

Εισαγωγή στην Βιοστατιστική και την Επιδημιολογία

Ιωάννης Ντζούφρας

Τμήμα Στατιστικής

Οικονομικό Πανεπιστήμιο Αθηνών

Άρης Περπέρογλου

Τμήμα Στατιστικής και Αναλογιστικών-Χρηματοοικονομικών Μαθηματικών

Πανεπιστήμιο Αιγαίου

Αθήνα, Σεπτέμβριος 2009,

4η Έκδοση (v.4.2 τελευταία ενημέρωση 23/9/2009)

ΠΡΟΛΟΓΟΣ

Οι σημειώσεις αυτές αποτελούν μια σύντομη εισαγωγή στις βασικές έννοιες της Επιδημιολογίας και Βιοστατιστικής. Η πρώτη πρόχειρη έκδοση των σημειώσεων αυτή έγινε το Μάρτιο του 2005 ενώ κάθε χρόνο γίνονταν μικρές διορθώσεις και προσθήκες νέου υλικού. Θα ήθελα να ευχαριστήσω θερμά το φοιτητή Στάθη Σουφλέρη για τις σημαντικές διορθώσεις και επισημάνσεις του στην 3η έκδοση των σημειώσεων.

Αθήνα, 16 Σεπτεμβρίου 2008

Ι.Νιζούφρας

Περιεχόμενα

1	Εισαγωγή	1
1.1	Τι είναι Επιδημιολογία	1
1.2	Ιστορική Αναδρομή.	2
1.3	Η Επιστήμη της Βιοστατιστικής.	3
2	Ειδικές Ιατρικές και Επιδημιολογικές Μελέτες	7
2.1	Κατηγορίες Ιατρικών Μελετών.	7
2.2	Πειραματικές Μελέτες και Μελέτες Παρατήρησης	8
2.2.1	Πειραματικές Μελέτες	8
2.2.2	Μελέτες Παρατήρησης	8
2.3	Περιγραφικές και Αιτιολογικές Μελέτες	9
2.3.1	Περιγραφικές Μελέτες	9
2.3.2	Αιτιολογικές Μελέτες ή Μελέτες Επαγωγικής Επιδημιολογίας	10
2.4	Σχέσεις Μεταβλητών και Αιτιότητα	10
2.4.1	Κριτήρια Αξιολόγησης Αιτιολογικών Σχέσεων	11
2.4.2	Δευτερογενής Πλασματική Σχέση δύο Μεταβλητών	13
2.4.3	Άμεσες και Έμμεσες Αιτιολογικές σχέσεις	14
2.5	Αιτιολογικές Μελέτες Παρατήρησης	14
2.5.1	Συγχρονικές - Διατμηματικές Μελέτες	14
2.5.2	Προοπτικές Μελέτες	15
2.5.2.1	Γενικές Έννοιες	15
2.5.2.2	Διαμόρφωση και Επιλογή των υπό Σύγκριση Ομάδων	18
2.5.2.3	Προβλήματα στην Παρακολούθηση	19
2.5.3	Αναδρομικές Μελέτες ή Μελέτες Μαρτύρων-Ασθενών	19
2.5.3.1	Γενικές Έννοιες	19
2.5.3.2	Στάδια Σχεδιασμού	20
2.5.3.3	Πλεονεκτήματα και Μειονεκτήματα	22

3 Δείκτες Νοσηρότητας, Μέτρα Κινδύνου και Διαγνωστικού Ελέγχου	23
3.1 Μέτρα Νοσηρότητας - Εμφάνιση μιας Νόσου	23
3.1.1 Επιπολασμός	23
3.1.2 Αθροιστική Επίπτωση	24
3.1.3 Ρυθμός Επίπτωσης	25
3.1.4 Διάρκεια Ασθένειας	25
3.2 Μέτρα Κινδύνου για Δίτιμα - Κατηγορικά Δεδομένα.	26
3.2.1 Εισαγωγικές Έννοιες για 2×2 Πίνακες Συνάφειας.	26
3.2.1.1 Πίνακες Συνάφειας	26
3.2.1.2 Δεσμευμένες Πιθανότητες	27
3.2.1.3 Ανεξαρτησία Δύο Κατηγορικών Μεταβλητών	28
3.2.2 Οφειλόμενος ή Αποδιδόμενος Κίνδυνος	29
3.2.3 Σχετικός Κίνδυνος (Relative Risk, RR)	29
3.2.4 Λόγος Σχετικών Πιθανοτήτων (Odds Ratio, OR)	31
3.2.4.1 Σχετική Πιθανότητα (odds)	31
3.2.4.2 Σύγκριση Σχετικών πιθανοτήτων με τη χρήση του Odds Ratio. . .	33
3.2.5 Εκτίμηση του σχετικού κινδύνου και του ΛΣΠ με τη Χρήση SPSS.	36
3.3 Σχέση Σχετικού Κινδύνου και ΛΣΠ (Odds Ratio)	36
3.4 Υπολογισμός του Σχετικού Κινδύνου στις Μελέτες Μαρτύρων-Ασθενών	37
3.5 Παραδείγματα	38
3.5.1 Μυοκαρδιακή Ανεπάρκεια και Αντισύλληψη (Μελέτη Μαρτύρων-Ασθενών).	38
3.5.1.1 Εισαγωγή των Δεδομένων στο SPSS.	39
3.5.1.2 Από κοινού και δεσμευμένες Συναρτήσεις Πιθανότητας.	39
3.5.1.3 Εκτίμηση ΛΣΠ και Συνάρτησης Κινδύνου.	41
3.5.2 Επίδραση Καπνίσματος στην Επιβίωση σε Ασθενείς με ένα επεισόδιο Καρδιακής ανακοπής (Προοπτική Μελέτη)	42
3.5.3 Παράδειγμα 3: Άσκηση για το Σπίτι.	44
3.6 Συμπερασματολογία και Έλεγχοι Υπόθεσης για 2×2 Πίνακες Συνάφειας	45
3.6.1 Ισότητα Δύο Διωνυμικών Ποσοστών	45
3.6.1.1 Αποδιδόμενος κίνδυνος	45
3.6.1.2 Σχετικός κίνδυνος.	49
3.6.1.3 Λόγος Σχετικών Πιθανοτήτων.	52
3.6.1.4 Παράδειγμα 3.2 (συνέχεια)	56
3.6.2 Έλεγχος Ανεξαρτησίας χ^2 του του Pearson.	58
3.6.3 Έλεγχος Μεγίστης Πιθανοφάνειας.	59

3.6.4	Ακριβής Έλεγχος Ανεξαρτησίας του Fisher.	61
3.6.5	Μόντε Κάρλο Έλεγχοι Ανεξαρτησίας.	63
3.7	Διαγνωστικοί έλεγχοι	64
3.7.1	Ευαισθησία και Ειδικότητα.	64
3.7.2	Ευαισθησία και ειδικότητα σε 2×2 πίνακες	67
3.7.3	Καμπύλες ROC	68
3.7.4	Υπολογισμός Επιπολασμού από Διαγνωστικές Μελέτες.	69
4	Κλινικές Δοκιμές (Clinical Trials)	71
4.1	Εισαγωγή και Γενικές Έννοιες	71
4.2	Σύντομη Ιστορική Αναδρομή.	72
4.3	Φάσεις Κλινικών Δοκιμών	73
4.3.1	Μελέτες Φάσης I	73
4.3.2	Μελέτες Φάσης II	75
4.3.3	Μελέτες Φάσης III.	76
4.3.4	Μελέτες Φάσης IV.	76
4.4	Τύποι Κλινικών Δοκιμών	76
4.5	Τυχαιοποίηση	77
4.6	Λεπτομέρειες Μελετών Φάσης III	77
4.6.1	Προσχεδιασμός της μελέτης	77
4.6.2	Ορισμός Ασθενών	80
4.6.3	Ορισμός θεραπευτικών αγωγών	80
4.6.4	Εκτίμηση και καθορισμός μεταβλητών απόκρισης	81
4.6.5	Εισαγωγή στη Μελέτη και Ανάθεση Αγωγών	81
4.7	Άλλες λεπτομέρειες Κλινικών Δοκιμών	82
4.7.1	Μέτα-ανάλυση	82
4.8	Πειραματικοί Σχεδιασμοί	83
4.8.1	Απλή τυχαιοποίηση	83
4.8.2	Στρωματοποιημένη τυχαιοποίηση	84
5	Αδρανοποίηση των Επιδράσεων Συγχυτικών Παραγόντων	85
5.1	Συγχυτικοί Παράγοντες	85
5.1.1	Τρόποι Αδρανοποίησης των Επιδράσεων των Συγχυτικών Παραγόντων	88
5.2	Ανάλυση σε Μελέτες με Εξομοιωμένα/Εξαρτημένα Δείγματα Δίτιμων Μεταβλητών	88

5.2.1 Έλεγχος Αναλογιών για Δύο Εξαρτημένα Δείγματα Δίτιμων Κατηγορικών Μεταβλητών	88
5.2.2 Ο Δείκτης Συμφωνίας Κάπα (Kappa).	93
5.3 Προτυποποίηση	96
5.4 Συμπερασματολογία για Στρωματοποιημένα Κατηγορικά Δεδομένα	96
5.4.1 Διαφοροποίηση Επιδράσεων (Effect Modification).	99
5.4.2 Εκτίμηση Κοινού ΛΣΠ για Στρωματοποιημένα Δεδομένα.	100
5.4.3 Εφαρμογή Στρωματοποιημένης Ανάλυσης με τη χρήση του SPSS.	101
5.4.4 Εκτίμηση ΛΣΠ σε Μελέτες εξομοιωμένων ζευγών.	103
6 Υπολογισμός Μεγέθους Δείγματος	107
6.1 Υπολογισμός Μεγέθους Δείγματος για 2 Διωνυμικά Ποσοστά	107
6.1.1 Ανεξάρτητα Δείγματα	107
6.1.2 Εξαρτημένα Δείγματα	108
6.1.3 Μέγεθος δείγματος σε Κλινικές Δοκιμές: Ανεξάρτητα Δείγματα, Δίτιμες Μεταβλητές με Αλλαγές στην Θεραπεία των Ασθενών	109
7 Λογιστική Παλινδρόμηση (Logistic Regression)	113
7.1 Εισαγωγή	113
7.2 Η λογιστική παλινδρόμηση	114
7.2.1 Εισαγωγή	114
7.2.2 Η Λογιστική Συνάρτηση	116
7.2.3 Το λογιστικό μοντέλο	117
7.3 Ερμηνεία Παραμέτρων	118
7.3.1 Παλινδρόμηση με μια ανεξάρτητη κατηγορική μεταβλητή	118
7.3.2 Παλινδρόμηση με μια ανεξάρτητη συνεχή μεταβλητή	119
7.3.3 Παλινδρόμηση με περισσότερες ανεξάρτητες μεταβλητές	120
7.4 Μέθοδοι εκτίμησης	123
7.4.1 Συνάρτηση πιθανοφάνειας	123
7.4.2 Μέθοδος μεγίστης πιθανοφάνειας	125
7.4.3 Μέθοδος Επαναλαμβανόμενων Σταθμισμένων Ελαχίστων Τετραγώνων	127
7.5 Έλεγχοι υποθέσεων	127
7.5.1 Έλεγχος λόγου πιθανοφανειών	128
7.5.2 Έλεγχοι Wald	133
7.5.3 Διαστήματα εμπιστοσύνης	134

7.6	Στρατηγικές μοντελοποίησης	135
7.7	Αλληλεπιδράσεις και συγχυτικοί παράγοντες	135
7.8	Πολύτομη λογιστική παλινδρόμηση	135
7.9	Ordinal	135
8	Ανάλυση Επιβίωσης (Survival Analysis)	137
8.1	Εισαγωγή	137
8.2	Βασικές συναρτήσεις	138
8.2.1	Η συνάρτηση επιβίωσης	139
8.2.2	Η συνάρτηση κινδύνου	140
8.2.3	Μέσος υπολειπόμενος χρόνος	142
8.3	Λογοκρισία	143
8.4	Μη παραμετρική εκτίμηση συναρτήσεων	147
8.4.1	Εκτιμητής για τη συνάρτηση επιβίωσης	147
8.4.2	Ανάλυση δεδομένων με τη χρήση του SPSS	150
8.4.2.1	Η συνάρτηση κινδύνου	152
8.5	Έλεγχοι υποθέσεων	154
8.6	Μοντέλο αναλογικών κινδύνων	156
8.6.1	Εκτίμηση των παραμέτρων	158
8.6.2	Έλεγχοι υποθέσεων	159
8.6.3	Εκτίμηση της βασικής συνάρτησης κινδύνου και της συνάρτησης επιβίωσης	164
8.7	Επεκτάσεις - Ειδικά θέματα	165
8.7.1	Στρωματοποιημένα μοντέλα	166
8.8	Άλλες προσεγγίσεις	167

Κεφάλαιο 1

Εισαγωγή

1.1 Τι είναι Επιδημιολογία

Ως Επιδημιολογία ορίζουμε την Επιστήμη που μελετάει την κατανομή και της εξέλιξη διαφόρων νοσημάτων ή χαρακτηριστικών στον ανθρώπινο πληθυσμό και των παραγόντων που τον διαμορφώνουν ή τον επηρεάζουν. Επίσης εξετάζει χαρακτηριστικά που συνδέονται τόσο με τον περιβαλλοντικό όσο και με το γενετικό έλεγχο. Αντικείμενο της Επιδημιολογίας είναι και η περιγραφική κατανομή των παραγόντων που επηρεάζουν τον πληθυσμό όπως για παράδειγμα ηλικία, φύλο, φυλή, γεωγραφική περιοχή και επάγγελμα (περιγραφική επιδημιολογία). Επιπλέον οι αιτίες που προκαλούν ασθένειες ή αλλαγές στον ανθρώπινο πληθυσμό είναι αντικείμενο της «επαγωγικής» ή «αναλυτικής Επιδημιολογίας». Στην Επιδημιολογία κάθε νόσος περιγράφεται κυρίως από τη συχνότητα εμφάνισης, το χρόνο διάρκειας και την εξέλιξη της.

Ως «Επιδημία» μπορούμε να ορίσουμε μία νόσο όταν η συχνότητα εμφάνισης της είναι πολύ μεγαλύτερη από τη συνηθισμένη. Έτσι αν μ είναι η μέση εμφάνιση μιας νόσου, τότε μπορούμε να χρησιμοποιήσουμε την Poisson κατανομή (εφόσον έχουμε συχνότητες) για να ορίσουμε ένα άνω όριο επιτρεπτών τιμών, που δίνεται από την ποσότητα $\mu + 3\sqrt{\mu}$, πέρα από το οποίο λέμε ότι έχουμε την εμφάνιση μιας επιδημίας.

Στόχοι της Επιδημιολογίας είναι οι ακόλουθοι:

1. Η μέτρηση και περιγραφή της νοσηρότητας και θνησιμότητας ενός πληθυσμού.
2. Ο έλεγχος της χρονικής εξέλιξης μίας νόσου.
3. Η ανακάλυψη αιτιολογικών παραγόντων που προκαλούν μια νόσο (επαγωγική Επιδημιολογία).

