Βαθμολογία ως προς την προτίμηση
-Ανεξάρτητη μεταβλητή με σταθερές επιδράσεις (fixed effects) = Επίδραση ακολουθίας (carry-over ή residual effects) και

-Ανεξάρτητη μεταβλητή με τυχαίες επιδράσεις (random effects) = O A/A των γευσισγνωστών (id ή subject).

Το αποτέλεσμα ήταν ότι και με τα 4 στατιστικά λογισμικά απορρίφθηκε – έστω και οριακά – η ύπαρξη σημαντικής επίδρασης ακολουθίας (carry-over effects) στα δεδομένα. Αυτό επέτρεψε τη συνέχιση της στατιστικής ανάλυσης με την εισαγωγή στο μοντέλο και άλλων παραγόντων όπως η περίοδος, η αλληλουχία δοκιμής σοκολατών κλπ.

Όπως αναφέρθηκε και παραπάνω, η ανάλυση ευαισθησίας είναι μια μέθοδος με την οποία αξιολογούμε τα ευρήματα της αρχικής ανάλυσης υπό διαφορετικές όμως συνθήκες. Στη συγκεκριμένη περίπτωση, παρεμβαίνουμε στις τιμές της εξαρτημένης μεταβλητής (απόκρισης) προσθέτοντας σ' αυτές τυχαίο «θόρυβο». Αυτό έχει ως αποτέλεσμα η μεταβλητή «Βαθμολογία προτίμησης» να μετατραπεί από διακριτή που λαμβάνει τιμές 1, 2, 3, 4 και 5 σε συνεχής τυχαία μεταβλητή. Ο τυχαίος «θόρυβος» προέρχεται από κανονική κατανομή και προστίθεται στις αρχικές τιμές της εξαρτημένης μεταβλητής (απόκρισης). Αξίζει να σημειωθεί ότι η διαδικασία επαναλαμβάνεται 4 φορές – αυξάνοντας κάθε φορά τη διακύμανση.

Συγκεκριμένα, ο τυχαίος «θόρυβος» που προστίθεται στις τιμές της «Βαθμολογίας προτίμησης» προέρχεται από τις ακόλουθες κανονικές κατανομές:

- α) $N(0,0.2)$
- β) $N(0,0.5)$
- γ) $N(0,1)$ και
- δ) $N(0,1.5)$

Για να προσθέσουμε στις αρχικές παρατηρήσεις τυχαίο «θόρυβο» χρησιμοποιούμε το στατιστικό λογισμικό R και ο κώδικας για να το επιτύχουμε αυτό δίνεται αμέσως παρακάτω (1^η περίπτωση $N(0,0.2)$):

```
cho["preferwithnoise1"]<-NA
head(cho)
cho$preferwithnoise1<-cho$prefer+rnorm(length(cho$prefer),0,0.2)
head(cho)
```

Ο παραπάνω κώδικας μας δίνει τις νέες τιμές της μεταβλητής «Βαθμολογία προτίμησης» στις οποίες έχει προστεθεί τυχαίος «θόρυβος» από κανονική κατανομή με μέση τιμή 0 και διακύμανση 0.2. Αυξάνοντας τη διακύμανση σε 0.5 και επαναλαμβάνοντας τον κώδικα προκύπτουν νέες τιμές για την εξαρτημένη μεταβλητή (απόκριση). Το ίδιο πραγματοποιούμε ακόμη δύο φορές αυξάνοντας περαιτέρω τη διακύμανση σε 1 και 1.5 αντίστοιχα.

Στον ακόλουθο πίνακα δίνονται οι αρχικές παρατηρήσεις καθώς και οι νέες όπως αυτές προκύπτουν από την εφαρμογή τυχαίου «θορύβου» (δίνονται οι 6 πρώτες παρατηρήσεις):

Πίνακας 5-1

Αρχικές και νέες τιμές μεταβλητής «Βαθμολογία προτίμησης»

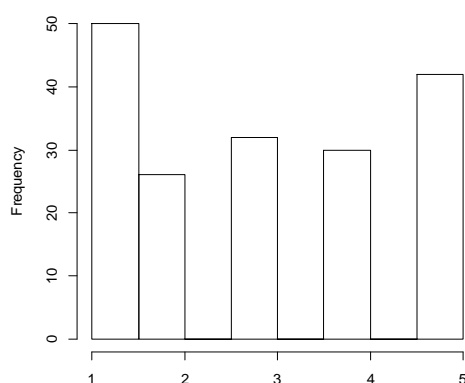
Τιμές μεταβλητής «Βαθμολογία προτίμησης»					
A/A	Αρχικές	N(0,0.2)	N(0,0.5)	N(0,1)	N(0,1.5)
1 ^η	3	2.57	3.82	3.71	2.91
2 ^η	4	4.08	3.55	4.86	4.27
3 ^η	1	0.68	0.88	0.56	0.85
4 ^η	5	4.89	4.72	4.50	2.97
5 ^η	1	0.79	0.45	2.28	-1.14
6 ^η	3	3.40	2.18	6.41	4.03
...

95

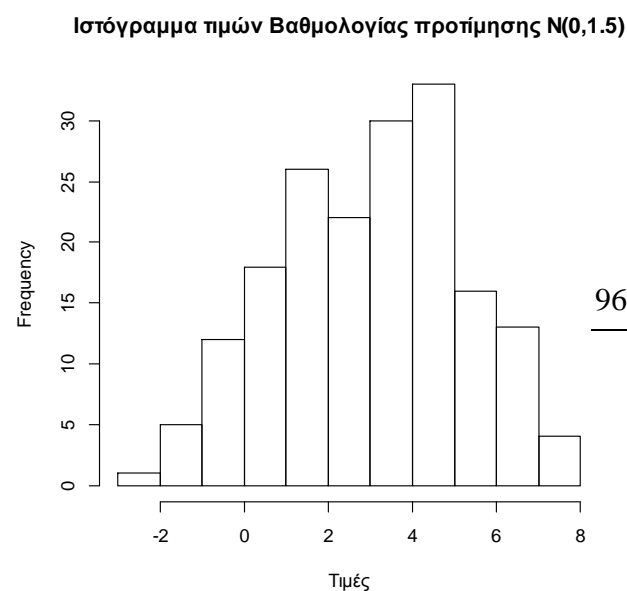
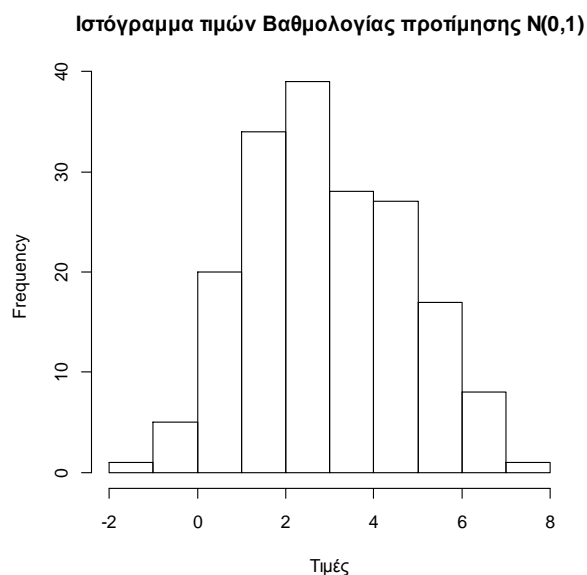
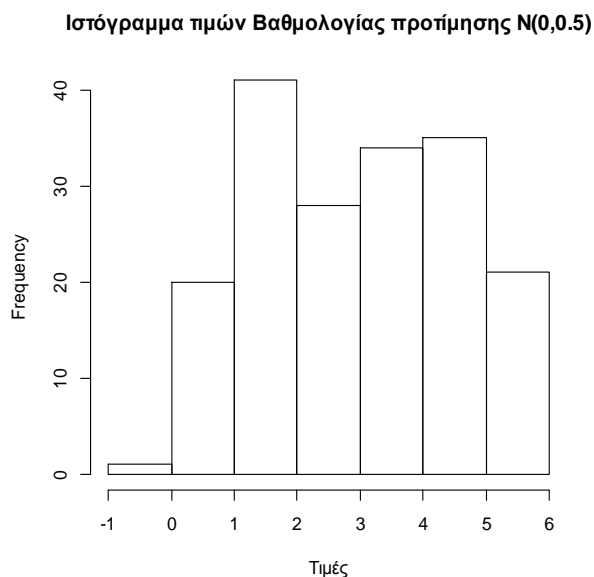
Ακολούθως δίνονται τα ιστογράμματα των 5 μεταβλητών (το πρώτο με τις αρχικές και τα επόμενα 4 με τις νέες τιμές). Παρατηρούμε ότι το πρώτο γράφημα που αφορά στις αρχικές τιμές της μεταβλητής «Βαθμολογία προτίμησης» είναι ραβδόγραμμα (διακριτή μεταβλητή).

Γράφημα 5-1

Ιστόγραμμα αρχικών τιμών Βαθμολογίας προτίμησης



ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»



Οι τιμές της εξαρτημένης μεταβλητής (απόκρισης) στα γραφήματα 4 και 5 φαίνεται ότι ακολουθούν κανονική κατανομή κάτι που επιβεβαιώνεται και από τον ακόλουθο πίνακα ($p_{N(0,1)} = 0.201$ και $p_{N(0,1.5)} = 0.128$).

Πίνακας 5-2

Έλεγχος κανονικότητας αρχικών και νέων τιμών «Βαθμολογίας προτίμησης»

Shapiro – Wilk Normality Test για «Βαθμολογία προτίμησης»					
Πιθανότητα	Αρχικές	N(0,0.2)	N(0,0.5)	N(0,1)	N(0,1.5)
p-value	0.000	0.000	0.000	0.201	0.128

Πίνακας 5-3

Βασικά περιγραφικά μέτρα αρχικών και νέων τιμών «Βαθμολογία προτίμησης»

Περιγραφικά μέτρα μεταβλητής «Βαθμολογία προτίμησης»					
Μέτρα	Αρχικές	N(0,0.2)	N(0,0.5)	N(0,1)	N(0,1.5)
Min	1.0	0.68	-0.51	-1.50	-2.20
1 st Quar.	1.0	1.3	1.59	1.49	1.42
Median	3.0	2.92	2.98	2.84	3.27
Mean	2.93	2.91	2.99	2.94	3.03
3 rd Quar.	4.0	4.24	4.38	4.19	4.44
Max	5.0	5.32	5.95	7.26	7.85
Stddev	1.54	1.53	1.59	1.78	2.20
SE	0.11	0.11	0.12	0.13	0.16
Range	4.0	4.64	6.47	8.77	10.06

Η προσθήκη τυχαίου «θορύβου» στις τιμές της μεταβλητής «Βαθμολογία προτίμησης» είχε ως αποτέλεσμα τη μετατροπή της από διακριτή σε συνεχή. Στον πίνακα 5-3 αποτυπώνονται ορισμένα βασικά περιγραφικά μέτρα για τις 5 μεταβλητές. Παρατηρούμε ότι όσο αυξάνεται η διακύμανση τόσο μεγαλώνει και το εύρος των παρατηρήσεων. Για παράδειγμα, οι τιμές της μεταβλητής στην οποία προσθέσαμε τυχαίο θόρυβο από N(0,1.5) «απλώνονται» από -2.20 έως 7.85.

Έχει ήδη αναφερθεί ότι σε δεδομένα που προέρχονται από crossover δοκιμή είναι ζωτικής σημασίας να εξεταστεί αν υφίσταται στατιστικά σημαντική επίδραση ακολουθίας (carry-over effects). Όπως είδαμε και στο κεφάλαιο 4 – με τη βοήθεια μεικτού γραμμικού μοντέλου ελέγξαμε αν υπάρχει σημαντική επίδραση ακολουθίας (carry-over effects) στις αρχικές τιμές της μεταβλητής «Βαθμολογία προτίμησης» (βλ. στήλη 2 – Πίνακα 5-4). Υπενθυμίζεται ότι η επίδραση ακολουθίας (carry-over effects) είναι η διατήρηση της επίδρασης μιας θεραπείας A που εφαρμόστηκε στην 1^η περίοδο και στην επόμενη (μεταγενέστερη) 2^η περίοδο. Τονίζεται ότι στην 2^η περίοδο εφαρμόζεται θεραπεία B (διαφορετική της A). [22]

Πίνακας 5-4

Προσαρμογή μεικτού μοντέλου
σε αρχικές και νέες τιμές «Βαθμολογίας προτίμησης»

Mixed model μεταβλητής «Βαθμολογία προτίμησης»						
Μέτρο	Αρχικές	N(0,0.2)	N(0,0.5)	N(0,1)	N(0,1.5)	
Model fit						
AIC	675.54	673.13	686.41	725.40	792.42	
BIC	697.73	695.33	708.6	747.59	814.61	
LogLik	-330.77	-329.56	-336.20	-355.70	-389.21	
-2LogLik	661.54	659.12	672.40	711.40	778.42	
Random effects						
StdDev	0.000	0.000	0.000	0.008	0.000	
Fixed effects						
t	A	-1.85	-2.10	-2.03	-2.27	-2.71
	B	-0.88	-1.01	-1.24	-0.98	-0.21
	C	-2.58	-2.69	-2.88	-2.72	-3.44
p-value	A	0.066	0.037	0.04	0.024	0.007
	B	0.376	0.311	0.21	0.324	0.831
	C	0.011	0.008	0.004	0.007	0.000
Anova table						
F	2.55	2.88	3.08	3.117	5.68	
p-value	0.058	0.038	0.03	0.028	0.001	

Συγκεκριμένα προσαρμόσαμε στις τιμές της εξαρτημένης μεταβλητής ένα μεικτό μοντέλο όπου ως μεταβλητή σταθερών επιδράσεων (fixed effects) θεωρήσαμε την επίδραση ακολουθίας (carry-over effects) και ως μεταβλητή τυχαίων επιδράσεων (random effects) τον A/A των γευσιγνωστών (id ή subject). Το αποτέλεσμα που προέκυψε ήταν το επιθυμητό, δεδομένου ότι δεν υφίστατο στατιστικά σημαντική επίδραση ακολουθίας στα δεδομένα ($p_A = 0.058$). Αξίζει να σημειωθεί ότι ίδιο αποτέλεσμα έδωσαν και τα υπόλοιπα στατιστικά λογισμικά (στον παραπάνω πίνακα παρουσιάζονται τα αποτελέσματα από το R). Η έλλειψη σημαντικής επίδρασης ακολουθίας (carry-over effects) επέτρεψε την περαιτέρω ανάλυση με την εξέταση και άλλων βασικών παραγόντων του crossover πειράματος.

Με την ανάλυση ευαισθησίας, προσθέσαμε στις τιμές της εξαρτημένης μεταβλητής (απόκρισης) τυχαίο «θόρυβο» από κανονική κατανομή 4 φορές - αυξάνοντας σταδιακά τη διακύμανση. Στις 4 νέες μεταβλητές - με τη βοήθεια του στατιστικού λογισμικού R – προσαρμόσαμε μεικτά μοντέλα για να ελέγξουμε την ύπαρξη στατιστικά σημαντικής επίδρασης ακολουθίας (carry-over effects).

Για κάθε μοντέλο υπολογίστηκαν τα μέτρα καλής προσαρμογής του μοντέλου (AIC, BIC και -2LogLikelihood) οι τιμές των οποίων παρατηρούμε ότι αυξάνονται όσο αυξάνεται και η διακύμανση του τυχαίου «θορύβου» που προσθέταμε στις αρχικές τιμές κάθε φορά. Στη συνέχεια, δίνεται για κάθε μοντέλο η εκτίμηση της μεταβλητότητας της εξαρτημένης μεταβλητής «προτίμηση» που εξηγείται από τον τυχαίο παράγοντα id γευσιγνωστών (random effects). Παρατηρούμε ότι για το σύνολο των μοντέλων η εκτίμηση αυτή είναι πάρα πολύ μικρή (πρακτικά μηδέν). Τούτο σημαίνει ότι ο τυχαίος παράγοντας δεν έχει καμιά ουσιαστική επίδραση στο μοντέλο και ενδεχομένως θα έπρεπε να εξαλειφθεί. Ακολουθώντας, δίνονται οι εκτιμήσεις της επίδρασης carryover (fixed effects) ως προς τις κατηγορίες A, B και C μαζί με τα τυπικά τους σφάλματα. Ως κατηγορία αναφοράς επιλέχθηκε η κατηγορία Null (έλλειψη επίδρασης).

Τη σημαντικότερη πληροφορία μας τη δίνει η ανάλυση διακύμανσης του μοντέλου όπου λαμβάνοντας υπόψιν την p-τιμή, μπορούμε να αποφανθούμε αν τελικά υφίσταται στατιστικά σημαντική επίδραση ακολουθίας ή όχι. Παρατηρούμε ότι και τα 4 μοντέλα με τα οποία πραγματοποιήσαμε ανάλυση ευαισθησίας μας δίνουν p-τιμές < 0.05. Τούτο σημαίνει ότι υπάρχει στατιστικά σημαντική επίδραση ακολουθίας (carry-over effects). Ως εκ τούτου περιορίζονται σημαντικά οι δυνατότητες μας για περαιτέρω στατιστική ανάλυση. Στην ουσία μόνο οι τιμές της 1^{ης} περιόδου (στις οποίες ούτως ή άλλως δεν υπάρχει επίδραση carry-over) μπορούν να χρησιμοποιηθούν για στατιστική επεξεργασία.

Στο παράρτημα του κεφαλαίου 5 δίνεται ο κώδικας R και για τα 4 μοντέλα. Επιπλέον, στο παράρτημα περιλαμβάνονται και οι εντολές υπολογισμού των υπολοίπων (residuals) και των εκτιμώμενων τιμών (fitted values) για όλα τα μοντέλα, καθώς και εντολή για τη δημιουργία γραφήματος των υπολοίπων ως προς τις εκτιμώμενες τιμές.

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

5-6. Συμπέρασμα

Είχε αναφερθεί ότι η σταθερότητα των αποτελεσμάτων μετά την ανάλυση ευαισθησίας δίνει σε αυτά μεγαλύτερη αξιοπιστία. Στη συγκεκριμένη περίπτωση τα αρχικά αποτελέσματα άρα και συμπεράσματα δεν επιβεβαιώθηκαν. Στις αρχικές τιμές της εξαρτημένης μεταβλητής (απόκρισης) προστέθηκε τυχαίος «θόρυβος» από κανονική κατανομή 4 φορές με μεγαλύτερη κάθε φορά διακύμανση. Στις 4 νέες μεταβλητές πραγματοποιήθηκε έλεγχος για την ύπαρξη (ή μη) στατιστικά σημαντικής επίδρασης ακολουθίας (carry-over effects) με τη χρήση μεικτού μοντέλου. Σε όλες τις περιπτώσεις διαπιστώθηκε ότι υφίσταται σημαντική επίδραση carry-over κάτι που δεν ίσχυε στο μοντέλο που προσαρμόσαμε στις αρχικές τιμές. Επομένως, ο ερευνητής δεν μπορεί να νιώθει σιγουριά με το πείραμα του και η ανάλυση ευαισθησίας που πραγματοποιήθηκε παραπάνω υποδεικνύει την ανάγκη για **χρήση μεγαλύτερου δείγματος για την εξαγωγή ασφαλέστερων συμπερασμάτων.**_