4. Η μελέτη των συνθηκών και των αιτιών που προκαλούν επιδημίες (monitoring)· για παράδειγμα κάπνισμα και καρκίνος των πνευμόνων.
5. Η κατανόηση της ιστορίας των νοσημάτων (για πρόληψη και έλεγχο-screening).
6. Η ταξινόμηση των νοσημάτων.
7. Ο προγραμματισμός, οργάνωση και αξιολόγηση των υπηρεσιών υγείας.

1.2 Ιστορική Αναδρομή.

Η επιδημιολογία διαμορφώθηκε ως επιστήμη τον μεσαίωνα για την μελέτη μεγάλων επιδημιών (χολέρα, ευλογιά, πανώλη). Ο πρώτος που ασχολήθηκε με τα αντικείμενα της «Επιδημιολογίας» είναι ο Ιπποκράτης (460-357 π.Χ.) στο «Περί αέρος, ύδατος και τόπων». Εκεί κατέγραψε τις εμπειρικές σχέσεις μεταξύ συγκεκριμένων ασθενειών και του τόπου εκδήλωσης ή εμφάνισης, των συνθηκών διαβίωσης, της διατροφής, κατοικίας, κλίματος και άλλων αιτιών. Όλα αυτά αποτελούν το αντικείμενο της σύγχρονης περιγραφικής Επιδημιολογίας.

Πολλά χρόνια αργότερα, ο Graunt (1620-1674) χρησιμοποίησε ποσοτικές μεθόδους για τη μελέτη της θνησιμότητας αγοριών και τις ιδιομορφίες της βρεφικής θνησιμότητας. Ο Farr (1807-1883) έβαλε τις βάσεις για την επιδημιολογική χρήση δημογραφικών δεδομένων εργαζόμενος ως υπεύθυνος Ιατρικών θεμάτων στη στατιστική υπηρεσία της Αγγλίας. Την ίδια εποχή, ο Snow (1813-1858) απέδειξε ότι η χολέρα προερχόταν από μικρόβιο που μεταδιδόταν από το νερό και την κοπριά. Στο αποτέλεσμα αυτό κατέληξε αφού μελέτησε το νερό από την κεντρική ύδρευση δύο περιοχών ίδιων χαρακτηριστικών (ηλικία, φύλο, κοινωνικοοικονομική κατάσταση): το Lambeth (με καθαρό νερό) και το Soutwork (όπου το νερό περιείχε απόβλητα). Το Lambeth είχε 8 φορές μικρότερη θνησιμότητα από χολέρα. Άρα το νερό ήταν το κλειδί και όχι ο αέρας όπως πίστευαν εκείνη την εποχή.

Στον 20ο αιώνα είχαμε σημαντικές εξελίξεις στην Επιδημιολογία. Μια από τις πιο σημαντικές είναι η σύνδεση του καπνίσματος με τον καρκίνο του πνεύμονα από τον Doll· για λεπτομέρειες βλ. Doll and Peto (1976). Πολύ σημαντικό γεγονός είναι και η δημιουργία Σχολή Δημόσιας Υγιεινής το 1922 στο Πανεπιστήμιο του Harvard η οποία περιλαμβάνει τμήματα Επιδημιολογίας και Βιοστατιστικής. Η σχολή αυτή συνέβαλε στον καθορισμό και στην προώθηση της επιστήμης της Επιδημιολογίας και καθιερώθηκε ως ένα από τα καλύτερα Σχολεία στον τομέα αυτό. Για πιο λεπτομερή ανάλυση παραπέμπουμε στα βιβλία του Τριχόπουλου (1982, σελ.4-7) και των McMahon & Trichopoulos (1996).

1.3 Η Επιστήμη της Βιοστατιστικής.

Ως Βιοστατιστική ορίζουμε τον εξειδικευμένο τομέα της Στατιστικής που ασχολείται αποκλειστικά με προβλήματα που εμφανίζονται στην ιατρική έρευνα.

Ένας τρόπος να καταλάβουμε την τη σημαντικότητα της Βιοστατιστικής είναι να ακολουθήσουμε μια μελέτη από τον σχεδιασμό της μέχρι την ολοκλήρωση της (συνήθως αποτελέσματα δημοσιευμένα).

Παράδειγμα 1.1 : Μελέτη σύγκρισης μηχανών και παραδοσιακών μετρήσεων. Στο παράδειγμα αυτό θέλουμε να συγκρίνουμε αυτόματα μηχανήματα μέτρησης πίεσης (*blood pressure*) που βρίσκονται στο δρόμο στις ΗΠΑ με τον παραδοσιακό τρόπο υπό την παρουσία ειδικού (συνήθως νοσηλεύτη). Η μελέτη αυτή έγινε στις ΗΠΑ και για λεπτομέρειες παραπέμπουμε στο βιβλίο του Rosner (1994, σελ. 1-4). Στη μελέτη αυτή για κάθε άτομο μετράμε την πίεση του με το μηχάνημα και για το ίδιο άτομο μετράμε την πίεση και με τον παραδοσιακό τρόπο με τη βοήθεια ενός ειδικευμένου νοσηλεύτη.

Σχεδιασμός της Μελέτης : Για σχεδιάσουμε την μελέτη πρέπει να απαντήσουμε στα ακόλουθα ερωτήματα :

1. Πόσα μηχανήματα θα εξετάσουμε και που θα είναι αυτά ;
2. Πόσους ανθρώπους θα εξετάσουμε ανά μηχανήματα ;
3. Ποια θα είναι η των σειρά μετρήσεων ;
4. Άλλα δεδομένα που θα πρέπει να μαζέψουμε ;
5. Πως θα γίνει η καταχώρηση και η κωδικοποίηση των δεδομένων ;
6. Πως θα γίνει ο έλεγχος της ακρίβειας και της ακεραιότητας των δεδομένων ;

Απαντήσεις στα ερωτήματα (σύμφωνα με τη μελέτη του Rosner, 1994):

1. Επιλέχθηκαν 4 μηχανήματα σε μη κοντινές περιοχές.
2. Χρησιμοποιήθηκαν στατιστικοί μέθοδοι καθορισμού του μεγέθους δείγματος.
3. Συνήθως η πρώτη μέτρηση συνήθως είναι πιο υψηλή. Για το λόγο αυτό αν ορίσουμε την πρώτη μέτρηση να λαμβάνεται συστηματικά από το μηχανήματα ή από τον ειδικευμένο νοσηλεύτη τότε θα δημιουργηθεί ένα συστηματικό σφάλμα που θα οφείλεται στη σειρά δειγματοληψίας και όχι στην πραγματική διαφορά των δύο μετρήσεων. Αυτό μπορεί να αποφευχθεί μέσω της τυχαιοποίησης (ανάθεση της σειράς με τυχαίο τρόπο).

4. Σίγουρα θα πρέπει να ρωτήσουμε τους συμμετέχοντες και επιπλέον ατομικά τους στοιχεία που επηρεάζουν την πίεση όπως βάρος, ύψος, body-mass index και φύλο. Επίσης άλλα δημογραφικά στοιχεία (π.χ. ηλικία, τόπος κατοικίας) ή ιατρικά στοιχεία (π.χ. ιστορικό υπέρτασης) που θα μας δώσουν πληροφόρηση ως προς τον πληθυσμό που χρησιμοποιεί τα μηχανήματα.
5. Στην κωδικοποίηση γίνεται πλήρης ορισμός των ονομάτων των μεταβλητών, των κωδικών των κατηγοριών στις ποιοτικές μεταβλητές, στις ελλειπείς τιμές (missing values). Επίσης τα δεδομένα εισήχθησαν 2 φορές και έγινε αυτόματη σύγκριση των αρχείων για εντοπισμό λαθών πληκτρολόγησης.
6. Στον έλεγχο των δεδομένων έγινε απλή περιγραφική ανάλυση, εκτύπωση ελαχίστων και μεγίστων τιμών ανά μεταβλητή και τυχαίος δειγματοληπτικός έλεγχος μερικών ερωτηματολογίων.

Στατιστική Ανάλυση.

1. **Περιγραφική Ανάλυση και Διαγραμματική Απεικόνιση.** Ενδεικτικά δίνουμε τον ακόλουθο πίνακα ανά τοποθεσία και ανάλογα με τον τρόπο μέτρησης:

Τοποθεσία	Αριθμός Ατόμων	Μηχάνημα		Ειδικός Νοσηλευτής		Διαφορά	
		Μέσος	Τυπ.Απόκλ.	Μέσος	Τυπ.Απόκλ.	Μέσος	Τυπ.Απόκλ.
A	98	142.5	21.0	142.0	18.1	0.5	11.2
B	84	134.1	22.5	133.6	23.2	0.5	12.1
Γ	98	147.9	20.3	133.9	18.3	14.0	11.7
Δ	62	135.4	16.7	128.5	19.0	6.9	13.6

Τα αρχικά δεδομένα ήταν υπολογισμένα να αποδώσουν 100 τιμές ανά τοποθεσία, αλλά όπως παρατηρείται υπήρξαν ελλειπείς τιμές (missing values), γεγονός συνηθισμένο στην Ιατρική Έρευνα και τη Βιοστατιστική. Ο αναμενόμενος αριθμός μηαπόκρισης πρέπει και λαμβάνεται υπ' όψιν στον καθορισμό του μεγέθους δείγματος.

2. **Επαγωγική Στατιστική.** Μας ενδιαφέρει να δούμε τις διαφορές μεταξύ ανθρώπου και μηχανής. Επειδή έχουμε ζευγάρια ποσοτικών μετρήσεων για κάθε άτομο (η πίεση με το μηχάνημα και με τον ειδικό) για το λόγο αυτό ελέγχουμε τη σημαντικότητα της διαφορά των δύο μετρήσεων με τη χρήση της δοκιμασίας t ανά ζεύγη (paired t-test).

Από την παραπάνω περιγραφή μια απλής σχετικά μελέτης βλέπουμε τις δυσκολίες, τα προβλήματα και τις απαιτήσεις της Ιατρικής έρευνας από τους στατιστικούς επιστήμονες. Πολλοί

επιστήμονες της στατιστικής δυσανασχετούν για το διαχωρισμό της Βιοστατιστικής από τη Στατιστική με την έννοια ότι ένας στατιστικός μπορεί να χειριστεί και Ιατρικά προβλήματα. Όμως η πράξη στις χώρες της Ευρώπης και τις ΗΠΑ έδειξε ότι χρειάζονται εξειδικευμένες γνώσεις (τόσο στην ορολογία όσο και στην ανάλυση) που ένας στατιστικός δεν γνωρίζει. Αυτό, σε συνδυασμό με αυξημένη ζήτηση Βιοστατιστικών επιστημόνων από Ιατρικές και φαρμακευτικές εταιρείες, οδήγησε στη δημιουργία εντατικών Μεταπτυχιακών προγραμμάτων Βιοστατιστικής ή Ιατρικής στατιστικής σε πολλά πανεπιστήμια της Ευρώπης και των ΗΠΑ εκ των οποίων αρκετά είναι άμεσα συνδεδεμένα με τη διαδικασία της παραγωγής και της Ιατρικής έρευνας.

Κεφάλαιο 2

Ειδικές Ιατρικές και Επιδημιολογικές Μελέτες

2.1 Κατηγορίες Ιατρικών Μελετών.

Οι Ιατρικές μελέτες χωρίζονται στις ακόλουθες κατηγορίες

1. «Πειραματικές» ή «παρεμβατικές» μελέτες (experimental - intervention studies). Εδώ συμπεριλαμβάνονται και οι συνηθισμένες κλινικές δοκιμές (clinical trials).
2. «Παρατήρησης» ή «μη πειραματικές» ή «μη παρεμβατικές» μελέτες (observational medical surveys). Αυτές με τη σειρά τους χωρίζονται σε
 - (α') «Περιγραφικής» ή «Διατμηματικές-Συγχρονικές» (Descriptive surveys, cross-sectional studies).
 - (β') «Αναλυτικής» ή «περιγραφικής» Επιδημιολογίας (analytic/aetiologic studies). Αυτές με τη σειρά τους χωρίζονται σε δύο μεγάλες κατηγορίες:
 - i. «Προπτικές Μελέτες» ή «μελέτες κοορτής» (prospective or cohort studies). Εναλλακτικές ονομασίες είναι «διαμηκείς-διαχρονικές μελέτες» (longitudinal studies) ή «μελέτες παρακολούθησης» (follow-up studies).
 - ii. «Αναδρομικές μελέτες» ή «μελέτες Μαρτύρων-Ασθενών» (retrospective or case-control studies).

Αναλυτική παρουσίαση των παραπάνω κατηγοριών ακολουθεί στις επόμενες ενότητες.

2.2 Πειραματικές Μελέτες και Μελέτες Παρατήρησης

Ο πρώτος μεγάλος διαχωρισμός των μελετών γίνεται ανάλογα αν ο ερευνητής παρεμβαίνει στη διαδικασία ή όχι. Έτσι, στις πειραματικές μελέτες ο ερευνητής παρεμβαίνει στο πείραμα με την έννοια ότι ελέγχει την κατανομή των ατόμων μέσα στις ομάδες (συνήθως αναθέτει τυχαία θεραπείες σε κάθε άτομο που συμμετέχει). Αντίθετα, στις μελέτες παρατήρησης ο ερευνητής δεν παρεμβαίνει, απλά παρατηρεί και καταγράφει. Σε αυτή την περίπτωση κυρίως σκοπός είναι η διερεύνηση των αιτιών που προκαλούν μία νόσο και η κατασκευή διαγνωστικών-προγνωστικών μοντέλων.

2.2.1 Πειραματικές Μελέτες

Στις μελέτες αυτές ο ερευνητής παρεμβαίνει ενεργητικά στον καθορισμό των ομάδων, του τρόπου διαβίωσης, της θεραπείας ή άλλων χαρακτηριστικών (π.χ. κλινικές δοκιμές). Αυτές οι μελέτες χωρίζονται σε τέσσερις υποκατηγορίες:

1. *Εκούσια και σχεδιασμένη εφαρμογή ενός πιθανού αιτιολογικού μοντέλου σε μια ομάδα και παρακολούθηση των αποτελεσμάτων.* Συνήθως γίνεται σύγκριση με μια ομάδα αναφοράς ή ελέγχου. Αυτή η ομάδα ορίζεται όσο το δυνατόν όμοια με την ήδη υπάρχουσα όσον αφορά κάποια χαρακτηριστικά του πληθυσμού. Αυτό επιτυγχάνεται με την *τυχαιοποίηση*, την τυχαία, δηλαδή, κατανομή ή ανάθεση των ατόμων στις ομάδες σύγκρισης.
2. *Εκούσια απομάκρυνση ενός παράγοντα.* Αυτές είναι δεοντολογικά και ηθικά πιο σωστές από τις παραπάνω. Προσοχή η απομάκρυνση πρέπει να γίνεται τυχαία, κάτι που είναι πολύ διαφορετικό από το να αφήσουμε τους συμμετέχοντες να αποφασίσουν σε ποια ομάδα επιθυμούν να συμμετάσχουν.
3. *Πειράματα φύσης.* Μελετούνται διαφορετικές ομάδες που δημιουργήθηκαν από ένα γεγονός.
4. *Σχεδιασμένη τροποποίηση παθογενών μηχανισμών σε μια ομάδα ανθρώπων.*

2.2.2 Μελέτες Παρατήρησης

Ήδη, είπαμε ότι σε αυτές τις μελέτες ο ερευνητής δε συμμετέχει αλλά απλά παρατηρεί και καταγράφει. Χωρίζονται σε τρεις μεγάλες κατηγορίες: τις διατμηματικές-συγχρονικές μελέτες οι οποίες είναι περιγραφικές κυρίως μελέτες και δε λαμβάνουν υπόψιν το παράγοντα χρόνο, τις προοπτικές μελέτες οι οποίες παρακολουθούν την πορεία μιας ομάδας ατόμων (κινούνται

μπροστά χρονικά) και τις αναδρομικές μελέτες οι οποίες πάνε πίσω χρονικά και ζητούν από τα άτομα να περιγράψουν συνήθειες τους στο παρελθόν. Αναλυτική παρουσίαση των μελετών αυτών ακολουθεί στις ενότητες 2.5.1, 2.5.2 και 2.5.3.