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Liu, M.B., Davis, K. (2010). *A Clinical Trial Manual from the Duke Clinical Research Institute* (2nd ed.), Chichester, West Sussex: Wiley – Blackwell
2. Pocock, S.J. (1983). *Clinical Trials – A Practical Approach* (1st ed.), Chichester, West Sussex: Wiley
3. Παπαϊωάννου,Τ. (2003). *Σημειώσεις μαθήματος Κλινικές Δοκιμές*, ΠΜΣ Εφαρμοσμένη Στατιστική, Τμήμα Στατιστικής & Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς
4. Λαθύρης, Δ. (2010). *Μελέτες Διασταυρούμενης Μετάβασης: Εμπειρική Μεταναλυτική Προσέγγιση*, Διδακτορική Διατριβή, Πανεπιστήμιο Ιωαννίνων
5. Senn, S.J. (2002). *Cross - over Trials in Clinical Research* (2nd ed.), Chichester, West Sussex: Wiley
6. Abas, N. (1995). *Crossover Trials in Medical Statistics*, MSc Thesis, University of Wales
7. Jones, B., Kenward, M.G. (2003). *Design and Analysis of Cross – over Trials* (2nd ed.), Washington, D.C: Chapman & Hall/ CRC
8. Cleophas, T.J., Zwinderman, A.H., Cleophas, T.F., Cleophas, E.P. (2006). *Statistics Applied to Clinical Trials* (1st ed.), Netherlands: Springer
9. Montgomery, D.C. (1997). *Design and Analysis of Experiments* (8th ed.), NJ: Wiley
10. Cox, D.R., Reid, N. (2000). *The Theory of the Design of Experiments*, Washington, D.C: Chapman & Hall/ CRC
11. Dean, A.M., Voss, D. (1999). *Design and Analysis of Experiments*, (1st Ed.), New York: Springer
12. Appa, G., Magos, D., Mourtos, I. (2004). *An LP- based proof for the non-existence of a pair of Orthogonal Latin Squares of order 6*, Operations Research Letters 32, 336-344
13. Stinson, D.R. (2004). *Combinatorial Designs–Constructions and Analysis*, (1st ed.), New York: Springer
14. Mckay, B.D., Wanless, I.M. (2005). *On the number of Latin Squares*, Ann. Comb. 9, 335 - 344

15. Kim, K., Kumar Prosanna, V.K. (1989). *Perfect latin Squares and parallel array access*, in: International Symposium on Computer Architecture, 372 -379
16. Kutner, M.H. Nachtsheim, C.J., Neter, J., Li, W. (2005). *Applied Linear Statistical Models*, (5th ed.), New York: Mc Graw –Hill /Irwin
17. Rosner, B. (2011). *Fundamentals of Biostatistics*, (7th ed.), Boston, MA: Brooks/ Cole
18. Καραβασίλης, Γ. (2012). *Σχεδιασμός, Ανάπτυξη και Αξιολόγηση Εναλλακτικού Διδακτικού Υλικού: Επεξεργασία και Ανάλυση Ερευνητικών Δεδομένων σε R*, Διπλωματική Εργασία, Ελληνικό Ανοικτό Πανεπιστήμιο
19. Thabane, L. et al. (2013). *A tutorial on sensitivity analysis in clinical trials: the what, why, when and how*, BMC Medical Research Methodology, 13: 92
20. Saltelli, A. et al. (2008). *Global Sensitivity Analysis–The Primer*, Chichester, West Sussex: Wiley
21. Tsiatis, A. (2006). *Semiparametric Theory and Missing Data*, New York : Springer
22. Senn, S.J. (1992). *Is the Simple Carry-over Model Useful?*, Statistics in Medicine, 11(6): p. 715-726
23. Ελληνικό Στατιστικό Ινστιτούτο (Ε.Σ.Ι), (2009). *Λεξικό Στατιστικής Ορολογίας*, (1^η Εκδ.), Αθήνα: Έκδοση Ελληνικού Στατιστικού Ινστιτούτου

ΠΗΓΕΣ

24. <https://en.wikipedia.org/wiki/Likert> scale
25. <https://onlinecourses.science.psu.edu/stat502/node/210>
26. [https:// en.wikipedia.org/wiki/SPSS](https://en.wikipedia.org/wiki/SPSS)
27. [https:// en.wikipedia.org/wiki/Minitab](https://en.wikipedia.org/wiki/Minitab)
28. [http://www.jmp.com /en.os/home.html](http://www.jmp.com/en.os/home.html)
29. [https://en.wikipedia.org/wiki/JMP_\(statistical_software\)](https://en.wikipedia.org/wiki/JMP_(statistical_software))

ΠΑΡΑΡΤΗΜΑ

ΚΕΦΑΛΑΙΟ 4

Testing 4 softwares

Με τη βοήθεια γενικού γραμμικού μοντέλου

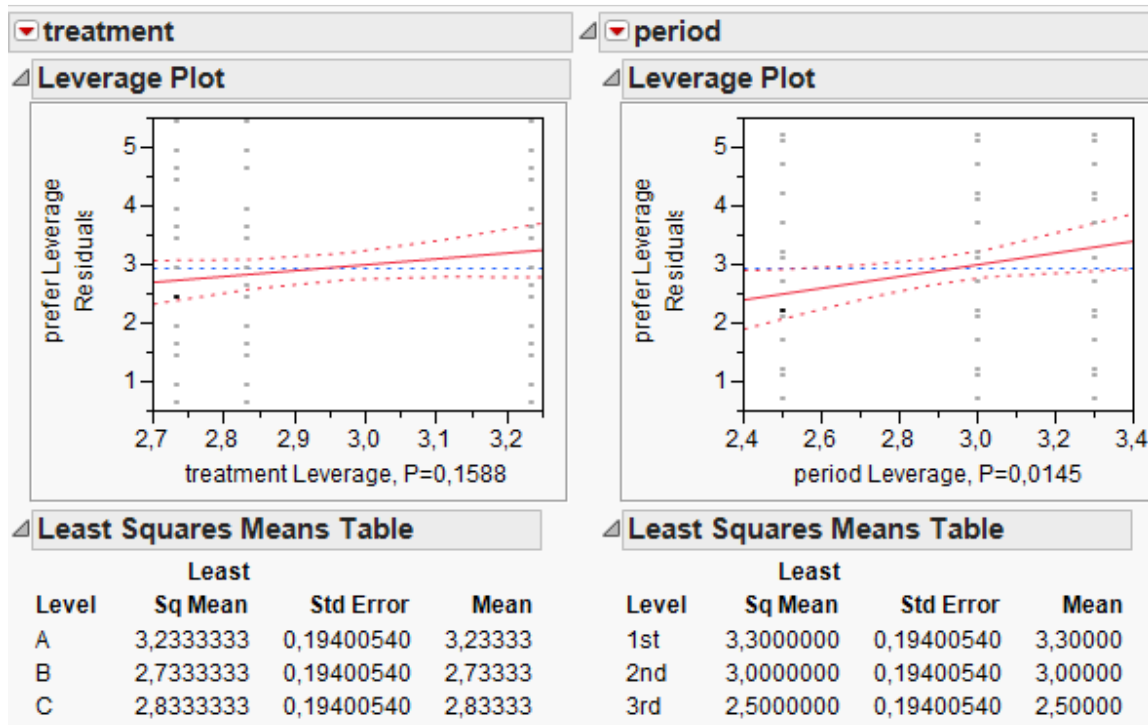
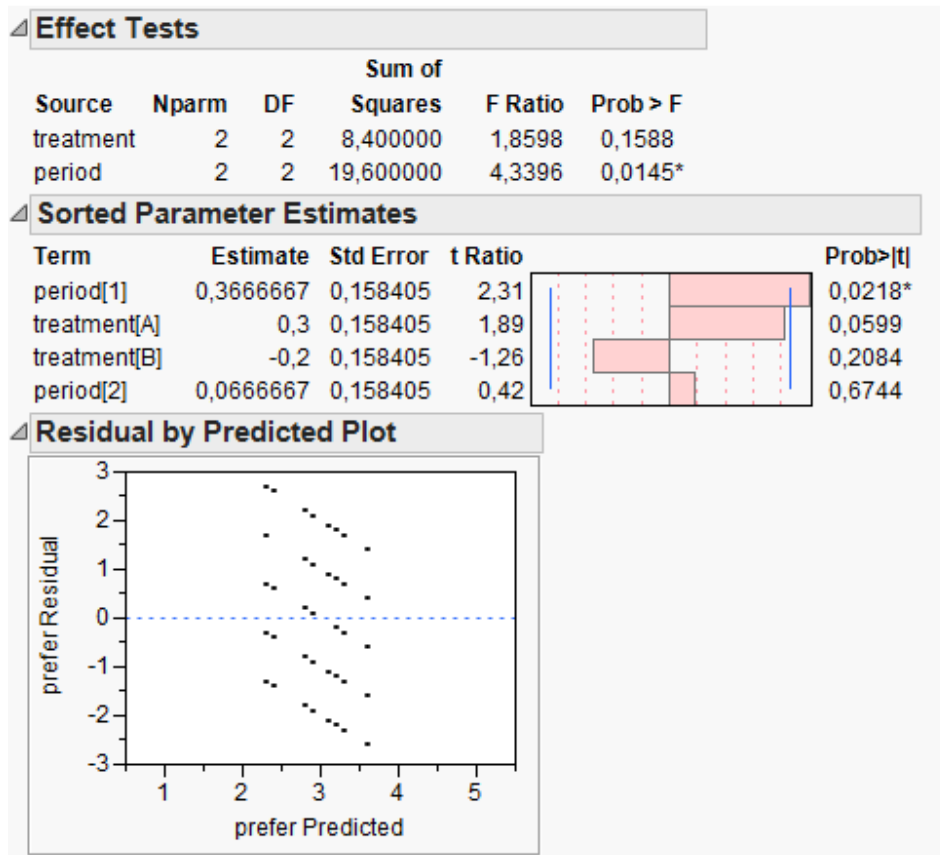
1) JMP

Summary of Fit				
RSquare		0,066163		
RSquare Adj		0,044818		
Root Mean Square Error		1,502759		
Mean of Response		2,933333		
Observations (or Sum Wgts)		180		

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	28,00000	7,00000	3,0997
Error	175	395,20000	2,25829	Prob > F
C. Total	179	423,20000		0,0170*

Lack Of Fit				
Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	4	41,00000	10,2500	4,9485
Pure Error	171	354,20000	2,0713	Prob > F
Total Error	175	395,20000		0,0008*

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	2,9333333	0,112009	26,19	<,0001*
treatment[A]	0,3	0,158405	1,89	0,0599
treatment[B]	-0,2	0,158405	-1,26	0,2084
period[1]	0,3666667	0,158405	2,31	0,0218*
period[2]	0,0666667	0,158405	0,42	0,6744



2) Minitab

Method

Factor coding (1; 0)

Factor Information

Factor	Type	Levels	Values
treatment	Fixed	3	A; B; C
period	Fixed	3	1st; 2nd; 3rd

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
treatment	2	8,400	4,200	1,86	0,159
period	2	19,600	9,800	4,34	0,014
Error	175	395,200	2,258		
Lack-of-Fit	4	41,000	10,250	4,95	0,001
Pure Error	171	354,200	2,071		
Total	179	423,200			

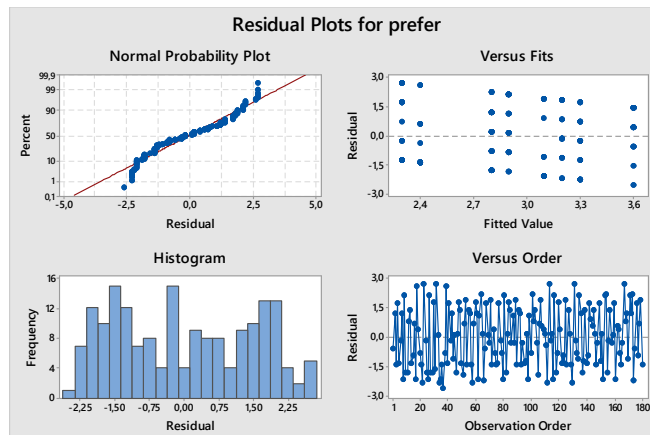
Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1,50276	6,62%	4,48%	1,20%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	2,400	0,250	9,58	0,000	
treatment					
A	0,400	0,274	1,46	0,147	1,33
B	-0,100	0,274	-0,36	0,716	1,33
period					
1st	0,800	0,274	2,92	0,004	1,33
2nd	0,500	0,274	1,82	0,070	1,33

Regression Equation

$$\text{prefer} = 2,400 + 0,400 \text{ treatment}_A - 0,100 \text{ treatment}_B + 0,0 \text{ treatment}_C + 0,800 \text{ period}_{1st} + 0,500 \text{ period}_{2nd} + 0,0 \text{ period}_{3rd}$$


ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

3) R

```
library(nlme)
library(lattice)
library(lme4)
library(MASS)
cho<-read.table("choco0.txt",header=T)
attach(cho)
names(cho)
head(cho)
str(cho)
cho$treatment<-relevel(cho$treatment, ref="C")
cho$period<-relevel(cho$period, ref="3rd")
test1<-lm(prefer~as.factor(treatment)+as.factor(period),data=cho)
summary(test1)
anova(test1)
```

4) SPSS

Descriptive Statistics

Dependent Variable: Choco Preference

Chocolate	Period	Mean	Std. Deviation	N
A	1	4,10	1,165	20
	2	2,85	1,424	20
	3	2,75	1,446	20
	Total	3,23	1,466	60
B	1	2,50	1,821	20
	2	2,70	1,261	20
	3	3,00	1,654	20
	Total	2,73	1,582	60
C	1	3,30	1,302	20
	2	3,45	1,432	20
	3	1,75	1,333	20
	Total	2,83	1,542	60
Total	1	3,30	1,576	60
	2	3,00	1,390	60
	3	2,50	1,557	60
	Total	2,93	1,538	180

Levene's Test of Equality of Error Variances^a

Dependent Variable: Choco Preference

F	df1	df2	Sig.
2,939	8	171	,004

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + treatment + period

Tests of Between-Subjects Effects

Dependent Variable: Choco Preference

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	28,000 ^a	4	7,000	3,100	,017
Intercept	1548,800	1	1548,800	685,830	,000
treatment	8,400	2	4,200	1,860	,159
period	19,600	2	9,800	4,340	,014
Error	395,200	175	2,258		
Total	1972,000	180			
Corrected Total	423,200	179			

a. R Squared = ,066 (Adjusted R Squared = ,045)

Parameter Estimates

Dependent Variable: Choco Preference

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	2,400	,250	9,582	,000	1,906	2,894
[treatment=1]	,400	,274	1,458	,147	-,141	,941
[treatment=2]	-,100	,274	-,364	,716	-,641	,441
[treatment=3]	0 ^a
[period=1]	,800	,274	2,916	,004	,259	1,341
[period=2]	,500	,274	1,822	,070	-,041	1,041
[period=3]	0 ^a

a. This parameter is set to zero because it is redundant.

Estimates

Dependent Variable: Choco Preference

Period	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	3,300	,194	2,917	3,683
2	3,000	,194	2,617	3,383
3	2,500	,194	2,117	2,883

Pairwise Comparisons

Dependent Variable: Choco Preference

(I) Period	(J) Period	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
1	2	,300	,274	,276	-,241	,841
	3	,800*	,274	,004	,259	1,341
2	1	-,300	,274	,276	-,841	,241
	3	,500	,274	,070	-,041	1,041
3	1	-,800*	,274	,004	-1,341	-,259
	2	-,500	,274	,070	-1,041	,041

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Univariate Tests

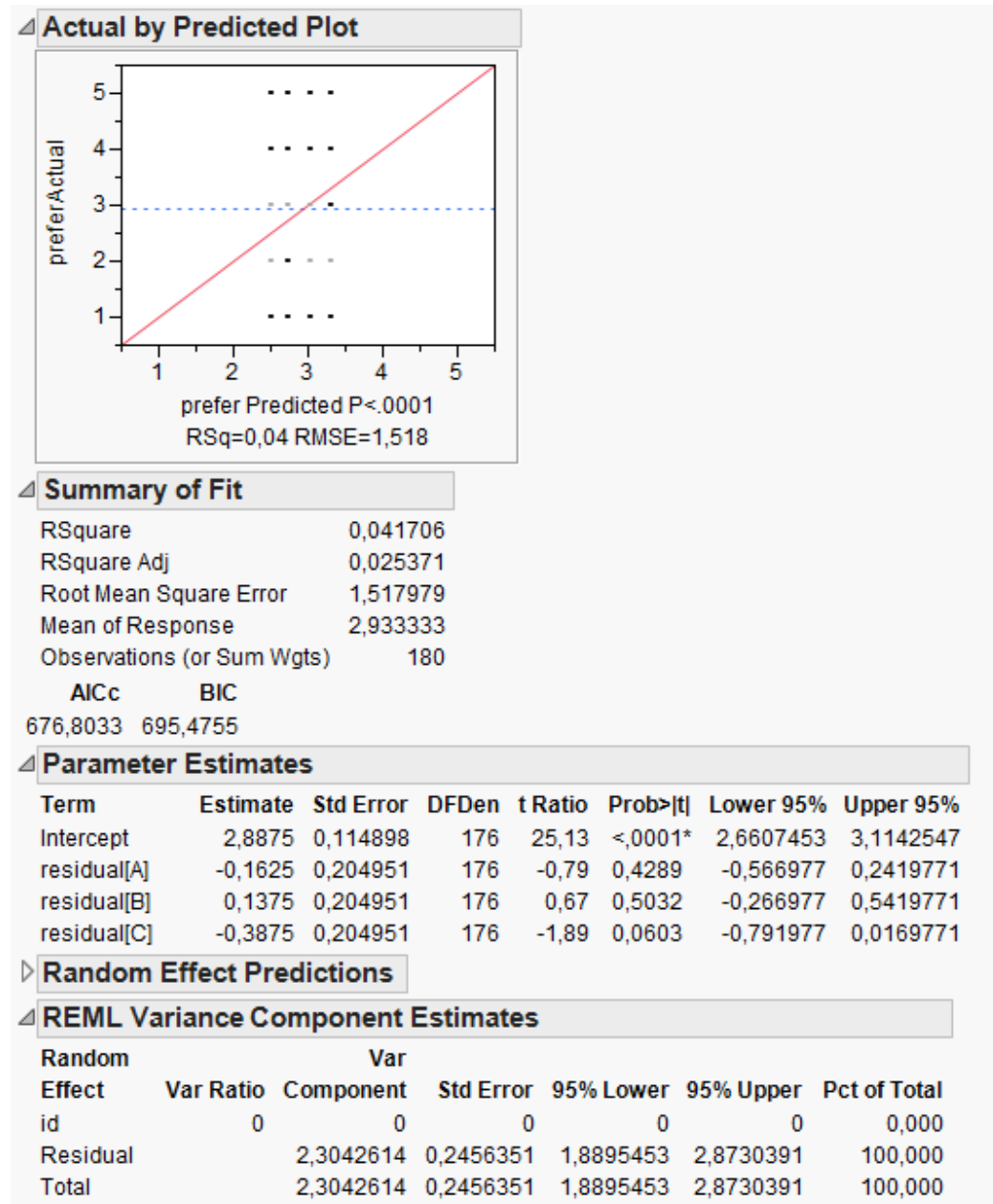
Dependent Variable: Choco Preference

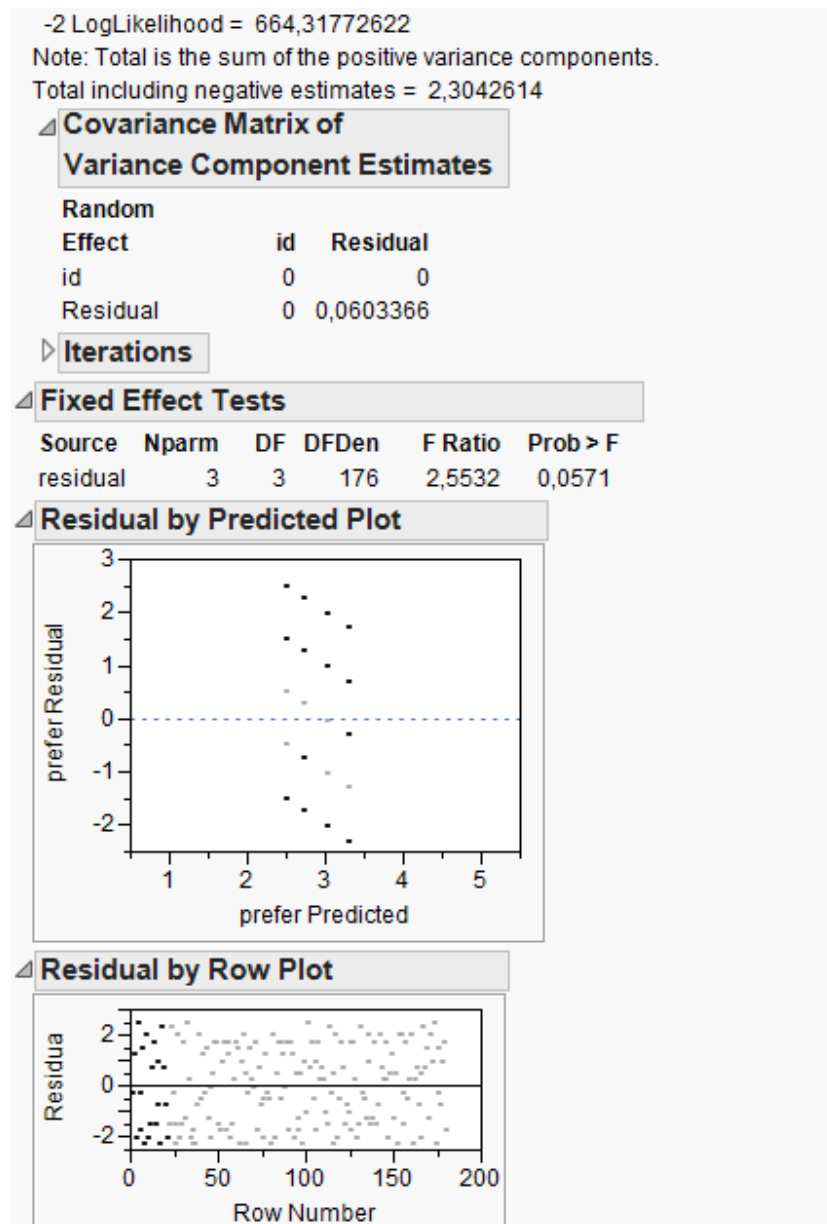
	Sum of Squares	df	Mean Square	F	Sig.
Contrast	19,600	2	9,800	4,340	,014
Error	395,200	175	2,258		