2.3 Περιγραφικές και Αιτιολογικές Μελέτες

Ο δεύτερος μεγάλος διαχωρισμός των Ιατρικών μελετών γίνεται ανάλογα με το αν ο σκοπός της μελέτης είναι να περιγράψουμε απλά ένα πληθυσμό ή αν θέλουμε να κατανοήσουμε τις σχέσεις μεταξύ χαρακτηριστικών και να εντοπίσουμε τα αίτια που προκαλούν ένα φαινόμενο, συνήθως μία ασθένεια. Ανάλογη είναι και η στατιστική ανάλυση που χρησιμοποιείται αλλά και ο σχεδιασμός της δειγματοληψίας. Οι περιγραφικές μελέτες περιορίζονται σε απλούς περιγραφικούς δείκτες ή δημογραφικούς πίνακες και σε διαγραμματική απεικόνιση των δεδομένων ενώ οι αιτιολογικές εφαρμόζουν στατιστικούς ελέγχους υποθέσεων βασισμένους σε βιολογικές θεωρίες ή σενάρια. Επιπλέον, είναι συνηθισμένη η χρήση πολύπλοκων στατιστικών υποδειγμάτων. Στη σύγχρονη Ιατρική επιστήμη ελάχιστες μελέτες είναι πλέον μόνο περιγραφικές.

2.3.1 Περιγραφικές Μελέτες

Όπως είπαμε και παραπάνω οι περιγραφικές μελέτες έχουν σκοπό απλά να περιγράψουν τον πληθυσμό σε ένα συγκεκριμένο τόπο και χρόνο. Για το λόγο αυτό χρησιμοποιούνται οι συνηθισμένοι τρόποι δειγματοληψίας όταν έχουμε μεγάλους πληθυσμούς ή γίνεται απογραφικά όταν ο πληθυσμός είναι περιορισμένος (συνήθως στις μελέτες δημογραφική φύσης). Η μελέτες αυτές ονομάζονται και διατμηματικές (cross-sectional). Όταν ο σκοπός περιορίζεται στην εκτίμηση του επιπολασμού (αναλογία ατόμων με τη νόσο στον πληθυσμο) τότε λέγονται και μελέτες δημογραφικές ή επιπολασμού (prevalence). Όπως είπαμε και παραπάνω, χρησιμοποιούν απλή μεθοδολογία και δίνουν μια ενδεικτική εικόνα του προβλήματος που μας ενδιαφέρει.

Οι μελέτες αυτές γίνονται σε πληθυσμό υπό φυσιολογικές συνθήκες και ονομάζονται και μελέτες πεδίου (field surveys). Παρουσιάζουν μια εκτίμηση ή αποτύπωση του επιπολασμού ή άλλων χαρακτηριστικών του πληθυσμού με στιγμιαίο και στατικό χαρακτήρα. Ένα πολύ σημαντικό στοιχείο είναι ότι μπορούν να δώσουν μια γενική εικόνα των σχέσεων μεταξύ ασθενειών κάτι που δεν μπορεί να γίνει με ενδοοσοκομιακές έρευνες.

Τα κύρια πλεονεκτήματα αυτών των μελετών είναι: η απλότητα τους, η ταχύτητα με την οποία μπορούν να γίνουν και το χαμηλό κόστος. Από την άλλη στα μειονεκτήματα συγκαταλέγονται: το γεγονός ότι δε λαμβάνουμε υπόψη τον παράγοντα του χρόνου (άρα δεν μπορεί να οριστεί η χρονική αλληλουχία των σχέσεων μεταξύ μεταβλητών), δεν μπορεί να γίνει δια-

χωρισμός της πιθανότητας εμφάνισης και της διάρκειας της νόσου και σπάνιες ασθένειες δεν μπορούν να μελετηθούν διότι το μέγεθος του δείγματος είναι συνήθως ανεπαρκές (άρα πρέπει να περιοριστούμε σε πιο κοινές ασθένειες).

2.3.2 Αιτιολογικές Μελέτες ή Μελέτες Επαγωγικής Επιδημιολογίας

Οι μελέτες αυτές αποτελούν τον κορμό της σύγχρονης Επιδημιολογίας. Αναφέρονται στις μελέτες που σκοπός τους είναι ο έλεγχος μιας επιδημιολογικής υπόθεσης και αποσκοπούν στην διερεύνηση της αιτιότητας ενός νοσήματος. Στις μελέτες αυτές χρησιμοποιούνται εξελιγμένες στατιστικές μέθοδοι ξεκινώντας από συνηθισμένες στατιστικές δοκιμασίες (για παράδειγμα t-test ή Wilcoxon test) μέχρι και πολύπλοκα υποδείγματα όπως τα ιεραρχικά γενικευμένα γραμμικά υποδείγματα (hierarchical generalised linear models) και μοντέλα πολυμεταβλητής ανάλυσης και Δομικών Εξισώσεων (structural equation models) στην Ψυχομετρία.

Στην κατηγορία αυτή μπορούμε να συμπεριλάβουμε τόσο μελέτες παρατήρησης (όπως οι προοπτικές και οι αναδρομικές μελέτες) όσο και παρεμβατικές μελέτες (κλινικές δοκιμές).

Κύριος σκοπός στις παραπάνω μελέτες παρατήρησης είναι να εντοπίσουν παράγοντες κινδύνου που αυξάνουν την εμφάνιση μιας νόσου (δηλαδή ο σκοπός είναι διαγνωστικός και προληπτικός) ενώ στις κλινικές δοκιμές σκοπός είναι να εντοπιστούν νέα φάρμακα που περιορίζουν τα συμπτώματα μίας νόσου ή οδηγούν στην πλήρη ίαση (θεραπεία).

2.4 Σχέσεις Μεταβλητών και Αιτιότητα

Το μεγάλο πρόβλημα γενικότερα στη στατιστική επιστήμη είναι το γεγονός ότι μια στατιστικά-σημαντική σχέση (association) ή συσχέτιση μεταξύ δύο μεταβλητών δε συνεπάγεται αυτόματα και μια πραγματική επαγωγική σχέση ή σχέση αιτιότητας. Αυτό συμβαίνει λόγω της πολυπλοκότητας των σχέσεων που υπάρχουν πραγματικά φαινόμενα που εξετάζουμε. Συνεπώς στην εμφάνιση ενός φαινομένου (π.χ. νόσου) συνεπιδρούν πολλοί παράγοντες και για το λόγο αυτό μπορεί η σχέση νόσου-παράγοντα κινδύνου να επηρεάζεται από τρίτους παράγοντες που ονομάζονται συγχυτικοί.

Ορισμός 2.1 Συγχυτικός ή συνεπιδρών παράγοντας (*confounder* ή *confounding variable*) ονομάζεται μια μεταβλητή η οποία, με οποιονδήποτε τρόπο, διαστρεβλώνει μια σχέση μεταξύ δύο μεταβλητών (στην Βιοστατιστική συνήθως εννοούμε την νόσο και ένα παράγοντα κινδύνου). Για παράδειγμα ως υποθέσουμε ότι επιθυμούμε να συγκρίνουμε τη θνησιμότητα σε δύο ομάδες με διαφορετική έκθεση σε ένα πιθανό παράγοντα κινδύνου (π.χ. κάπνισμα). Αν η πρώτη ομάδα περιέχει άτομα

πιο μεγάλης ηλικίας τότε σαφώς και περιμένουμε μεγαλύτερη θνησιμότητα σε αυτή την ομάδα η οποία δε θα οφείλεται όμως στο γεγονός ότι κάπνιζαν απλά απλά στο ότι έχουν μεγαλύτερη ηλικία. Η ηλικία εδώ είναι συγχυτικός παράγοντας (confounder) της σχέσης της θνησιμότητας και του καπνίσματος. Ένας συγχυτικός παράγοντας μπορεί να εμφανίζει ως στατιστικά σημαντικές σχέσεις που στην πραγματικότητα δεν υπάρχουν ή να καλύπτει υπάρχουσες. Ο έλεγχος και αδρανοποίηση των συγχυτικών παραγόντων γίνεται με κατάλληλο σχεδιασμό της μελέτης ή με ειδική στατιστική ανάλυση.

Ακόμα και αν οι συγχυτικοί παράγοντες εξαλειφθούν, ο εντοπισμός πραγματικών σχέσεων-ναιτιότητας είναι δύσκολος και ξεφεύγει από τον απλό υπολογισμό στατιστικών δεικτών. Για το λόγο αυτό έχουν καθοριστεί συγκεκριμένα κριτήρια τα οποία μπορούμε να χρησιμοποιήσουμε για να εντοπίσουμε και να ελέγξουμε αν οι στατιστικά σημαντικές συσχετίσεις είναι σχέσεις αιτιότητας. Τα κριτήρια αυτά είναι τα ακόλουθα:

1. **Συνέπεια** (consistency).
2. **Ισχύς** (strength).
3. **Ειδικότητα** (specificity).
4. **Χρονική αλληλουχία** (temporality).
5. **Λογική αλληλουχία** (coherency).

Όταν μια σχέση ικανοποιεί όλα τα παραπάνω κριτήρια, τότε έχουμε μια ισχυρή (αν όχι απόλυτη) ένδειξη ότι αυτή είναι μία σχέσης αιτιότητας.

2.4.1 Κριτήρια Αξιολόγησης Αιτιολογικών Σχέσεων

α) Συνέπεια μιας σχέσης: Μία σχέση ονομάζεται συνεπής αν εμφανιστεί επαναλαμβανόμενα σε μελέτες που έχουν γίνει με διαφορετικό σχεδιασμό και σε διαφορετικούς πληθυσμούς. Φυσικά στις περιπτώσεις αυτές μπορούν να υπάρχουν και εξαιρέσεις, δηλαδή μελέτες που δεν επιβεβαιώνουν τον κανόνα, αλλά θα πρέπει να είναι περιορισμένες και να οφείλονται σε τυχαίους παράγοντες (σφάλματα τύπου I και II των στατιστικών ελέγχων υποθέσεων). Ένα παράδειγμα συνεπούς σχέσης είναι η θετική σχέση καπνίσματος - καρκίνου του πνεύμονα η οποία έχει επιβεβαιωθεί σε μεγάλο αριθμό μελετών τα τελευταία χρόνια.

β) Ισχύς μιας σχέσης: Μια σχέση ονομάζεται ισχυρή όταν η επίδραση της παρουσίας του παράγοντα κινδύνου είναι μεγάλη. Η επίδραση συνήθως μετριέται με την αναλογική αύξηση της πιθανότητας θανάτου ή εμφάνισης της νόσου. Σε αυτή την περίπτωση, σημαντική είναι και η ύπαρξη της δοσολογικής επίδρασης (dose-response effect) που δεν είναι τίποτα άλλο από την αναλογική μεταβολή της μεταβλητής απόκρισης (π.χ. εμφάνιση της νόσου) και της δοσολογίας ενός φαρμάκου ή του μεγέθους έκθεσης σε ένα παράγοντα κινδύνου.

Παράδειγμα: η πιθανότητα εμφάνισης του καρκίνου του πνεύμονα είναι αυξημένη στους καπνιστές και αυξάνει ανάλογα με τον αριθμό τσιγάρων που καπνίζει κάθε άτομο.

γ) Ειδικότητα μιας σχέσης: Μια σχέση παράγοντα κινδύνου-νόσου ονομάζεται ειδική αν η παρουσία του παράγοντα κινδύνου οδηγεί με μεγάλη πιθανότητα στην εμφάνιση της νόσου ενώ η απουσία του οδηγεί με μεγάλη πιθανότητα στην αποφυγή της εμφάνισης της νόσου. Αν ένας παράγοντας είναι «ειδικός» μιας νόσου (δηλαδή, στην ουσία, είναι η κύρια αιτία που προκαλεί τη νόσο) τότε είναι πολύ πιθανό και η σχέση να είναι αιτιολογική. Στην πράξη όμως, λόγω της πολυπλοκότητας των σχέσεων, ειδικές σχέσεις είναι σπάνιες. Και όχι μόνο αυτό αλλά σε μια πραγματική αιτιολογική σχέση μπορεί η πιθανότητα εμφάνισης της νόσου να αυξάνει σημαντικά με την παρουσία του παράγοντα κινδύνου αλλά όχι τόσο πολύ ώστε να τη χαρακτηρίσουμε ειδική. Η ύπαρξη μιας στατιστικής σχέσης αιτιότητας δε συνεπάγεται ότι αυτή θα επαληθευτεί σε όλες τις περιπτώσεις. Παράδειγμα, αν κάποιος καπνίζει δε σημαίνει ότι θα εμφανίσει τη νόσο σίγουρα (αν συνέβαινε αυτό δε θα μιλάγαμε για στατιστική σχέση αλλά για σχέση φυσική).

δ) Χρονική Αλληλουχία: Σε πολλές περιπτώσεις η στατιστική ανάλυση δε λαμβάνει υπόψη της τη χρονική αλληλουχία ή σειρά με την οποία εμφανίζονται κάποια ενδεχόμενα αλλά και μεταβλητές που συνδέονται με αυτά. Αυτό συνήθως γίνεται με τη χρήση της κοινής λογικής. Για το λόγο αυτό η μεταβλητή απόκρισης στα στατιστικά μοντέλα θα πρέπει να έπεται χρονικά των επεξηγηματικών (κάτι το οποίο δεν είναι πάντα προφανές). Έτσι λοιπόν στην Βιοστατιστική είναι απαραίτητο ο παράγοντας κινδύνου να προϋπάρχει χρονικά της εμφάνισης της νόσου. Σε μερικές περιπτώσεις αυτό είναι εύκολο να το κρίνεις (π.χ. κάπνισμα και καρκίνος του πνεύμονα). Τι γίνεται όμως για περιπτώσεις όπως το παθητικό κάπνισμα για το οποίο μπορεί να μην είναι εύκολα προσδιορίσιμος ο χρόνος έκθεσης σε αυτόν τον κίνδυνο; Ο καθορισμός της χρονικής αλληλουχίας είναι επίσης δύσκολος στις περιπτώσεις που η νόσος εντοπίζεται πολύ αργότερα από την πραγματική της εμφάνιση.