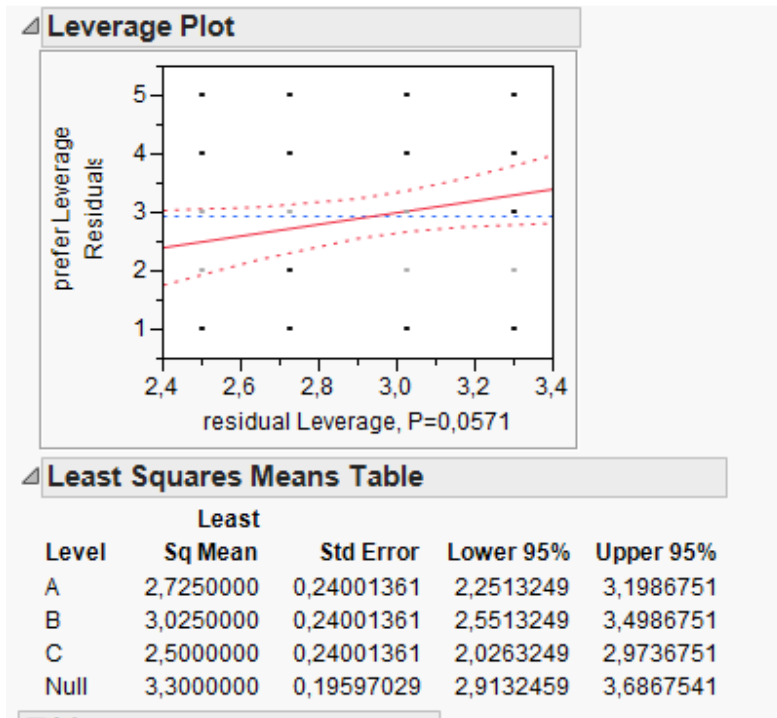
The F tests the effect of Period. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Έλεγχος επίδρασης ακολουθίας (carry-over effects)

1) JMP







2) Minitab

General Linear Model: Βαθμολογία προτίμησης vs επίδραση carryover

Method

Factor coding (1; 0)

Factor Information

Factor	Type	Levels	Values
carryover	Fixed	4	0; 1; 2; 3
id	Random	60	1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 17; 18; 19; 20; 21; 22; 23; 24; 25; 26; 27; 28; 29; 30; 31; 32; 33; 34; 35; 36; 37; 38; 39; 40; 41; 42; 43; 44; 45; 46; 47; 48; 49; 50; 51; 52; 53; 54; 55; 56; 57; 58; 59; 60

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
carryover	3	17,65	5,883	2,55	0,057
Error	176	405,55	2,304		
Total	179	423,20			

Model Summary

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Παράρτημα

S R-sq R-sq(adj) R-sq(pred)
 1,51798 4,17% 2,54% 0,00%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	3,300	0,196	16,84	0,000	
carryover					
1	-0,575	0,310	-1,86	0,065	1,30
2	-0,275	0,310	-0,89	0,376	1,30
3	-0,800	0,310	-2,58	0,011	1,30

id	Coef	SE Coef	T-Value	P-Value	VIF
2	0,55	1,43	0,38	0,702	*
3	0,75	1,43	0,52	0,602	*
4	-0,58	1,43	-0,41	0,685	*
5	0,22	1,43	0,15	0,880	*
6	0,67	1,43	0,47	0,641	*
7	-0,33	1,43	-0,23	0,816	*
8	0,55	1,43	0,38	0,702	*
.
.
.

Regression Equation

prefer = 3,300 + 0,0 carryover_0 - 0,575 carryover_1 - 0,275 carryover_2
 - 0,800 carryover_3 + 0,0 id_1 + 0,55 id_2 + 0,75 id_3 - 0,58 id_4 + 0,22 id_5
 + 0,67 id_6 etc

Equation treats random terms as though they are fixed.

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 carryover	(3) + Q[1]
2 id	(3) + 2,9831 (2)
3 Error	(3)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 carryover	117,00	3,0490	(3)
2 id	117,00	3,0490	(3)

Means

Term	Fitted Mean
carryover	
0	3,30000
1	2,70000
2	3,10000
3	2,45000
id	
1	2,52083
2	3,07083

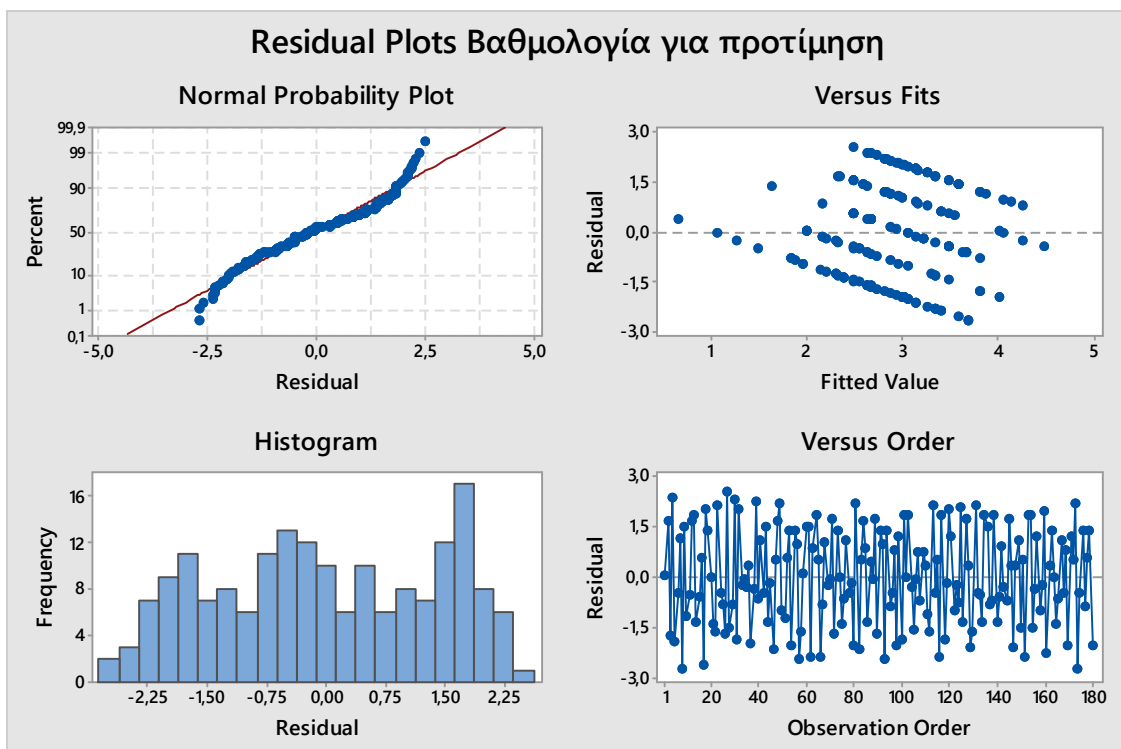
ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Παράρτημα

```
3          3,27083
4          1,93750
5          2,73750
6          3,18750
.          .
.          .
.          .
```

Variance Components, using Adjusted SS

Source	Variance	% of Total	StDev	% of Total
id	-0,744741*	0,00%	0,00000	0,00%
Error	3,049	100,00%	1,74614	100,00%
Total	3,049		1,74614	



3) R

```
install.packages("nlme")
```

```
install.packages("lattice")
```

```
install.packages("lme4")
```

```
install.packages("MASS")
```

```
install.packages("car")
```

```
install.packages("psych")
```

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Παράρτημα

```
install.packages("MuMIn")
install.packages("Ismeans")
library(nlme)
library(lattice)
library(lme4)
library(MASS)
library(car)
library(psych)
library(MuMIn)
library(Ismeans)
cho<-read.table("choco0.txt",header=T)
attach(cho)
names(cho)
head(cho)
str(cho)
cho$residual<-relevel(cho$residual, ref="Null")
mix1<-lme(prefer~as.factor(residual), data=cho,random=~1|id/residual, method="REML")
summary(mix1)
anova(mix1) # Type I SS
anova.lme(mix1,type="marginal",adjustSigma=F) # Type III SS
Ismeans(mix1,"residual")
resid(mix1)
fitted(mix1)
plot(mix1,main="Γράφημα τυποποιημένων υπολοίπων ως προς τις εκτιμώμενες τιμές")
mix2<-lmer(prefer~as.factor(residual)+(1|id), data=cho)
summary(mix2)
r.squaredGLMM(mix2)
confint(mix2)
```

115

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

4) SPSS

Mixed model analysis

Model Dimension^a

		Number of Levels	Covariance Structure	Number of Parameters
Fixed Effects	Intercept	1	Identity	1
	residual	4		3
Random Effects	Intercept + id(residual) ^b	181		1
Residual				1
Total		186		6

a. Dependent Variable: Choco Preference.

b. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.

Information Criteria^a

-2 Restricted Log Likelihood	661,545
Akaike's Information Criterion (AIC)	665,545
Hurvich and Tsai's Criterion (AICC)	665,615
Bozdogan's Criterion (CAIC)	673,886
Schwarz's Bayesian Criterion (BIC)	671,886

The information criteria are displayed in smaller-is-better form.

a. Dependent Variable: Choco Preference.

Type III Tests of Fixed Effects^a

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	1290252,150	7,376	,007
residual	3	176,000	2,553	,057

a. Dependent Variable: Choco Preference.

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	3,300000	1,074956	159336,086	3,070	,002	1,193110	5,406890
[residual=1]	-,575000	,309856	176,000	-1,856	,065	-1,186512	,036512
[residual=2]	-,275000	,309856	176,000	-,888	,376	-,886512	,336512
[residual=3]	-,800000	,309856	176,000	-2,582	,011	-1,411512	-,188488
[residual=4]	0 ^b	0

a. Dependent Variable: Choco Preference.

b. This parameter is set to zero because it is redundant.

Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1,187136	,245635	4,833	,000	,791361	1,780844
Intercept + id(residual) Variance	1,117126 ^b	,000000

a. Dependent Variable: Choco Preference.

b. This covariance parameter is redundant. The test statistic and confidence interval cannot be computed.

Estimates^a

Carryover effect	Mean	Std. Error	df	95% Confidence Interval	
				Lower Bound	Upper Bound
A	2,725	,240	176,000	2,251	3,199
B	3,025	,240	176,000	2,551	3,499
C	2,500	,240	176,000	2,026	2,974
Null	3,300	,196	176,000	2,913	3,687

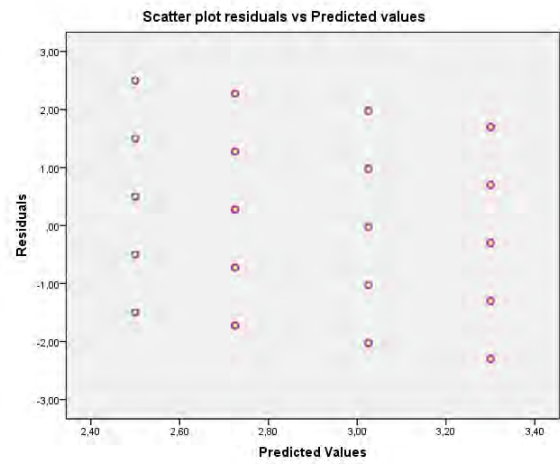
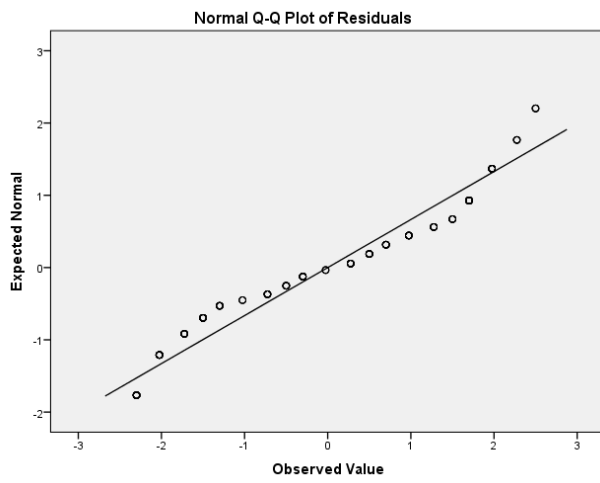
a. Dependent Variable: Choco Preference.

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residuals	,123	180	,000	,925	180	,000

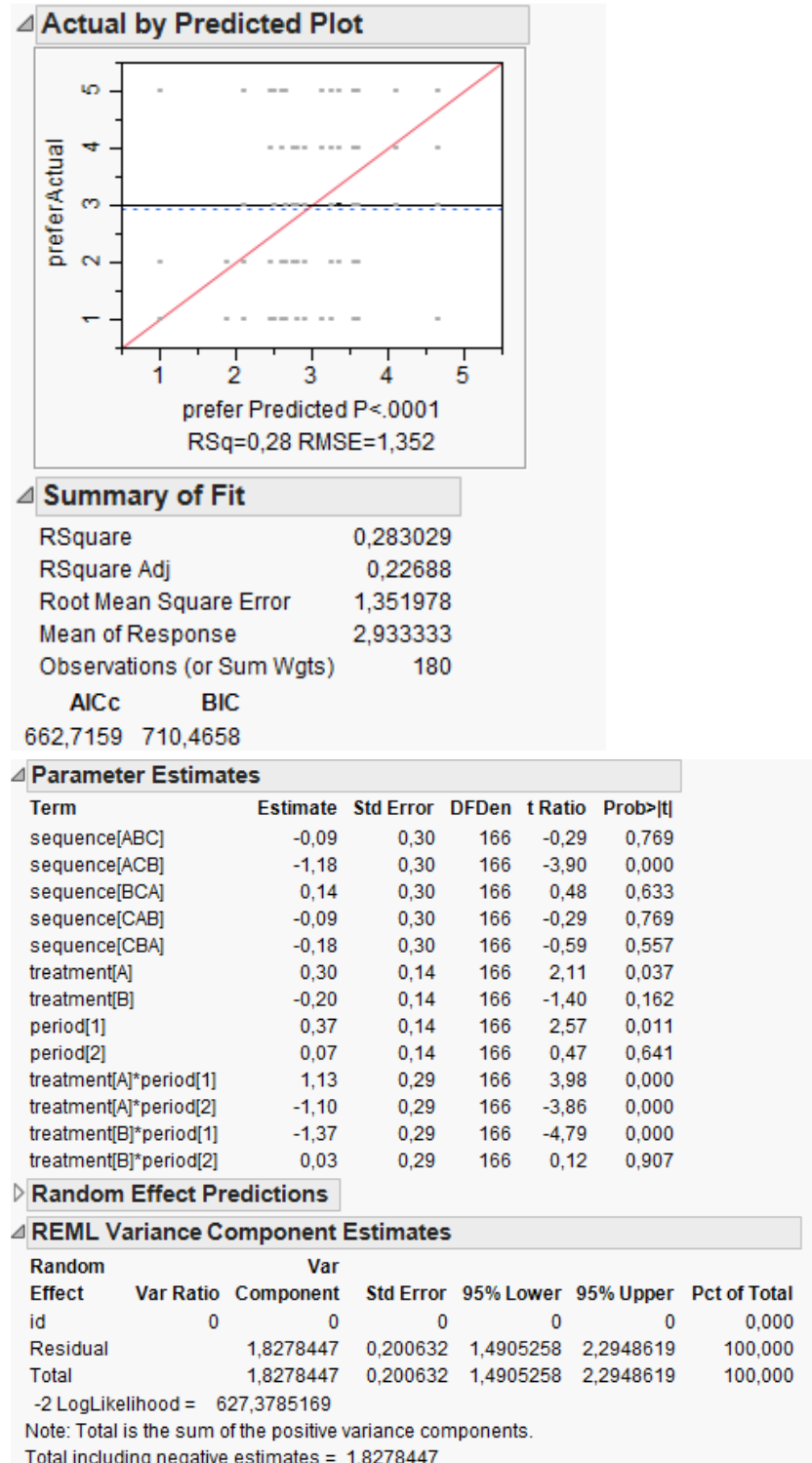
a. Lilliefors Significance Correction

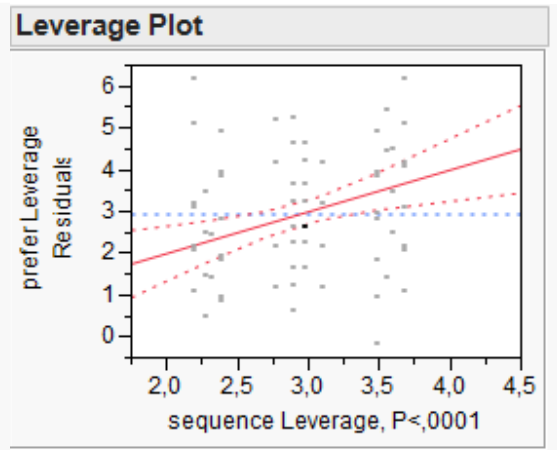
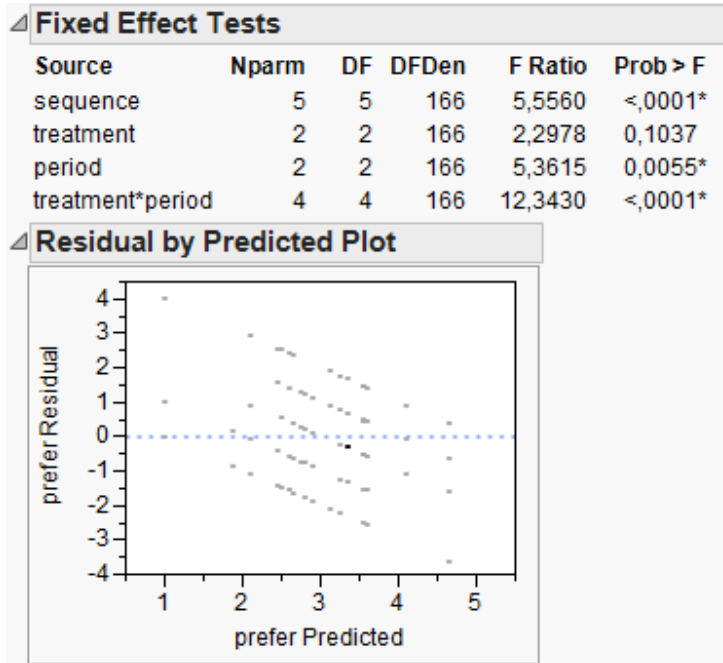
Παράρτημα