ε) Λογική Αλληλουχία: Για να έχει μια σχέση λογική αλληλουχία θα πρέπει να μην αντικρούει αποδεδειγμένες αλήθειες των φυσικών και βιολογικών επιστημών (π.χ. της Φυσικής, Χημείας,

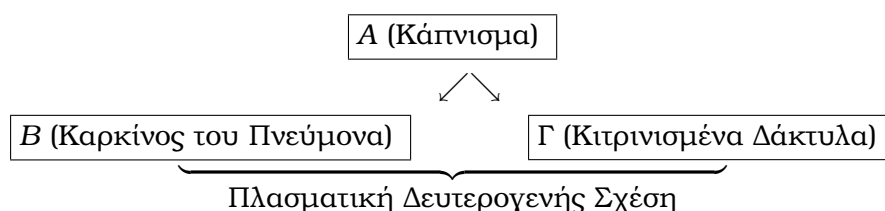
Ιατρικής, Γενετικής και Βιολογίας). Σε περίπτωση που μια στατιστικά σημαντική σχέση έρχεται σε αντίθεση με μια ήδη αναγνωρισμένη θεωρία, τότε θα πρέπει να υπάρξει μια ανάλογη επιστημονική τεκμηρίωση και θεωρία που θα υποστηρίζει και θα αιτιολογεί την ύπαρξη της. Επιπλέον θα πρέπει να γίνουν και ανάλογες μελέτες έτσι ώστε να δούμε αν ικανοποιείται το κριτήριο της συνέπειας που τελικά θα οδηγήσει στην επιστημονική επικράτηση και καθιέρωση του ευρήματος ή της νέας θεωρίας.

Γενικά τα παραπάνω πέντε κριτήρια είναι αρκετά αυστηρά και πολλές φορές δεν ικανοποιούνται όλα από μια πραγματική αιτιολογική σχέση.

2.4.2 Δευτερογενής Πλασματική Σχέση δύο Μεταβλητών

Όπως αναφέραμε και στην προηγούμενη ενότητα, όταν υπάρχει μια πραγματική αιτιολογική σχέση μεταξύ δύο μεταβλητών τότε περιμένουμε να βρεθεί και μια στατιστικά σημαντική σχέση μεταξύ τους. Το αντίστροφο όμως δεν ισχύει πάντα. Δηλαδή μπορεί να έχουμε μια στατιστικά σημαντική σχέση η οποία όμως δεν οφείλεται σε πραγματικά υπαρκτή σχέση.

Ας πάρουμε το ακόλουθο παράδειγμα. Έστω ότι θέλουμε να εξετάσουμε το κάπνισμα σε σχέση με τον καρκίνο του πνεύμονα. Το κάπνισμα προκαλεί αυξημένη εμφάνιση του καρκίνου αλλά και κιτρινίζει τα δάκτυλα σε περίπτωση όπου η δοσολογία και ο χρόνος καπνίσματος είναι πολύ αυξημένος. Αν εμείς απομονώσουμε τις μεταβλητή που υποδεικνύει τα κιτρινισμένα δάκτυλα (αγνοώντας το κάπνισμα που τα προκαλεί) και τη νόσο, τότε είναι πολύ πιθανό να βγάλουμε μια στατιστικά σημαντική σχέση (δηλαδή ότι τα άτομα με κιτρινισμένα δάκτυλα έχουν μεγαλύτερη πιθανότητα εμφάνισης καρκίνου ή ότι ο καρκίνος προκαλεί, ως σύμπτωμα της νόσου, την εμφάνιση κιτρινισμένων δακτύλων). Στην πραγματικότητα η σχέση είναι εικονική - πλασματική και συμβαίνει απλά διότι αυτοί που έχουν κιτρινισμένα δάκτυλα καπνίζουν πολύ περισσότερο σε σχέση με τους υπόλοιπους άρα προφανώς το κάπνισμα είναι που προκαλεί τη νόσο.



Συνεπώς σε μία πλασματική σχέση μεταξύ δύο ανεξάρτητων μεταβλητών B και Γ, οι μεταβλητές αυτές εμφανίζουν στατιστικά σημαντική εξάρτηση λόγω της σχέσης τους με μια μεταβλητή A που είναι η αιτία και των δύο. Αν μεταβληθεί η κατανομή της μεταβλητής Γ δεν είναι απαραίτητο

ότι θα αλλάξει και η κατανομή της μεταβλητής B όπως θα άλλαζε αν η σχέση ήταν πραγματικά αιτιολογική.

Κάθε παράγοντας που συνδέεται στατιστικά σημαντικά και η παρουσία του αυξάνει την πιθανότητα εμφάνισης μιας ασθένειας ονομάζεται παράγοντας κινδύνου (risk factor). Σε αντίθετη περίπτωση ο παράγοντας ονομάζεται προστατευτικός παράγοντας (protective factor). Η παραπάνω μεταβλητή A που επηρεάζει τις μεταβλητές B και Γ, δρα ως συγχυτικός παράγοντας για τη σχέση των B και Γ.

2.4.3 Άμεσες και Έμμεσες Αιτιολογικές σχέσεις

Αν μία μεταβλητή A επηρεάζει την εμφάνιση μιας μεταβλητής Δ και αυτή με τη σειρά της τη μεταβλητή B τότε λέμε ότι οι μεταβλητές A και B συνδέονται μεταξύ τους με έμμεση αιτιολογική σχέση. Αυτή η σχέση μπορεί περιγραφεί από το ακόλουθο απλό διάγραμμα:

$$\boxed{A} \longrightarrow \boxed{\Delta} \longrightarrow \boxed{B} .$$

Παραδοσιακά οι παράγοντες που συνδέονται με τη νόσο με άμεσες αιτιολογικές σχέσεις θεωρούνται πιο επικίνδυνοι και σημαντικοί, ενώ οι έμμεσοι παράγοντες κινδύνου θεωρούνται επικουρικοί ή προδιάθεσης. Ως παράδειγμα μπορούμε να αναφέρουμε την παχυσαρκία η οποία προκαλεί (μεταξύ άλλων) υπέρταση και διαβήτη και αυτά με τη σειρά τους στεφανιαία νόσο. Έτσι λοιπόν η σχέση στεφανιαίας νόσου και παχυσαρκίας είναι έμμεση που σημαίνει ότι αν κάποιος είναι παχύσαρκος αλλά δεν εμφανίζει τις επιμέρους νόσους ή συμπτώματα τότε δε θα έχει αυξημένη πιθανότητα στεφανιαίας νόσου. Παρόλα αυτά, και επειδή η παχυσαρκία συνδέεται ισχυρά με τις επιμέρους νόσους, τα τελευταία χρόνια έγινε μεγάλη εκστρατεία για την καταπολέμηση της.

2.5 Αιτιολογικές Μελέτες Παρατήρησης

2.5.1 Συγχρονικές - Διατμηματικές Μελέτες

Οι μελέτες αυτές δεν λαμβάνουν υπόψη τους τη χρονική αλληλουχία μεταξύ μεταβλητών (temporal relationship). Μελετάνε ένα τυχαίο δείγμα σε μια συγκεκριμένη χρονική στιγμή και σε συγκεκριμένο γεωγραφικό χώρο. Πολλές φορές παρομοιάζεται με φωτογραφία του πληθυσμού (snapshot). Στις μελέτες αυτές χρησιμοποιούνται οι συνηθισμένοι τρόποι δειγματοληψίας (απλή τυχαία, στρωματοποιημένη, συστηματική κλπ.) ενώ σε περιορισμένες περιπτώσεις μπορεί να είναι απογραφική. Οι μελέτες αυτές μπορεί να είναι είτε αιτιολογικές ή απλά περιγραφικές.

Οι διατμηματικές-συγχρονικές μελέτες έχουν ως στόχο τη μελέτη των σχέσεων μεταξύ μεταβλητών (κυρίων ασθενειών και παραγόντων κινδύνου), χωρίς να λαμβάνεται υπόψη η χρονική τους αλληλουχία. Στις μελέτες αυτές μπορούμε να εκτιμήσουμε με ακρίβεια τον επιπολασμό μιας ασθένειας (και σε πολλές από αυτές αυτός είναι ο κύριος στόχος τους). Τέλος, αυτές οι μελέτες μας δίνουν πληροφορίες για την αιτιολογία των ασθενειών αλλά σε περιορισμένο επίπεδο.

Ορισμός 2.2 Επιπολασμός (*prevalence*) μιας νόσου ονομάζεται το ποσοστό/αναλογία ατόμων στον πληθυσμό που έχουν τη νόσο σε μια συγκεκριμένη χρονική στιγμή ή κατά τη διάρκεια ενός μικρού χρονικού διαστήματος.

Οι διατμηματικές-συγχρονικές μελέτες μπορούν να χρησιμοποιηθούν και για συμπερασματολογία που αφορά τα υγιή άτομα που συμμετέχουν στη μελέτη (και συνεπώς για τους πληθυσμούς υγιών που αντιπροσωπεύουν). Αυτά τα αποτελέσματα μπορούν να χρησιμοποιηθούν για να πάρουμε μια εικόνα των αναγκών του συστήματος Υγείας, της στάσης των πολιτών έναντι στο σύστημα υγείας, για το σχεδιασμό πολιτικών υγείας και για την κατάρτιση προγραμμάτων επιμόρφωσης υγείας.

2.5.2 Προοπτικές Μελέτες

2.5.2.1 Γενικές Έννοιες

Για τις προοπτικές μελέτες έχουν χρησιμοποιηθεί αρκετοί διαφορετικοί όροι όπως μελέτες κοορτής (cohort studies), παρακολούθησης (follow-up) και διαχρονικές-διαμήκεις (longitudinal). Η ονομασία κοορτή (cohort) είναι στρατιωτικός όρος και αναφέρεται στο ένα δέκατο μια Ρωμαϊκής λεγεώνας. Στις προοπτικές μελέτες έχει την έννοια της ομάδας ή σειράς ατόμων την οποία παρακολουθούμε κατά διάρκεια του χρόνου. Η ονομασία διαχρονική-διαμήκης (longitudinal) έρχεται σε αντιδιαστολή με τη συγχρονική-διατμηματική (cross-sectional) και σημαίνει μια μελέτη που γίνεται σε πολλά διαφορετικά χρονικά και (μερικές φορές) γεωγραφικά σημεία. Κυρίως όμως αναφέρεται στην επαναλαμβανόμενη χρονικά διενέργεια μιας μελέτης στον ίδιο αριθμό ατόμων (είναι σαν να κάνουμε πολλές συγχρονικές μελέτες). Κύριο τους χαρακτηριστικό είναι οι επαναλαμβανόμενες μετρήσεις στα ίδια άτομα σε διαφορετικές χρονικές στιγμές.

Το βασικό στοιχείο στις προοπτικές μελέτες είναι ότι παίρνουμε μια ή περισσότερες ομάδες ατόμων και τις παρακολουθούμε στο χρόνο με σκοπό να δούμε την εξέλιξη τους ως προς κάποιες ασθένειες και πιθανούς παράγοντες κινδύνου. Στις περισσότερες φορές, στις μελέτες έχουμε επαναλαμβανόμενες μετρήσεις των οποίων η στατιστική ανάλυση δεν είναι απλή. Στην ουσία πρόκειται για επαναλαμβανόμενες συγχρονικές μελέτες στον ίδιο αριθμό ατόμων σε διαφορετικές χρονικές στιγμές (φανταστείτε σαν ένα αθλητή που τον βγάζουμε πολλές διαδοχικές

φωτογραφίες). Αυτό σημαίνει ότι μπορούμε να πάρουμε τα στοιχεία μιας χρονικής στιγμής (συνήθως της πρώτης) και να τα χειριστούμε σαν απλή διατμηματική-συγχρονική μελέτη.

Ο πιο απλός τρόπος προοπτικής μελέτης αποτελείται από μια συγχρονική μελέτη όπου καταγράφονται τα βασικά στοιχεία του ασθενή ή του ατόμου που εξετάζουμε και μετά απλά καταγράφουμε το χρόνο θανάτου ή τον τελευταίο χρόνο όπου γνωρίζαμε ότι ήταν ζωντανός (μελέτες θνησιμότητας ή χρόνου επιβίωσης).

Οι προοπτικές μελέτες αποτελούν ένα ιδιαίτερα αξιόπιστο τρόπο διερεύνησης αιτιολογικών σχέσεων στην επιδημιολογία. Τα βασικά χαρακτηριστικά τους είναι ότι εκμηδενίζονται τα σφάλματα λόγω καταγραφής ή ανάκλησης μνήμης, λαμβάνεται υπόψη ο χρόνος και χρονική αλληλουχία των σχέσεων και μπορεί να εκτιμηθεί η μεταβλητότητα που οφείλεται στα ίδια τα άτομα.

Μια ειδική περίπτωση προοπτικών μελετών είναι αυτές που βασίζονται σε ιστορικά αρχεία (historical records cohort studies). Σε αυτές τις μελέτες παίρνουμε την ομάδα που μας ενδιαφέρει και εξετάζουμε την πορεία τους μέσα από «αντικειμενικά» ιστορικά αρχεία. Αν και η προσέγγιση είναι στην ουσία αναδρομική, συμπεριλαμβάνεται στις προοπτικές με την έννοια ότι τα στοιχεία τα λαμβάνουμε από «αντικειμενικά» αρχεία που δεν έχουν σφάλμα ανάκλησης μνήμης. Σε αυτές τις μελέτες όμως υπάρχει πιθανότητα σφαλμάτων που να οφείλονται στην αξιοπιστία και την ακρίβεια των αρχείων· για λεπτομέρειες βλέπε Armitage (1994, σελ. 184).

Επιγραμματικά, τα πλεονεκτήματα των προοπτικών μελετών συνοψίζονται στα ακόλουθα :

1. Μπορούμε να μελετήσουμε την επίδραση σπάνιων παραγόντων κινδύνου. Αυτό γίνεται λαμβάνοντας μία ομάδα με άτομα που έχουν εκτεθεί στον (σπάνιο) κίνδυνο (συνήθως απογραφικά) και μία δεύτερη ομάδα (συνήθως από την ίδια γεωγραφική περιοχή) με άτομα που δεν έχουν εκτεθεί στον κίνδυνο (ομάδα ελέγχου).
2. Είναι ιδιαίτερα αξιόπιστες καθώς περιορίζονται τα σφάλματα λόγω ανάκλησης μνήμης, καταγραφής, και προκατάληψης.
3. Μπορούμε να υπολογίσουμε και απόλυτα και σχετικά μέτρα κινδύνου (αυτό σε αντιδιαστολή με τις αναδρομικές που ακολουθούν).
4. Μπορούμε να μελετήσουμε τη σχέση της νόσου που μας ενδιαφέρει και με άλλα νοσήματα.

Υπάρχουν όμως και σημαντικά μειονεκτήματα που δεν αφήνουν τις προοπτικές μελέτες να γίνουν το κύριο εργαλείο έρευνας στην Ιατρική. Αυτά συνοψίζονται στα ακόλουθα :

1. Απαιτούν μεγάλο αριθμό ατόμων ακόμα και για σχετικά συχνές ασθένειες. Είναι εντελώς ανέφικτες και αναποτελεσματικές στην περίπτωση σπάνιων νόσημάτων.