Πλήρες μοντέλο

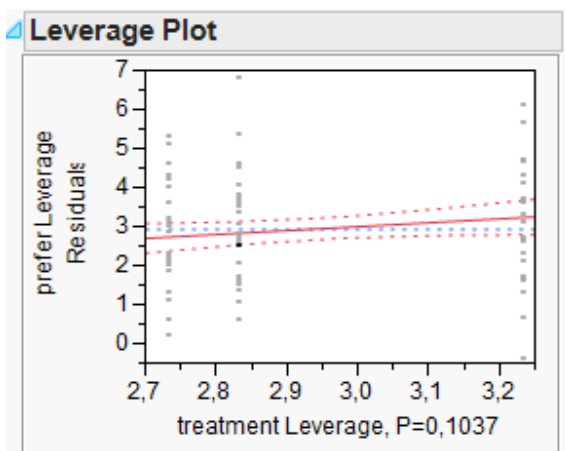
1) JMP





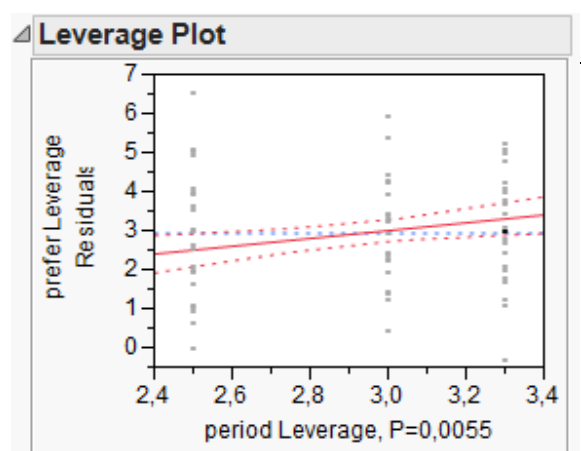
Least Squares Means Table

Level	Sq Mean	Std Error
ABC	2,8444444	0,31866429
ACB	1,7555556	0,31866429
BCA	3,0777778	0,31866429
CAB	2,8444444	0,31866429
CBA	2,7555556	0,31866429
BAC	4,3222222	0,31866429



Least Squares Means Table

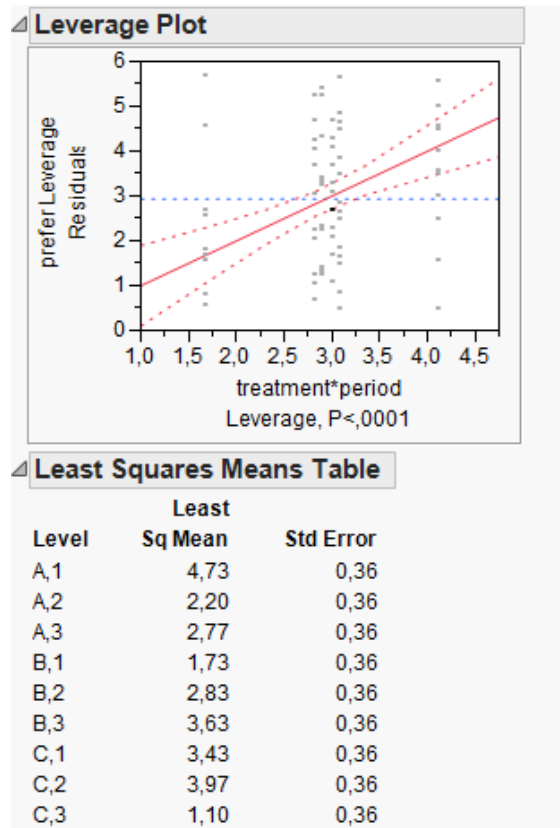
Level	Sq Mean	Std Error
A	3,2333333	0,17453962
B	2,7333333	0,17453962
C	2,8333333	0,17453962



Least Squares Means Table

Level	Sq Mean	Std Error
1	3,3000000	0,17453962
2	3,0000000	0,17453962
3	2,5000000	0,17453962

Παράρτημα



121

2) Minitab

Method

Factor coding (1; 0)

Factor Information

Factor	Type	Levels	Values
sequence	Fixed	6	ABC; ACB; BAC; BCA; CAB; CBA
treatment	Fixed	3	A; B; C
period	Fixed	3	1st; 2nd; 3rd

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
sequence	5	50,78	10,156	5,56	0,000
treatment	2	39,79	19,893	10,88	0,000
period	2	55,79	27,893	15,26	0,000
treatment*period	4	90,24	22,561	12,34	0,000
Error	166	303,42	1,828		
Total	179	423,20			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
---	------	-----------	------------

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Παράρτημα

1,35198 28,30% 22,69% 15,70%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	2,489	0,377	6,60	0,000	
sequence					
ABC	-1,478	0,451	-3,28	0,001	2,78
ACB	-2,567	0,494	-5,20	0,000	3,33
BAC	0,000000	0,000000	*	*	*
BCA	-1,244	0,451	-2,76	0,006	2,78
CAB	-1,478	0,451	-3,28	0,001	2,78
CBA	-1,567	0,494	-3,17	0,002	3,33
treatment					
A	1,667	0,552	3,02	0,003	6,67
B	2,533	0,552	4,59	0,000	6,67
C	0,000000	0,000000	*	*	*
period					
1st	2,333	0,552	4,23	0,000	6,67
2nd	2,867	0,552	5,19	0,000	6,67
3rd	0,000000	0,000000	*	*	*
treatment*period					
A 1st	-0,367	0,855	-0,43	0,669	7,11
A 2nd	-3,433	0,855	-4,02	0,000	7,11
A 3rd	0,000000	0,000000	*	*	*
B 1st	-4,233	0,855	-4,95	0,000	7,11
B 2nd	-3,667	0,855	-4,29	0,000	7,11
B 3rd	0,000000	0,000000	*	*	*
C 1st	0,000000	0,000000	*	*	*
C 2nd	0,000000	0,000000	*	*	*
C 3rd	0,000000	0,000000	*	*	*

122

Regression Equation

prefer = 2,489 - 1,478 sequence_ABC - 2,567 sequence_ACB + 0,0 sequence_BAC
 - 1,244 sequence_BCA - 1,478 sequence_CAB - 1,567 sequence_CBA
 + 1,667 treatment_A
 + 2,533 treatment_B + 0,0 treatment_C + 2,333 period_1st + 2,867 period_2nd
 + 0,0 period_3rd - 0,367 treatment*period_A 1st - 3,433 treatment*period_A
 2nd
 + 0,0 treatment*period_A 3rd - 4,233 treatment*period_B 1st
 - 3,667 treatment*period_B 2nd + 0,0 treatment*period_B 3rd
 + 0,0 treatment*period_C
 1st + 0,0 treatment*period_C 2nd + 0,0 treatment*period_C 3rd

Fits and Diagnostics for Unusual Observations

Obs	prefer	Fit	Resid	Std Resid	
4	5,000	2,111	2,889	2,23	R
37	1,000	4,644	-3,644	-2,81	R
39	5,000	1,011	3,989	3,07	R

R Large residual

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Παράρτημα

1	sequence	(5) + Q[1]
2	treatment	(5) + Q[2; 4]
3	period	(5) + Q[3; 4]
4	treatment*period	(5) + Q[4]
5	Error	(5)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 sequence	166,00	1,8278	(5)
2 treatment	166,00	1,8278	(5)
3 period	166,00	1,8278	(5)
4 treatment*period	166,00	1,8278	(5)

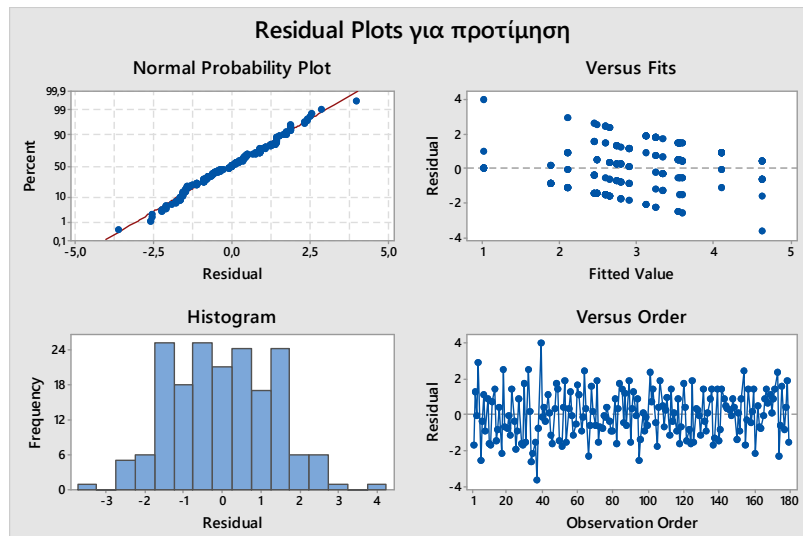
Means

Term	Fitted Mean	SE Mean
sequence		
ABC	2,844	0,319
ACB	1,756	0,319
BAC	4,322	0,319
BCA	3,078	0,319
CAB	2,844	0,319
CBA	2,756	0,319
treatment		
A	3,233	0,175
B	2,733	0,175
C	2,833	0,175
period		
1st	3,300	0,175
2nd	3,000	0,175
3rd	2,500	0,175

123

Variance Components, using Adjusted SS

Source	Variance	% of Total	StDev	% of Total
Error	1,82784	100,00%	1,35198	100,00%
Total	1,82784		1,35198	



ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

3) R

```
library(nlme)
```

```
library(lattice)
```

```
library(lme4)
```

```
library(MASS)
```

```
library(foreign)
```

```
library(psych)
```

```
library(car)
```

```
library(MuMIn)
```

```
library(Ismeans)
```

```
cho<-read.table("choco0.txt",header=T)
```

```
attach(cho)
```

```
names(cho)
```

```
head(cho)
```

```
str(cho)
```

```
cho$sequence<-relevel(cho$sequence, ref="BAC")
```

```
cho$treatment<-relevel(cho$treatment, ref="C")
```

```
cho$period<-relevel(cho$period, ref="3rd")
```

```
mixF<-
```

```
lme(prefer~as.factor(sequence)+treatment+period+treatment*period,random=~1|id/sequence,data=cho,method="REML")
```

```
anova.lme(mixF,type="marginal",adjustSigma=F)
```

```
summary(mixF)
```

```
r.squaredGLMM(mixF)
```

```
Ismeans(mixF, "sequence")
```

```
Ismeans(mixF, "treatment")
```

```
Ismeans(mixF, "period")
```

```
plot(mixF,main="Γράφημα τυποποιημένων υπολοίπων ως προς τις εκτιμώμενες τιμές")
```

Παράρτημα

`resid(mixF)`

`fitted(mixF)`

`interaction.plot(sequence,period,prefer)`

`interaction.plot(period,sequence,prefer)`

4) SPSS

Model Dimension^a

		Number of Levels	Covariance Structure	Number of Parameters
Fixed Effects	Intercept	1	Identity	1
	sequence	6		5
	treatment	3		2
	period	3		2
	treatment * period	9		4
Random Effects	Intercept + id ^b	61	1	1
Residual				1
Total		83		16

a. Dependent Variable: Choco Preference.

b. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.