2. Είναι ιδιαίτερα χρονοβόρες (μερικές μεγάλες μελέτες κρατάνε και 50-60 έτη ενώ οι πιο απλές μπορούν να είναι 2-3 ετών).
3. Απαιτούν μεγάλο κόστος και πόρους (ανθρώπινους και σε εξοπλισμό).
4. Η ανάλυση των δεδομένων είναι ιδιαίτερα δύσκολη λόγω του όγκου των δεδομένων αλλά και της φύσης τους (επαναλαμβανόμενες μετρήσεις σε μεγάλο αριθμό ατόμων, πολλές ελλιπείς τιμές). Αυτό συνεπάγεται απασχόληση εξειδικευμένου προσωπικού για αρκετό χρόνο εργασίας (άρα και ανάλογο κόστος).

Όπως είναι κατανοητό, ακόμα και μια μικρής εμβέλειας προοπτική μελέτη είναι δύσκολο να γίνει. Για το λόγο αυτό είναι περιορισμένες και γίνονται από μεγάλα Πανεπιστήμια ή Ερευνητικά κέντρα και συνήθως χρηματοδοτούνται από κυβερνήσεις κρατών ή μη-κερδοσκοπικών οργανισμών.

Άλλο χαρακτηριστικό που μπορεί να αντιμετωπίσουμε σε μια προοπτική μελέτη είναι το γεγονός ότι όλοι οι συμμετέχοντες δεν εισέρχονται στη μελέτη (ούτε εξέρχονται από αυτή) την ίδια χρονική στιγμή. Επιπλέον είναι σύνηθες το φαινόμενο των ελλিপών τιμών γιατί κάποιος ασθενής μπορεί να έχασε μία από τις επισκέψεις στο Ιατρικό κέντρο της έρευνας ή δεν μπορούσε να εντοπιστεί κατά τη διάρκεια μίας επανεξέτασης. Στην περίπτωση που η μεταβλητή απόκριση είναι ο χρόνος επιβίωσης (ή γενικότερα μέχρι την εμφάνιση ενός γεγονότος) τότε πληροφορία μπορούμε να πάρουμε και από την τελευταία χρονική στιγμή όπου γνωρίζουμε ότι κάποιος από τους συμμετέχοντες ήταν εν ζωή (ή γενικότερα τη τελευταία στιγμή που δεν είχε συμβεί το γεγονός που εξετάζουμε). Αυτή η περίπτωση ελλিপών δεδομένων ονομάζεται λογοκρισία (censoring) και τα δεδομένα αυτά λογοκριμένα (censored) δεδομένα.

Τέλος, δύο είναι τα κύρια χαρακτηριστικά μιας προοπτικής μελέτης: οι ομάδες-κοορτές έκθεσης στον κίνδυνο και ο χρόνος παρακολούθησης.

Παράδειγμα 2.1 *Το 1951 έγινε μια μεγάλη προοπτική ταχυδρομική έρευνα στη Βρετανία. Ερωτηματολόγια ταχυδρομήθηκαν σε 59600 γιατρούς εγγεγραμμένους στους Ιατρικούς συλλόγους της Βρετανίας και έλαβα 41000 απαντήσεις. Η μελέτη έγινε σε αυτό τον περιορισμένο πληθυσμό έτσι ώστε να επιτευχθεί υψηλό ποσοστό απόκρισης. Η πεποίθηση ότι το ποσοστό απόκρισης θα ήταν σχετικά υψηλό (πράγμα δύσκολο σε ταχυδρομικές έρευνες) οφειλόταν στο γεγονός ότι ο πληθυσμός των συμμετεχόντων ήταν εκπαιδευμένοι και σχετιζόμενοι με την Ιατρική έρευνα πράγμα που σημαίνει ότι θα συμπλήρωναν το ερωτηματολόγιο ενσυνείδητα και σωστά χωρίς σημαντικά σφάλματα και απώλειες.*

Το ερωτηματολόγιο ήταν σχετικά απλό και αναφερόταν στις συνήθειες καπνίσματος με διαχωρισμένες τρεις ομάδες έκθεσης: καπνιστές, πρώην καπνιστές και μη-καπνιστές. Μετά κατέγραφαν

το χρόνο θανάτου, ο οποίος ήταν σχετικά εύκολος να βρεθεί από τους αντίστοιχους Ιατρικούς συλλόγους. Το 1976 ανακοινώθηκε η πρώτη αναφορά όπου το κύριο αποτέλεσμα ήταν: «οι θάνατοι από καρκίνο του πνεύμονα και άλλες ασθένειες ήταν περισσότεροι για τους καπνιστές». Ως δευτερεύον αποτέλεσμα ανέφεραν και την δοσολογική θετική σχέση καπνίσματος και δείκτη γενικής θνησιμότητας· για λεπτομέρειες βλ. Doll & Peto (1976).

Δύο είναι τα σημαντικά χαρακτηριστικά της μελέτης:

1. Το μέγεθος του δείγματος είναι μεγάλο (41000). Αυτό ήταν θεμιτό αφού η κύρια ασθένεια που εξετάζαν (καρκίνος του πνεύμονα) είναι σπάνια ασθένεια με επιπολασμό μικρότερο του 8/10000.
2. Το δείγμα δεν ήταν τυχαίο. Ήταν απογραφικό του πληθυσμού των Ιατρών της Βρετανίας. Τα αποτελέσματα δεν μπορούν να γενικευτούν σε όλο τον πληθυσμό της Βρετανίας.

2.5.2.2 Διαμόρφωση και Επιλογή των υπό Σύγκριση Ομάδων

Στις περισσότερες προοπτικές μελέτες ξεκινάμε με δύο ή περισσότερες ομάδες-κοορτές με διαφορετική έκθεση σε ένα ή περισσότερους παράγοντες κινδύνου. Συνήθως έχουμε δύο ομάδες έκθεσης και μη έκθεσης σε ένα παράγοντα κινδύνου (εκτεθειμένοι/exposed και μη εκτεθειμένοι/non-exposed στον κίνδυνο).

Οι πιο συνηθισμένες περιπτώσεις είναι οι ακόλουθες

1. Αν ο παράγοντας κινδύνου εμφανίζεται με μεγάλη συχνότητα ή είναι ποσοτική μεταβλητή τότε μπορούμε να χρησιμοποιήσουμε απλή τυχαία δειγματοληψία σε όλο τον πληθυσμό που μας ενδιαφέρει να εξετάσουμε.
2. Σε αρκετές περιπτώσεις επιλέγουμε ομάδες τις οποίες μπορούμε να παρακολουθήσουμε εύκολα. Για παράδειγμα ασθενείς ενός Ιατρείου ή ασφαλισμένοι σε μια εταιρεία ή νεοσύλλεκτοι στρατιώτες. Με αυτό τον τρόπο καταφέρνουμε να μειώσουμε το ποσοστό μη απόκρισης και τις ελλειπείς τιμές. Δυστυχώς όμως τα αποτελέσματα αυτά συνήθως αναφέρονται στον συγκεκριμένο πληθυσμό από τον οποίο πήραμε το δείγμα της μελέτης και δύσκολα γενικεύονται έτσι ώστε να βγάλουμε συμπεράσματα για το γενικό πληθυσμό.
3. Μελέτες με εθελοντές. Είναι πιο σπάνια περίπτωση και είναι επικίνδυνες για μεροληψίες. Σε καμιά περίπτωση δεν είναι αντιπροσωπευτικές του πληθυσμού. Γενικά πρέπει να αποφεύγονται. Οι μόνες μελέτες με εθελοντές που θεωρούνται αξιόπιστες και αποδεκτές, αλλά είναι παρεμβατικής φύσης, είναι οι κλινικές δοκιμές που θα δούμε στο κεφάλαιο 8.

4. Απογραφικές μελέτες ενός πληθυσμού μικρής κλίμακας (κάτοικοι ενός χωριού ή πόλης, φοιτητές ενός πανεπιστημίου κλπ.).

Μελέτες με μέγεθος κοντά στα 10000 άτομα μπορούν να χρησιμοποιηθούν για τη μελέτη σχετικά συχνών ασθενειών όπως η στεφανιαία νόσος.

Επίσης πολύ συχνά χρησιμοποιείται και η εξομοίωση (matching) για να αποφευχθούν οι επιδράσεις συγχυτικών παραγόντων. Έτσι για κάθε άτομο στη μία ομάδα με συγκεκριμένα χαρακτηριστικά αντιστοιχούμε ένα άτομο στην άλλη με ακριβώς τα ίδια χαρακτηριστικά. Με αυτό τον τρόπο, οι δύο ομάδες είναι ακριβώς ίδιες ως προς τους συγχυτικούς παράγοντες αδρανοποιώντας έτσι την επίδρασή τους. Οι πιο σημαντικοί συγχυτικοί παράγοντες είναι η ηλικία και το φύλο. Εάν οι συγχυτικοί παράγοντες δεν αδρανοποιηθούν μέσω του σχεδιασμού της μελέτης, τότε μπορούμε να τους αδρανοποιήσουμε λαμβάνοντας τους ύποψη στην κατάλληλη στατιστική ανάλυση. Τρεις είναι οι σημαντικές προσεγγίσεις: η στρωματοποιημένη ανάλυση, η προτυποποίηση/τυποποίηση δεικτών νοσηρότητας και θνησιμότητας και τα παλινδρομικά μοντέλα.

2.5.2.3 Προβλήματα στην Παρακολούθηση

Ένα βασικό πρόβλημα που αντιμετωπίζουμε στις μελέτες παρακολούθησης είναι η διαγνωστική προκατάληψη. Αν ο Γιατρός γνωρίζει την ομάδα έκθεσης τότε επηρεάζεται και εντοπίζει λανθασμένα αυξημένο αριθμό αρρώστων, σε σχέση με τον πραγματικό, στην ομάδα κινδύνου.

Άλλο σημαντικό πρόβλημα είναι το αυξημένο ποσοστό μη απόκρισης. Εδώ είναι σημαντικό να μην διαφοροποιούνται τα άτομα που τελικά συμμετέχουν στη μελέτη σε σχέση με αυτά που δεν ανταποκρίνονται και δε συμμετέχουν. Αυτό στην πράξη είναι πολύ δύσκολο να ελεγχθεί. Απλοί έλεγχοι μπορούν να γίνουν με διάφορα δημογραφικά στοιχεία τα όποια είναι διαθέσιμα (π.χ. τόπος κατοικίας, επάγγελμα, ηλικία κλπ.).

Τέλος, ένα πολύ συχνό πρόβλημα είναι οι ελλιπείς ή λογοκριμένες τιμές καθώς επίσης η απώλεια ατόμων κατά τη διάρκεια της μελέτης. Το ποσοστό αυτών των περιπτώσεων μπορεί να σχετίζεται και με την ίδια τη φύση της ασθένειας.

2.5.3 Αναδρομικές Μελέτες ή Μελέτες Μαρτύρων-Ασθενών

2.5.3.1 Γενικές Έννοιες

Οι αναδρομικές μελέτες συλλέγουν, από το παρόν, δεδομένα που αφορούν το παρελθόν, δηλαδή κοιτάζουν πίσω χρονικά. Αυτό σημαίνει ότι οι συμμετέχοντες ερωτώνται για συνήθειες που είχαν στο παρελθόν. Έτσι αυτές οι μελέτες βασίζονται σε αναμνηστική ανάκληση πληροφοριών (recall information) και για το λόγο αυτό εμπεριέχουν σφάλμα μέτρησης.

Οι πιο συνηθισμένες αναδρομικές μελέτες είναι αυτές των μαρτύρων-ασθενών (case-control studies). Συνήθως αποτελούνται από δύο ομάδες: των περιπτώσεων (cases) ή ασθενών (δηλαδή των ατόμων που έχουν τη νόσο) και των μαρτύρων (controls, ατόμων δηλαδή που δεν έχουν τη νόσο ή άλλη σχετική). Οι περιπτώσεις αποτελούν την ομάδα των ασθενών (disease group) ενώ οι μάρτυρες αποτελούν την ομάδα ελέγχου (control group). Τα άτομα που έχουν τη νόσο δεν επιλέγονται τυχαία αλλά συνήθως απογραφικά από μια (ή περισσότερες) νοσοκομειακές μονάδες που ειδικεύονται στη νόσο που θέλουμε να εξετάσουμε. Οι μάρτυρες μπορούν να επιλεγούν από το γενικό πληθυσμό. Συνήθως όμως επιλέγονται από άλλα τμήματα των ίδιων νοσοκομειακών μονάδων από τα οποία έχουμε επιλέξει τους ασθενείς. Φυσικά δεν πρέπει να έχουν κάποια ασθένεια σχετική με αυτή που εξετάζουμε. Ο κύριος λόγος για την επιλογή των μαρτύρων από το ίδιο νοσοκομείο είναι ότι έτσι διατηρούμε διάφορους κοινωνικο-οικονομικούς συγχυτικούς παράγοντες σταθερούς. Στόχος αυτών των μελετών είναι να εξετάσουμε αν υπάρχει διαφοροποίηση της έκθεσης στον κίνδυνο για τις δύο ομάδες (ασθενών και μαρτύρων). Αυτό γιατί αν ένας παράγοντας συνδέεται με μία νόσο τότε περιμένουμε να εμφανίζεται και πιο συχνά στην ομάδα των ασθενών.

Το κύριο της πλεονέκτημα είναι το χαμηλό κόστος και η σχετική ευκολία διενέργειας της. Ένα άλλο βασικό πλεονέκτημα είναι ότι μπορεί να χρησιμοποιηθεί για σπάνιες ασθένειες σε αντίθεση με τις προοπτικές μελέτες.

2.5.3.2 Στάδια Σχεδιασμού

Τα κύρια στάδια του σχεδιασμού μιας αναδρομικής μελέτης είναι τα ακόλουθα

1. Η επιλογή των ασθενών.
2. Η επιλογή των μαρτύρων.
3. Ο τρόπος ορισμού και μέτρησης του παράγοντα κινδύνου.
4. Η ανάλυση της σχέσης νόσου-παράγοντα κινδύνου.

α) Επιλογή των ασθενών: Πρώτα από όλα θα πρέπει να γίνει αναλυτικός ορισμός του «ασθενή» έτσι ώστε να είναι ξεκάθαρο ποια άτομα θα μπορούν να εισαχθούν στην ομάδα ασθενών. Ένας βασικός κανόνας είναι να συμπεριλαμβάνουμε άτομα στα οποία η νόσος έχει πρόσφατα εντοπιστεί. Άτομα που έχουν άλλες σημαντικές και χρόνιες ασθένειες δε θα πρέπει να συμπεριλαμβάνονται στην ομάδα ασθενών. Όπως είπαμε και παραπάνω, τους ασθενείς τους λαμβάνουμε απογραφικά για ένα χρονικό διάστημα από εξειδικευμένες Ιατρικές μονάδες. Αυτό κυρίως συμβαίνει διότι ο πληθυσμός των ασθενών είναι περιορισμένος και μικρός.

β) Επιλογή των μαρτύρων: Η επιλογή των μαρτύρων είναι σχετικά πιο δύσκολη. Οι μάρτυρες θα πρέπει να αντιπροσωπεύουν το γενικό πληθυσμό. Σε αυτή την περίπτωση θα μπορούσαμε να κάνουμε απλή τυχαία ή στρωματοποιημένη δειγματοληψία. Όμως μερικά πρακτικά προβλήματα (υψηλό ποσοστό μη απόκρισης και υψηλό κόστος) οδηγούν στην επιλογή των μαρτύρων από τις ίδιες νοσοκομειακές μονάδες από όπου επιλέξαμε και τους ασθενείς. Προσοχή όμως δε θα πρέπει να έχουν καμία ασθένεια σχετική με τη νόσο που εξετάζουμε ή άλλη σημαντική ασθένεια. Επίσης άλλος σημαντικός παράγοντας που συνηγορεί στην χρήση μαρτύρων από τις ίδιες μονάδες είναι το γεγονός ότι εξασφαλίζει μεγαλύτερη συγκρισιμότητα των μαρτύρων με τους ασθενείς όσον αφορά κοινωνικο-οικονομικούς και γεωγραφικούς παράγοντες.

Για να αδρανοποιήσουμε συγχυτικούς παράγοντες όπως η ηλικία και το φύλο, πολλές φορές οι μάρτυρες επιλέγονται έτσι ώστε να έχουν ίδια χαρακτηριστικά με τους ασθενείς. Όταν αυτό γίνεται για το κάθε άτομο ξεχωριστά τότε έχουμε ατομική εξομοίωση (*individual matching*) ενώ όταν η εξομοίωση γίνεται έτσι ώστε οι δύο ομάδες να έχουν ίδιες περιθώριες κατανομές ως προς τους συγχυτικούς παράγοντες τότε έχουμε την εξομοίωση συχνοτήτων (*frequency matching*).

Τέλος ένα βασικό στοιχείο είναι ότι το άτομο που κάνει την επιλογή των μαρτύρων δε θα πρέπει να γνωρίζει τους παράγοντες κινδύνου έτσι ώστε να αποφύγουμε τη μεροληψία επιλογής.

γ) Μέτρηση της έκθεσης στον κίνδυνο: Ο τρόπος που μετριέται η έκθεση στον κίνδυνο βασίζεται αποκλειστικά σε αναμνηστική ανάκληση της πληροφορίας. Για το λόγο αυτό εμπειριέχει και σημαντικό σφάλμα μέτρησης. Στις μελέτες αυτές συνήθως η έκθεση μετριέται με κατηγορικές μεταβλητές λίγων επιπέδων (π.χ. κάπνιζες πριν ένα έτος; ΝΑΙ/ΟΧΙ). Ακόμα και αν η έκθεση στον κίνδυνο μετριέται ποσοτικά, οι αποκρίσεις σε αυτές τις μελέτες πρέπει να δίνονται σε κατηγορίες/ομάδες. Αυτό έτσι ώστε να ελαχιστοποιούνται σφάλματα μέτρησης λόγω της ανάκλησης πληροφορίας από τη μνήμη. Σε αυτές τις μελέτες είναι πολύ πιο δύσκολο να εντοπιστούν δοσολογικές σχέσεις νόσου-παράγοντα κινδύνου. Σε πολλές περιπτώσεις, η έκθεση στον κίνδυνο μπορεί να ληφθεί από αξιόπιστους καταλόγους νοσοκομείων ή να οριστεί αντικειμενικά έτσι ώστε να ελαχιστοποιηθεί το σφάλμα ανάκλησης μνήμης (π.χ. απόσταση κατοικίας τα 5 τελευταία έτη από ένα εργοστάσιο χημικής επεξεργασίας).

Τέλος η συνέντευξη θα πρέπει να γίνεται από το ίδιο άτομο το οποίο να έχει καταρτιστεί κατάλληλα. Επιπλέον πρέπει να γίνεται με τον ίδιο ακριβώς τρόπο και στις δύο ομάδες. Για να αποφευχθούν απρόβλεπτα προβλήματα μεροληψίας καλό θα είναι να μη γνωρίζει την ομάδα στην οποία ανήκει το κάθε άτομο.

δ) Ανάλυση της σχέσης νόσου-παράγοντα κινδύνου: Εδώ ορίζουμε το τρόπο μέτρησης του κινδύνου. Παραδείγματα κυρίως σε κατηγορικά δεδομένα θα δούμε στην επόμενη ενότητα.

2.5.3.3 Πλεονεκτήματα και Μειονεκτήματα

Τα κύρια πλεονεκτήματα των αναδρομικών μελετών είναι τα ακόλουθα:

1. Απαιτούν μικρό δείγμα.
2. Είναι περιορισμένες χρονικά (συνήθως κάθε άτομο εξετάζεται μία φορά μόνο).
3. Είναι μελέτες σχετικά εύκολες και με χαμηλό κόστος.
4. Μελετούν ικανοποιητικά τη σχέση νόσου και παράγοντα κινδύνου.
5. Μπορούν να χρησιμοποιηθούν για τη μελέτη σπάνιων ασθενειών.

Τα μειονεκτήματα είναι αρκετά και μπορούν να συνοψιστούν στα ακόλουθα:

1. Είναι επιρρεπείς σε μεροληψίες και σφάλματα μέτρησης (κυρίως λόγω της ανάκλησης μνήμης).
2. Πρέπει να γίνεται προσεκτική επιλογή της ομάδας ελέγχου.
3. Μπορούμε να υπολογίσουμε μόνο σχετικά μέτρα κινδύνου και όχι απόλυτα (βλέπε επόμενη ενότητα).
4. Πλασματικές σχέσεις εμφανίζονται πιο συχνά από ότι στις προοπτικές.
5. Αν δεν προσέξουμε μπορεί να έχουμε μεροληπτική επιλογή των συμμετεχόντων (selection bias).
6. Εξετάζουν τις σχέσεις ανάποδα όσον αφορά τη χρονική τους αλληλουχία και για το λόγο αυτό δεν έχουν ισχυρό αποδεικτικό κύρος.

Παρόλα αυτά οι αναδρομικές μελέτες είναι ένα πολύ σημαντικό εργαλείο το οποίο έχει επικρατήσει στην Ιατρική έρευνα. Ένας απλός κανόνας είναι ότι πρέπει να εμπιστευόμαστε κυρίως αποτελέσματα μεγάλων προοπτικών μελετών αλλά «καμία προοπτική έρευνα δεν πρέπει να ξεκινάει χωρίς να έχει προηγηθεί μια αναδρομική». Για το λόγο αυτό μπορούν να λειτουργήσουν ως διερευνητικές και επικουρικές των μεγάλων προοπτικών μελετών.

Κεφάλαιο 3

Δείκτες Νοσηρότητας, Μέτρα Κινδύνου και Διαγνωστικού Ελέγχου

3.1 Μέτρα Νοσηρότητας - Εμφάνιση μιας Νόσου

3.1.1 Επιπολασμός

Ως επιπολασμό (prevalence) μιας ασθένειας ονομάζουμε το ποσοστό (ή αναλογία) του πληθυσμού που έχει την νόσο τη χρονική στιγμή t (στιγμιαίος επιπολασμός). Ως επιπολασμό περιόδου $[t_0, t_1)$ μιας ασθένειας ονομάζουμε το ποσοστό (ή αναλογία) του πληθυσμού που έχει την νόσο σε μία οποιαδήποτε χρονική στιγμή t μέσα στο διάστημα $[t_0, t_1)$ (δηλαδή $t_0 \leq t < t_1$). Πιο απλά ο (στιγμιαίος) επιπολασμός είναι η πιθανότητα ένα άτομο του πληθυσμού να έχει την υπό εξέταση νόσο d τη χρονική στιγμή t . Δίνεται από τον τύπο

$$\text{Prevalence}_t^d = \frac{N_t^d}{N_t},$$

όπου N_t^d είναι ο αριθμός των ατόμων του πληθυσμού με την ασθένεια d τη χρονική στιγμή t και $N(x)$ είναι το μέγεθος του συνολικού πληθυσμού τη χρονική στιγμή t . Ο δείκτης του επιπολασμού θεωρείται μέτρο συγχρονικών-διατμηματικών μελετών (cross-sectional measure) διότι δε λαμβάνει υπόψη του το χρόνο και εκτιμάται μόνο από τέτοιου είδους μελέτες όπου η δειγματοληψία είναι τυχαία και αντιπροσωπευτική του γενικού πληθυσμού. Η εκτίμηση του επιπολασμού από μια συγχρονική-διατμηματική μελέτη μεγέθους n δίνεται απλά από τον τύπο :

$$\widehat{\text{Prevalence}}_t^d = \frac{1}{n} \sum_{i=1}^n Y_i^d = \frac{n^d}{n},$$

όπου Y_i^d είναι μια δίτιμη-δείκτρια μεταβλητή που παίρνει τιμές ένα (1) και μηδέν (0) αν το i άτομο έχει ή όχι τη νόσο d αντίστοιχα και n^d είναι ο συνολικός αριθμός των ατόμων στο δείγμα

που έχουν τη νόσο d .

Να σημειώσουμε ότι οι παραπάνω τύποι αναφέρονται στο στιγμιαίο επιπολασμό ο οποίος χρησιμοποιείται πιο συχνά. Στην περίπτωση υπολογισμού του επιπολασμού περιόδου τότε τα μεγέθη N_t και n συμβολίζουν το μέσο αριθμό ατόμων του πληθυσμού και δείγματος κατά το χρονικό διάστημα που εξετάζουμε. Για λεπτομέρειες παραπέμπουμε στα βιβλία των Daly *et. al* (1991, σελ. 281-282) και Pereira-Maxwell (1998, σελ. 62).

3.1.2 Αθροιστική Επίπτωση

Σε αντίθεση με τον επιπολασμό που είναι ένας στιγμιαίος δείκτης που εκτιμάει την έκταση μιας νόσου στον πληθυσμό εδώ θα ορίσουμε κάποιους δείκτες που λαμβάνουν υπόψη τους και το χρόνο. Πριν όμως προχωρήσουμε στον ορισμό της αθροιστικής επίπτωσης πρέπει να ορίσουμε την έννοια του επεισοδίου ή περιστατικού μίας νόσου.

Μια περίπτωση (case) ονομάζεται «επεισόδιο» ή «περιστατικό» (incident) τη στιγμή ακριβώς που εμφανίζει για πρώτη φορά τα συμπτώματα της νόσου.

«Αθροιστική επίπτωση» (cumulative incidence) μίας νόσου για το χρονικό διάστημα $[t_0, t_1)$ είναι η πιθανότητα ένα άτομο που δεν έχει εκδηλώσει προηγουμένως την ίδια νόσο να εμφανίσει για πρώτη φορά τα συμπτώματα της νόσου μέσα σε αυτό το διάστημα (Rosner, 1994, σελ. 61). Συνεπώς δίνεται από την πιθανότητα

$$I_{t_0, t_1}^d = P(t_0 < D^d < t_1 | D^d > t_0) = \frac{NI_{t_0, t_1}^d}{\overline{NI}_{t_0}^d}$$

όπου D^d είναι ο χρόνος εμφάνισης της νόσου d , NI_{t_0, t_1}^d είναι ο αριθμός των νέων επεισοδίων της νόσου d στο χρονικό διάστημα $[t_0, t_1)$ και $\overline{NI}(t_0)$ είναι αριθμός των ατόμων του πληθυσμού χωρίς να έχει εκδηλώσει τη νόσο d πριν τον χρόνο t_0 . Η αθροιστική επίπτωση εκτιμάται από προοπτικές μελέτες (σε αντίθεση με τον επιπολασμό) από το τύπο:

$$\hat{I}_{t_0, t_1}^d = \frac{\sum_{i=1}^n Y_i^{Id(t_0, t_1)}}{\sum_{i=1}^n Y_i^{\overline{Id}(t_0)}} = \frac{n\hat{i}_{t_0, t_1}^d}{\overline{n\hat{i}}_{t_0}^d}$$

όπου $n\hat{i}^d(t_0, t_1)$ και $\overline{n\hat{i}}^d(t_0)$ είναι οι αντίστοιχες δειγματικές ποσότητες των $NI^d(t_0, t_1)$ και $\overline{NI}(t_0)$. Επιπλέον $Y_i^{Id(t_0, t_1)}$ και $Y_i^{\overline{Id}(t_0)}$ είναι δίτιμες-δείκτριες (0 – 1) συναρτήσεις που παίρνουν τη τιμή ένα αν το i άτομο εμφανίσει επεισόδιο μίας νόσου στο χρονικό διάστημα $[t_0, t_1)$ και αν το i άτομο δεν έχει προϋστορία της νόσου μέχρι το χρόνο t_0 αντίστοιχα.

Συνάρτηση ή πιθανότητα αποφυγής της νόσου (ή επιβίωσης) μέχρι τον χρόνο t είναι η πιθανότητα ένα άτομο να μην εμφανίσει την νόσο πριν τη χρονική στιγμή t (ή να επιβιώσει τουλάχιστον μέχρι τη χρονική στιγμή t) και δίνεται από τον τύπο $S(t) = P(D^d > t)$.

Χρησιμοποιώντας τη συνάρτηση επιβίωσης μπορούμε να γράψουμε:

$$I(t_0, t_1) = \frac{S(t_0) - S(t_1)}{S(t_0)} = 1 - \frac{S(t_1)}{S(t_0)}.$$

3.1.3 Ρυθμός Επίπτωσης

Η αθροιστική επίπτωση λαμβάνει υπόψη του το χρόνο ως προς τα επεισόδια της νόσου που εμφανίζονται μέσα σε ένα χρονικό διάστημα. Δε λαμβάνει υπόψη του το χρόνο διάρκειας μιας νόσου ο οποίος είναι σημαντικός παράγοντας για τη κατανόηση της επίδρασης της νόσου στον πληθυσμό. Για παράδειγμα μια νόσος με μικρή πιθανότητα εμφάνισης αλλά μεγάλη διάρκεια μπορεί να εμφανίζει υψηλό επιπολασμό δηλαδή να υπάρχουν πολλά άτομα στον πληθυσμό με τη νόσο και για το λόγο αυτό να είναι σημαντική η επίδρασή της στο σύνολο του πληθυσμού.