125

Information Criteria^a

-2 Restricted Log Likelihood	610,612
Akaike's Information Criterion (AIC)	614,612
Hurvich and Tsai's Criterion (AICC)	614,685
Bozdogan's Criterion (CAIC)	622,836
Schwarz's Bayesian Criterion (BIC)	620,836

The information criteria are displayed in smaller-is-better form.

a. Dependent Variable: Choco Preference.

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Type III Tests of Fixed Effects^a

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	166	847,337	,000
sequence	5	166	5,556	,000
treatment	2	166	2,298	,104
period	2	166	5,362	,006
treatment * period	4	166	12,343	,000

a. Dependent Variable: Choco Preference.

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	2,49	,38	166,000	6,601	,000	1,74	3,23
[sequence=ABC]	-1,48	,45	166	-3,279	,001	-2,37	-,59
[sequence=CAB]	-1,48	,45	166	-3,279	,001	-2,37	-,59
[sequence=BCA]	-1,24	,45	166	-2,761	,006	-2,13	-,35
[sequence=CBA]	-1,57	,49	166	-3,173	,002	-2,54	-,59
[sequence=ACB]	-2,57	,49	166	-5,199	,000	-3,54	-1,59
[period=1st]	2,33	,55	166,000	4,227	,000	1,24	3,42
[period=2nd]	2,87	,55	166,000	5,194	,000	1,78	3,96
[treatment=A] * [period=1]	-,37	,86	166	-,429	,669	-2,05	1,32
[treatment=A] * [period=2]	-3,43	,86	166	-4,015	,000	-5,12	-1,75
[treatment=B] * [period=1]	-4,23	,86	166	-4,951	,000	-5,92	-2,55
[treatment=B] * [period=2]	-3,67	,86	166	-4,288	,000	-5,35	-1,98

a. Dependent Variable: Choco Preference.

Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error
Residual	1,827845	,200632
Intercept + id Variance	,000000 ^b	,000000

a. Dependent Variable: Choco Preference.

b. This covariance parameter is redundant.

Covariance Matrix for Estimates of Covariance Parameters^a

Parameter	Residual	Intercept + id
		Variance
Residual	,040253	,000000
Intercept + id	Variance	,000000

a. Dependent Variable: Choco Preference.

Estimates^a

Sequence	Mean	Std. Error	df	95% Confidence Interval	
				Lower Bound	Upper Bound
ABC	2,844	,319	166	2,215	3,474
CAB	2,844	,319	166	2,215	3,474
BCA	3,078	,319	166	2,449	3,707
CBA	2,756	,319	166	2,126	3,385
ACB	1,756	,319	166	1,126	2,385
BAC	4,322	,319	166	3,693	4,951

a. Dependent Variable: Choco Preference.

Estimates^a

Period	Mean	Std. Error	df	95% Confidence Interval	
				Lower Bound	Upper Bound
1st	3,300	,175	166	2,955	3,645
2nd	3,000	,175	166	2,655	3,345
3rd	2,500	,175	166	2,155	2,845

a. Dependent Variable: Choco Preference.

3. Chocolate * Period^a

Chocolate	Period	Mean	Std. Error	df	95% Confidence Interval	
					Lower Bound	Upper Bound
A	1st	4,733	,363	166	4,016	5,451
	2nd	2,200	,363	166	1,483	2,917
	3rd	2,767	,363	166	2,049	3,484
B	1st	1,733	,363	166	1,016	2,451
	2nd	2,833	,363	166	2,116	3,551
	3rd	3,633	,363	166	2,916	4,351
C	1st	3,433	,363	166	2,716	4,151
	2nd	3,967	,363	166	3,249	4,684
	3rd	1,100	,363	166	,383	1,817

a. Dependent Variable: Choco Preference.

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

ΚΕΦΑΛΑΙΟ 5

Sensitivity Analysis

Αρχικά πρέπει να αποθηκεύσετε το αρχείο choco1.txt στον δίσκο C του υπολογιστή σας και ειδικότερα στο φάκελο «Τα έγγραφά μου». Επισημαίνεται το γεγονός ότι κάθε φορά που «τρέχετε» τον παρακάτω κώδικα προκύπτουν νέες τιμές στις μεταβλητές $N(0,0.2)$, $N(0,0.5)$, $N(0,1)$ και $N(0,1.5)$ λόγω του τυχαίου θορύβου που προστίθεται στις αρχικές τιμές.

Reproducible code

```
install.packages("nlme")
install.packages("lattice")
install.packages("lme4")
install.packages("MASS")
install.packages("car")
install.packages("psych")

library(nlme)
library(lattice)
library(lme4)
library(MASS)
library(car)
library(psych)

cho<-read.table("choco1.txt",header=T)
attach(cho)
names(cho)
head(cho)
str(cho)

# Τυχαίος Θόρυβος N(0,0.2)
cho["preferwithnoise1"]<-NA
head(cho)
```

Παράρτημα

```
cho$preferwithnoise1<-cho$prefer+rnorm(length(cho$prefer),0,0.2)
head(cho)
summary(cho)
describe(cho)
shapiro.test(cho$preferwithnoise1)
hist(cho$preferwithnoise1,main="Ιστόγραμμα τιμών Βαθμολογίας προτίμησης N(0,0.2)",xlab="Τιμές")
cho$residual<-relevel(cho$residual, ref="Null") #Select reference group
mix1<-lme(preferwithnoise1~as.factor(residual), data=cho,random=~1|id/residual, method="REML")
summary(mix1)
anova(mix1)
fixef(mix1) #Estimates of fixed effects
ranef(mix1) #Estimates of random effects
fit<-fitted(mix1) #Predicted or fitted values
summary(fit)
residuals<-resid(mix1) #Residuals
summary(residuals)
plot(residuals~fit, main="Residuals vs Fitted values N(0,0.2)",ylab="Residuals", xlab="Fitted")
```

129

Τυχαίος Θόρυβος N(0,0.5)

```
cho["preferwithnoise2"]<-NA
head(cho)
cho$preferwithnoise2<-cho$prefer+rnorm(length(cho$prefer),0,0.5)
head(cho)
summary(cho)
describe(cho)
shapiro.test(cho$preferwithnoise2)
hist(cho$preferwithnoise2,main="Ιστόγραμμα τιμών Βαθμολογίας προτίμησης N(0,0.5)",xlab="Τιμές")
cho$residual<-relevel(cho$residual, ref="Null")
mix2<-lme(preferwithnoise2~as.factor(residual), data=cho,random=~1|id/residual, method="REML")
```

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Παράρτημα

```
summary(mix2)
```

```
anova(mix2)
```

```
fixef(mix2) #Estimates of fixed effects
```

```
ranef(mix2) #Estimates of random effects
```

```
fit<-fitted(mix2) #Predicted or fitted values
```

```
summary(fit)
```

```
residuals<-resid(mix2) #Residuals
```

```
summary(residuals)
```

```
plot(residuals~fit, main="Residuals vs Fitted values N(0,0.5)",ylab="Residuals", xlab="Fitted")
```

Τυχαίος Θόρυβος N(0,1)

```
cho["preferwithnoise3"]<-NA
```

```
head(cho)
```

```
cho$preferwithnoise3<-cho$prefer+rnorm(length(cho$prefer),0,1)
```

```
head(cho)
```

```
summary(cho)
```

```
describe(cho)
```

```
shapiro.test(cho$preferwithnoise3)
```

```
hist(cho$preferwithnoise3,main="Ιστόγραμμα τιμών Βαθμολογίας προτίμησης N(0,1)",xlab="Τιμές")
```

```
cho$residual<-relevel(cho$residual, ref="Null")
```

```
mix3<-lme(preferwithnoise3~as.factor(residual), data=cho,random=~1|id/residual, method="REML")
```

```
summary(mix3)
```

```
anova(mix3)
```

```
fixef(mix3) #Estimates of fixed effects
```

```
ranef(mix3) #Estimates of random effects
```

```
fit<-fitted(mix3) #Predicted or fitted values
```

```
summary(fit)
```

```
residuals<-resid(mix3) #Residuals
```

```
summary(residuals)
```

ΠΜΣ «Μεθοδολογία Βιοϊατρικής Έρευνας, Βιοστατιστική και Κλινική Βιοπληροφορική»

Παράρτημα

```
plot(residuals~fit, main="Residuals vs Fitted values N(0,1)",ylab="Residuals", xlab="Fitted")
```

```
# Τυχαίος Θόρυβος N(0,1.5)
```

```
cho["preferwithnoise4"]<-NA
```

```
head(cho)
```

```
cho$preferwithnoise4<-cho$prefer+rnorm(length(cho$prefer),0,1.5)
```

```
head(cho)
```

```
summary(cho)
```

```
describe(cho)
```

```
shapiro.test(cho$preferwithnoise4)
```

```
hist(cho$preferwithnoise4,main="Ιστόγραμμα τιμών Βαθμολογίας προτίμησης N(0,1.5)",xlab="Τιμές")
```

```
cho$residual<-relevel(cho$residual, ref="Null")
```

```
mix4<-lme(preferwithnoise4~as.factor(residual), data=cho,random=~1|id/residual, method="REML")
```

```
summary(mix4)
```

```
anova(mix4)
```

```
fixef(mix4) #Estimates of fixed effects
```

```
ranef(mix4) #Estimates of random effects
```

```
fit<-fitted(mix4) #Predicted or fitted values
```

```
summary(fit)
```

```
residuals<-resid(mix4) #Residuals
```

```
summary(residuals)
```

```
plot(residuals~fit, main="Residuals vs Fitted values N(0,1.5)",ylab="Residuals", xlab="Fitted")
```