Έτσι ορίζουμε το «ρυθμό επίπτωσης» (incidence rate) ως το λόγο του αριθμού των νέων επεισοδίων μιας νόσου σε ένα χρονικό διάστημα $[t_0, t_1]$ προς «ανθρωποχρόνο κινδύνου» για την ίδια περίοδο (Pereira-Maxwell, 1998, σελ. 29). Ο «ανθρωποχρόνος κινδύνου» δεν είναι τίποτα άλλο από το χρόνο παρακολούθησης αθροιστικά για όλα τα άτομα του πληθυσμού. Έτσι, εδώ ο παρανομαστής είναι χρόνος σε αντίθεση με τηναθροιστική επίπτωση όπου παρανομαστής είναι ο αριθμός των ατόμων του πληθυσμού. Όσον αφορά τον πληθυσμό, ο ρυθμός επίπτωσης δίνεται από τον τύπο:

$$IR_{t_0, t_1}^d = \frac{NI_{t_0, t_1}^d}{T_{(t_0, t_1)}} = \frac{NI_{t_0, t_1}^d}{\sum_{i=1}^{N_{t_0}} T_{i, (t_0, t_1)}}$$

όπου N_{t_0} είναι όλα τα άτομα στον πληθυσμό το χρόνο t_0 , $T_{(t_0, t_1)}$ είναι ο συνολικός ανθρωποχρόνος παρακολούθησης και $T_{i, (t_0, t_1)}$ είναι ο χρόνος παρακολούθησης μέσα στο χρονικό διάστημα $[t_0, t_1]$ για το άτομο i . Αν δεν υπάρχουν απώλειες στο χρόνο παρακολούθησης και ως συνέπεια όλα τα άτομα παρακολουθούνται για συνολική διάρκεια ίση με $(t_1 - t_0)$ τότε ο ρυθμός επίπτωσης δίνεται ως

$$IR_{t_0, t_1}^d = \frac{NI_{t_0, t_1}^d}{N_{t_0}(t_1 - t_0)}.$$

3.1.4 Διάρκεια Ασθένειας

Ως «διάρκεια ασθένειας» (duration) ορίζουμε το χρονικό διάστημα από τη πρώτη εμφάνιση της νόσου ως την τελική θεραπεία και τη συμβολίζουμε ως D (ή D^d). Όταν αναφερόμαστε σε ένα πληθυσμό ή ένα δείγμα τότε μας ενδιαφέρει η μέση διάρκεια της νόσου για τα άτομα που έχουν εμφανίσει τη νόσο και τη συμβολίζουμε ως \bar{D} (ή \bar{D}^d). Όταν μια ασθένεια είναι θανατηφόρα τότε ως διάρκεια ζωής ονομάζεται ο χρόνος μέχρι την κατάληξη του ασθενή. Όπως είπαμε η διάρκεια είναι σημαντικό μέτρο της επίδρασης μίας ασθένειας σε ένα πληθυσμό καθώς ασθένειες

με μεγάλη μέση διάρκεια οδηγούν σε υψηλό επιπολασμό (δηλαδή πολλά άτομα στον πληθυσμό έχουν τη νόσο) ακόμα και αν ο ρυθμός επίπτωσης είναι μικρός.

3.2 Μέτρα Κινδύνου για Δίτιμα - Κατηγορικά Δεδομένα.

Συνήθως στις Ιατρικές μελέτες, η κύρια μεταβλητή απόκρισης είναι δίτιμη (η ασθένεια είναι παρούσα ή όχι). Επιπλέον οι πιο συνηθισμένοι παράγοντες κινδύνου είναι επίσης κατηγορικοί όπως για παράδειγμα το φύλο. Έτσι λοιπόν σε αυτή την ενότητα θα ορίσουμε δείκτες μέτρησης του κινδύνου εμφάνισης μίας νόσου όταν έχουμε δίτιμες κατηγορικές μεταβλητές. Στην πιο απλή περίπτωση μας ενδιαφέρει να συγκρίνουμε τον κίνδυνο εμφάνισης μιας νόσου σε ένα άτομο που έχει εκτεθεί στον παράγοντα κινδύνου με τον αντίστοιχο κίνδυνο ενός ατόμου που δεν έχει εκτεθεί στον ίδιο παράγοντα κινδύνου.

Ενώ όταν συγκρίνουμε δύο ή περισσότερες ομάδες ως προς κάποιες ποσοτικές μεταβλητές μας ενδιαφέρει κυρίως η σύγκριση των μέτρων κεντρικής τους θέσης (κυρίως μέση τιμή και/ή διάμεσος), στην αντίστοιχη σύγκριση κάποιων ομάδων ως προς κάποιες κατηγορικές μεταβλητές μας ενδιαφέρει η σύγκριση των δεσμευμένων κατανομών ή πιθανοτήτων. Αυτό πρακτικά μεταφράζεται στη σύγκριση των ποσοστών (ή πιθανοτήτων ή αναλογιών) για κάθε διαφορετική ομάδα έκθεσης σε ένα παράγοντα κινδύνου.

Στην ενότητα αυτή θα ασχοληθούμε με την πιο απλή περίπτωση της σύγκρισης δίτιμων (Bernoulli) παραγόντων κινδύνου X σε δύο ομάδες μιας δεύτερης κατηγορικής μεταβλητής Y που συνήθως είναι η εμφάνιση της νόσου. Δηλαδή θα επικεντρωθούμε στη σύγκριση κατηγορικών μεταβλητών που περιγράφονται από 2×2 πίνακες συνάφειας.

3.2.1 Εισαγωγικές Έννοιες για 2×2 Πίνακες Συνάφειας.

3.2.1.1 Πίνακες Συνάφειας

Αν θέλουμε να περιγράψουμε 2 κατηγορικές μεταβλητές μαζί τότε πρέπει να κάνουμε ομαδοποιημένα ραβδοδιαγράμματα (barchart) και πίνακες συνάφειας διπλής εισόδου (two - way contingency tables). Οι πίνακες συνάφειας διπλής εισόδου δεν είναι τίποτα άλλο από κατανομές συχνοτήτων για δύο κατηγορικές μεταβλητές δηλαδή για όλους τους συνδυασμούς των κατηγοριών των 2 μεταβλητών που περιγράφουμε. Εκτιμούν την από κοινού συνάρτηση πιθανότητας των δύο μεταβλητών που εξετάζουμε. Έστω δύο μεταβλητές X και Y με επίπεδα I και J τότε έχουμε $I \times J$ συνδυασμούς. Η απεικόνιση του γίνεται με ένα ορθογώνιο πίνακα με I γραμμές και J στήλες. Κάθε συνδυασμός εμφανίζεται στον πίνακα διπλής εισόδου ως ένα κουτί το οποίο ονομάζεται κελί (cell). Ο αριθμός των εμφανίσεων κάθε συνδυασμού $(X, Y) = (i, j)$ στο δείγμα

ονομάζεται συχνότητα (count) και συμβολίζεται με n_{ij} . Το συνολικό μέγεθος του δείγματος συμβολίζεται με n και είναι ίσο με $n = \sum_{i=1}^I \sum_{j=1}^J n_{ij}$. Η αντίστοιχη πιθανότητα $P(X = i, Y = j)$ συμβολίζεται ως π_{ij} και εκτιμάται από το $p_{ij} = n_{ij}/n$.

Γενικά, όταν λέμε ότι έχουμε έναν $I \times J$ πίνακα στην ουσία εννοούμε ένα πίνακα συνάφειας διπλής εισόδου με I γραμμές και J στήλες. Έτσι ένας 2×2 πίνακας διπλής εισόδου θα έχει την ακόλουθη μορφή:

X: Παρ. Κινδύνου	Y: Νόσος		Περ.Κατ.
	1 (Ασθενής)	2 (Υγιής)	X
1 (Παρόν)	n_{11}	n_{12}	$n_{1\bullet}$
2 (Απόν)	n_{21}	n_{22}	$n_{2\bullet}$
Περ.Κατ. Y	$n_{\bullet 1}$	$n_{\bullet 2}$	$n_{\bullet\bullet} = n$

Οι μονομεταβλητές κατανομές των X και Y δίνονται στην τελευταία στήλη και γραμμή του παραπάνω πίνακα και ονομάζονται περιθώριες κατανομές (marginal distributions). Η περιθώρια συχνότητα μιας κατηγορίας i της μεταβλητής X συμβολίζεται ως $n_{i\bullet}$ (ή εναλλακτικά n_{i+}) ενώ η περιθώρια συχνότητα μιας κατηγορίας j της μεταβλητής Y συμβολίζεται ως $n_{\bullet j}$ (ή εναλλακτικά n_{+j}) και δίνονται από τους τύπους $n_{i\bullet} = \sum_{j=1}^J n_{ij}$ και $n_{\bullet j} = \sum_{i=1}^I n_{ij}$. Γενικά η τελεία « \bullet » ή εναλλακτικά το σύμβολο της πρόσθεσης « $+$ » στους δείκτες υποδηλώνει άθροισμα. Έτσι με $n_{\bullet\bullet}$ ή n_{++} πολύ συχνά συμβολίζεται το συνολικό άθροισμα του δείγματος. Με παρόμοιο τρόπο ορίζονται και οι περιθώριες πιθανότητες $\pi_{i\bullet}$ και $\pi_{\bullet j}$ των X και Y και οι αντίστοιχες δειγματικές τους εκτιμήσεις $p_{i\bullet}$ και $p_{\bullet j}$. Οι παραπάνω πίνακες διπλής εισόδου ονομάζονται πίνακες συνάφειας (contingency tables) όταν τα κελιά περιέχουν συχνότητες (counts). Εναλλακτική ονομασία είναι «πίνακες σταυρωτής καταχώρησης» (cross-classified tables).

3.2.1.2 Δεσμευμένες Πιθανότητες

Σε πολλούς πίνακες η μία μεταβλητής (π.χ. Y) είναι το λογικό αποτέλεσμα - απόκριση (response) και η άλλη (π.χ. X) είναι η ανεξάρτητη ή επεξηγηματική μεταβλητή (explanatory variable). Σε αυτή την περίπτωση η ανεξάρτητη μεταβλητή X μπορεί να είναι σταθερή ή προκαθορισμένη από το σχεδιασμό του πειράματος και όχι τυχαία μεταβλητή. Τότε η πιθανότητα $\pi_{ij} = P(X = i, Y = j)$ δεν έχει νόημα (αφού η X δεν είναι τυχαία μεταβλητή). Όμως μας ενδιαφέρει η κατανομή της μεταβλητής Y για κάθε δεδομένη και διαφορετική τιμή (επίπεδο) της μεταβλητής X . Συνεπώς μας ενδιαφέρει να εξετάσουμε πως η μεταβλητή Y αλλάζει για διαφορετικές τιμές του X . Έτσι, λοιπόν, έχει νόημα η ανάλυση των υπό-συνθήκη πιθανοτήτων

$$\pi_{j|i} = P(Y = j | X = i)$$

με $\sum_{j=1}^J \pi_{j|i} = 1$. Οι πιθανότητες $(\pi_{1|i}, \dots, \pi_{J|i})$ δίνουν τη δεσμευμένη (ή υπό συνθήκη) κατανομή (conditional distribution) της Y για το i επίπεδο της X .

Στους πίνακες 2×2 ο έλεγχος ανεξαρτησίας μπορεί να γίνει με τον έλεγχο της ισότητας δύο και μόνο δεσμευμένων πιθανοτήτων (βλέπε εξίσωση 3.1). Έτσι, λοιπόν, μπορούμε να φτιάξουμε διάφορα μέτρα τα όποια δεν είναι τίποτα άλλο από συναρτήσεις των πιθανοτήτων $P(\text{Ασθενής}|\text{Κίνδυνος Παρόν}) = \pi_{j=1|i=1}$ και $P(\text{Ασθενής}|\text{Κίνδυνος Απών}) = \pi_{j=1|i=2}$ και στην ουσία θα μετρούν την σχέση και απόκλιση των δύο αυτών πιθανοτήτων.

3.2.1.3 Ανεξαρτησία Δύο Κατηγορικών Μεταβλητών

Η ανεξαρτησία μεταξύ δύο μεταβλητών ορίζεται στατιστικά ως

$$\pi_{ij} = P(X = i, Y = j) = P(X = i)P(Y = j) = \pi_{i \bullet} \pi_{\bullet j}.$$

Η ανεξαρτησία των δύο μεταβλητών συνεπάγεται ότι και οι δεσμευμένες κατανομές $P(Y|X = i)$ είναι ίσες για όλα τα i εφόσον

$$\begin{aligned} \pi_{j|i} &= P(Y = j|X = i) = \frac{P(Y = j, X = i)}{P(X = i)} \\ &= \frac{P(Y = j)P(X = i)}{P(X = i)} = P(Y = j) = \pi_{\bullet j}. \end{aligned}$$

Αυτό πρακτικά σημαίνει ότι η εμφάνιση της νόσου (Y) παραμένει σταθερή ανεξάρτητα της έκθεσης ή όχι στον παράγοντα κινδύνου X . Με τον ίδιο τρόπο μπορούμε να βρούμε τα ίδια αποτελέσματα για τις δεσμευμένες κατανομές $\pi_{i|j} = P(X = i|Y = j)$.

Συνεπώς ο έλεγχος της ανεξαρτησίας είναι ισοδύναμος με τον έλεγχο

$$H_0 : \pi_{j|i} = \pi_{\bullet j} \quad \forall i = 1, \dots, I \text{ και } j = 1, \dots, J.$$

Επειδή υπάρχει ο περιορισμός $\sum_{j=1}^J \pi_{j|i} = 1$, στον παραπάνω έλεγχο ο δείκτης j μπορεί να πάρει τιμές μέχρι το $J - 1$ εφόσον ισότητα των πιθανοτήτων των πρώτων $J - 1$ επιπέδων θα συνεπάγεται αυτόματα και ισότητα των πιθανοτήτων στο J επίπεδο. Για 2×2 πίνακες ο έλεγχος μπορεί να περιοριστεί στην απλή σχέση

$$H_0 : \pi_{j=1|i=1} = \pi_{j=1|i=2} \quad \text{έναντι της εναλλακτικής } H_1 : \pi_{j=1|i=1} \neq \pi_{j=1|i=2}, \quad (3.1)$$

που για τη περίπτωση νόσου και παράγοντα κινδύνου μπορεί να γραφτεί ως

$$H_0 : P(A|E) = P(A|\bar{E}) \quad \text{έναντι της εναλλακτικής } H_1 : P(A|E) \neq P(A|\bar{E}),$$

όπου A , E και \bar{E} είναι τα ενδεχόμενα να είναι κάποιος ασθενής (δηλαδή να εμφανίσει τη νόσο), να έχει εκτεθεί και να μην έχει εκτεθεί στον παράγοντα κινδύνου αντιστοίχως.

Στην ουσία όλα τα μέτρα κινδύνου που θα περιγράψουμε πιο κάτω συγκρίνουν τον κίνδυνο, ο οποίος μετριέται με την πιθανότητα εμφάνισης της νόσου, σε κάθε ομάδα έκθεσης σε ένα παράγοντα κινδύνου. Τα μέτρα κινδύνου στην ουσία μετράνε τις διαφορές στον κίνδυνο (πιθανότητα εμφάνισης της νόσου) και μας δίνουν ένα συνολικό δείκτη εξάρτησης μεταξύ του παράγοντα κινδύνου που εξετάζουμε και της νόσου.

3.2.2 Οφειλόμενος ή Αποδιδόμενος Κίνδυνος

Ως αποδιδόμενο κίνδυνο (Δαφνή, 1997 - attributable risk) ή διαφορά κινδύνου (risk difference, Rosner, 1995, σελ. 362-363) ορίζουμε τη διαφορά μεταξύ των ρυθμών επίπτωσης (ή των δεικτών θνησιμότητας) των ομάδων με άτομα εκτεθειμένα και μη εκτεθειμένα σε ένα παράγοντα κινδύνου. Εναλλακτικά ονομάζεται και οφειλόμενος ή αποδοτέος κίνδυνος (Τριχόπουλος, 1982, σελ. 189). Όταν αναφερόμαστε σε ένα πίνακα συνάφειας 2×2 τότε ο αποδιδόμενος κίνδυνος ορίζεται ως η διαφορά των δεσμευμένων πιθανοτήτων $\pi_{j=1|i=1}$ και $\pi_{j=1|i=2}$. Έτσι μπορούμε να ορίσουμε

$$AR = \pi_E - \pi_{\bar{E}} = P(A|E) - P(A|\bar{E}) = \pi_{j=1|i=1} - \pi_{j=1|i=2} = \pi_{j=2|i=2} - \pi_{j=2|i=1}$$

όπου $\pi_E = P(A|E)$ και $\pi_{\bar{E}} = P(A|\bar{E})$ ορίζουμε για λόγους συντομίας τις πιθανότητες εμφάνισης της νόσου ανάλογα αν κάποιος έχει εκτεθεί στον κίνδυνο (E) ή όχι (\bar{E}). Η ποσότητα αυτή μπορεί να εκτιμηθεί ως

$$\widehat{AR} = p_E - p_{\bar{E}} = p_{j=1|i=1} - p_{j=1|i=2} = \frac{n_{11}}{n_{1\bullet}} - \frac{n_{21}}{n_{2\bullet}} = \frac{n_{11}}{n_{11} + n_{12}} - \frac{n_{21}}{n_{21} + n_{22}},$$

όπου p_E και $p_{\bar{E}}$ αντιστοιχούν στις δειγματικές αναλογίες εμφάνισης της νόσου ανάλογα αν κάποιος έχει εκτεθεί στον κίνδυνο ή όχι. Ο έλεγχος ανεξαρτησίας (3.1) μπορεί να γραφτεί απλά ως $H_0 : AR = 0$ έναντι της εναλλακτικής $H_1 : AR \neq 0$.

Το πρόβλημα με τον παραπάνω δείκτη είναι ότι μετράει απόλυτες διαφορές. Έτσι οι πιθανότητες 0.10 και 0.20 θα μας δώσουν ίδιο AR με τις πιθανότητες 0.6 και 0.7 αγνοώντας το γεγονός ότι στην πρώτη περίπτωση η πιθανότητα της μίας ομάδας είναι διπλάσια από την άλλη. Αυτό μπορεί να διορθωθεί με τον ποσοστιαίο αποδιδόμενο κίνδυνο (Τριχόπουλος, 1982, σελ. 190) που δίνεται από τον τύπο

$$p_{AR} = \frac{\pi_E - \pi_{\bar{E}}}{\pi_E} = \frac{P(A|E) - P(A|\bar{E})}{P(A|E)} = \frac{\pi_{j=1|i=1} - \pi_{j=1|i=2}}{\pi_{j=1|i=1}} = \frac{AR}{\pi_E}.$$

3.2.3 Σχετικός Κίνδυνος (Relative Risk, RR)

Το επόμενο φυσικό μέτρο που μπορούμε να χρησιμοποιήσουμε είναι ο «σχετικός κίνδυνος» (relative risk) ο οποίος ορίζεται ως ο λόγος της επίπτωσης ή της θνησιμότητας δύο ομάδων

με διαφορετική έκθεση στον παράγοντα κινδύνου. Σε πίνακες 2×2 δίνεται από το λόγο των πιθανοτήτων $\pi_{j=1|i=1}$ προς $\pi_{j=1|i=2}$, δηλαδή

$$RR = \frac{\pi_E}{\pi_{\bar{E}}} = \frac{P(A|E)}{P(A|\bar{E})} = \frac{\pi_{j=1|i=1}}{\pi_{j=1|i=2}} = \frac{1 - \pi_{j=2|i=1}}{1 - \pi_{j=2|i=2}}.$$

Εκτιμάται από την ποσότητα

$$\widehat{RR} = \frac{p_E}{p_{\bar{E}}} = \frac{p_{j=1|i=1}}{p_{j=1|i=2}} = \frac{n_{11}/n_{1\bullet}}{n_{21}/n_{2\bullet}}.$$

Χρησιμοποιώντας το σχετικό κίνδυνο, ο έλεγχος ανεξαρτησίας (3.1) μπορεί να γραφτεί απλά ως $H_0 : RR = 1$ έναντι της εναλλακτικής $H_1 : RR \neq 1$. Συνεπώς, αν μοναδιαίος σχετικός κίνδυνος συνεπάγεται ανεξαρτησία μεταξύ των δύο μεταβλητών. Στην Ιατρική, όταν η μεταβλητή απόκρισης Y εκφράζει την εμφάνιση μιας νόσου και η μεταβλητή X την έκθεση στον παράγοντα κινδύνου τότε, αν ο σχετικός κίνδυνος είναι κοντά στο ένα, μπορούμε να πούμε ότι η πιθανότητα εμφάνισης της νόσου δεν επηρεάζεται από τη διαφορετική έκθεση στον παράγοντα κινδύνου X (ή δεν αλλάζει για τις διαφορετικές κατηγορίες της μεταβλητής X).

Αν $RR = a$, τότε η ερμηνεία του σχετικού κινδύνου μπορεί να γίνει με οποιαδήποτε από τις παρακάτω προτάσεις:

1. Η πιθανότητα εμφάνισης της νόσου ($Y = 1$) όταν ένα άτομο εκτεθεί στον παράγοντα κινδύνου ($X = 1$) είναι ίση με a φορές την ίδια πιθανότητα όταν δεν εκτεθεί στον παράγοντα κινδύνου ($X = 2$).
2. Αν $a > 1$ τότε: Η πιθανότητα εμφάνισης της νόσου ($Y = 1$) όταν ένα άτομο εκτεθεί στον παράγοντα κινδύνου ($X = 1$) είναι ίση με $a - 1$ φορές μεγαλύτερη (ή $(a - 1)100\%$ φορές μεγαλύτερη) από την ίδια πιθανότητα όταν δεν εκτεθεί στον κίνδυνο ($X = 2$).
3. Αν $a < 1$ τότε: Η πιθανότητα εμφάνισης της νόσου ($Y = 1$) όταν ένα άτομο εκτεθεί στον παράγοντα κινδύνου ($X = 1$) είναι ίση με $1 - a$ φορές μικρότερη (ή $(1 - a)100\%$ φορές μικρότερη) από την ίδια πιθανότητα όταν δεν εκτεθεί στον κίνδυνο ($X = 2$) (η μεταβλητή μας εδώ λέγεται προστατευτικός παράγοντας).
4. Αν $a = 1$ τότε: Η πιθανότητα εμφάνισης της νόσου ($Y = 1$) όταν ένα άτομο εκτεθεί στον παράγοντα κινδύνου ($X = 1$) είναι ίση με την ίδια πιθανότητα όταν δεν εκτεθεί στον κίνδυνο ($X = 2$) (δηλαδή η μεταβλητή X δεν επηρεάζει την εμφάνιση της νόσου συνεπώς δεν αποτελεί παράγοντα κινδύνου).

3.2.4 Λόγος Σχετικών Πιθανοτήτων (Odds Ratio, OR)

Πριν ορίσουμε την έννοια του Λόγου Σχετικών πιθανοτήτων χρειαζόμαστε να εισάγουμε και να εξηγήσουμε πλήρως την έννοια της σχετικής πιθανότητας.

3.2.4.1 Σχετική Πιθανότητα (odds)

Ο όρος “odds” χρησιμοποιείται πολύ συχνά στην Αγγλική γλώσσα υποκαθιστώντας το όρο “probability” (πιθανότητα) και συνήθως δίνει μια εκτίμηση της τύχης που έχει κάποιος να κερδίσει σε ένα αγώνα ή στοίχημα. Μαθηματικά το odds ενός ενδεχομένου A δίνεται από τον τύπο

$$\text{Odds}(A) = \frac{P(A)}{1 - P(A)}$$

και δεν είναι τίποτα άλλο από τον λόγο της πιθανότητας εμφάνισης ενός ενδεχομένου έναντι της πιθανότητας μη εμφάνισης του (δηλαδή την πιθανότητα του συμπληρωματικού, ως προς το δειγματικό χώρο, ενδεχομένου).

Η απόδοση του όρου odds στα Ελληνικά δεν ήταν εύκολη υπόθεση. Ανοίγοντας το λεξικό Αγγλικής γλώσσας του Longman (1987) βρίσκουμε τις ακόλουθες δύο ερμηνείες (σχετικές με πιθανότητες):

1. «Η πιθανότητα ότι κάτι θα συμβεί ή όχι»
2. «Πιθανότητα εκφρασμένη σε αριθμούς για χρήση σε στοιχήματα».

Πράγματι ο όρος odds χρησιμοποιείται κυρίως για στοιχήματα και δίνει τις «ευκαιρίες» νίκης μιας ομάδας ή ενός διαγωνιζόμενου. Για παράδειγμα η Ελληνική ομάδα ποδοσφαίρου πριν το Ευρωπαϊκό πρωτάθλημα ποδοσφαίρου του 2004 είχε, σύμφωνα με τα διεθνή γραφεία στοιχημάτων, πιθανότητα κατάκτησης (ή πιο σωστά αποτυχίας κατάκτησης) του πρωταθλήματος 80 προς 1 (80 : 1) δηλαδή odds = 1/80. Υπό αυτή τη δεύτερη προσέγγιση, το odds δεν είναι τίποτα άλλο από τον λόγο ή την αναλογία του αριθμού εμφάνισης ενός ενδεχομένου (εκφρασμένο είτε σε απόλυτους αριθμούς είτε ως πιθανότητες) προς τον αριθμό μη εμφάνισης του ενδεχομένου αυτού (Pereira-Maxwell, 1998, σελ. 49).

Κοιτάζοντας τώρα ένα Αγγλο-ελληνικό λεξικό, το odds μεταφράζεται απλά σαν πιθανότητα το οποίο όμως δεν αποδίδει την ακριβή έννοια του όρου. Από τα παραπάνω είδαμε ότι δεν είναι τίποτα άλλο από το λόγο της πιθανότητας ενός ενδεχομένου έναντι της πιθανότητας του συμπληρωματικού ενδεχομένου. Για το λόγο αυτό η σωστότερη και ακριβέστερη απόδοση του στα ελληνικά θα μπορούσε να γίνει με τον όρο «Λόγος πιθανοτήτων σχετικών ενδεχομένων». Επειδή η απόδοση αυτή είναι αρκετά δύσχρηστη μπορούμε να χρησιμοποιούμε τον πιο σύντομο όρο

«συμπληρωματική πιθανότητα» ο οποίος είναι πιο εύχρηστος και αποδίδει τις δύο πιο σημαντικές έννοιες του odds : την άμεση σχέση του με την πιθανότητα και τη συμπληρωματικότητα των πιθανοτήτων που συγκρίνουμε. Ο ελληνικός όρος «συμπληρωματική πιθανότητα» εμφανίστηκε στην Ελληνική βιβλιογραφία από τη συνάδελφο Ουρανία Δαφνή (βλέπε Δαφνή, 1997, σελ. 30).

Παρόλα αυτά υπάρχει ένα πρόβλημα με την ονομασία «συμπληρωματική πιθανότητα». Αναφέρεται μόνο στη σύγκριση πιθανοτήτων συμπληρωματικών ενδεχομένων. Το γεγονός αυτό κάνει απαγορευτικό την χρήση του παραπάνω όρου για τα odds που προκύπτουν από σύγκριση επιπέδων μη-δίτιμων κατηγορικών μεταβλητών εφόσον αυτά αναφέρονται σε λόγους πιθανοτήτων μη-συμπληρωματικών ενδεχομένων (δηλ. $odds_{12} = \pi_1/\pi_2$ με $\pi_1 + \pi_2 < 1$). Για το λόγο αυτό τα τελευταία χρόνια έχει επικρατήσει ο όρος «σχετική πιθανότητα» τον οποίο θα χρησιμοποιήσουμε και εδώ. Ο όρος αυτός αποδίδει δύο σημαντικά στοιχεία του odds : α) ότι είναι ένα μέτρο πιθανότητας ενός ενδεχομένου και β) ότι το μέτρο αυτό το βλέπουμε συγκριτικά με ένα δεύτερο ενδεχόμενο.

Πρακτικά η σχετική πιθανότητα είναι $1 - 1$ μετασχηματισμός της αρχικής πιθανότητας εφόσον

$$odds = \frac{\pi}{1 - \pi} \Leftrightarrow \pi = \frac{odds}{1 + odds}.$$

που σημαίνει ότι μεταφέρει ακριβώς την ίδια πληροφορία με διαφορετικό τρόπο. Η ερμηνεία της θεωρείται πιο εύκολη από την απλή πιθανότητα διότι συγκρίνει την πιθανότητα εμφάνισης ενός ενδεχομένου με την πιθανότητα μη εμφάνισης. Έτσι όταν η Ελληνική ομάδα είχε σχετική πιθανότητα να μην κερδίσει το ευρωπαϊκό πρωτάθλημα του 2004 80 : 1 αυτό σήμαινε ότι πιθανότητα να μην κερδίσει το πρωτάθλημα ήταν ίση με 80 φορές την πιθανότητα να κερδίσει. Πιο απλά η πιθανότητα να κερδίσει ήταν μόλις $(1/(80 + 1) =) 0.0123$.

Στην Ιατρική η σχετική πιθανότητα επικεντρώνεται στην πιθανότητα εμφάνισης της νόσου σε σχέση με τη μη εμφάνιση της. Η ερμηνεία της σχετικής πιθανότητας δίνεται από τις ακόλουθες προτάσεις

1. Αν $odds = 1$ (ή 1 : 1) τότε οι πιθανότητες εμφάνισης ή μη εμφάνισης του ενδεχομένου που εξετάζουμε είναι ίσες (δηλαδή 50%).
2. Αν $odds = a$ τότε η πιθανότητα εμφάνισης του ενδεχομένου που εξετάζουμε είναι ίση με a φορές την πιθανότητα μη εμφάνισής του.
3. Αν $odds = a$ και
 - (α') $a > 1$ τότε η πιθανότητα εμφάνισης του ενδεχομένου που εξετάζουμε είναι $a - 1$ (ή $(a - 1)100\%$) φορές μεγαλύτερη από την πιθανότητα μη εμφάνισης του.

(β) $\alpha < 1$ τότε η πιθανότητα εμφάνισης του ενδεχομένου που εξετάζουμε είναι $1 - \alpha$ (ή $(1 - \alpha)100\%$) φορές μικρότερη από την πιθανότητα μη εμφάνισης του.

Συνεπώς όταν μιλάμε για την σχετική πιθανότητα εμφάνισης της νόσου τότε:

- $odds = 1$ συνεπάγεται ίση πιθανότητα εμφάνισης και μη εμφάνισης της νόσου.
- $odds > 1$ συνεπάγεται η εμφάνιση της νόσου είναι πιο πιθανή από την πιθανότητα μη εμφάνισης της νόσου.
- $odds < 1$ συνεπάγεται η εμφάνιση της νόσου είναι λιγότερο πιθανή από την πιθανότητα μη εμφάνισης της νόσου.

Επιστρέφοντας στους 2×2 πίνακες συνάφειας, μπορούμε να ορίσουμε δύο σχετικές πιθανότητες ή odds (ένα για κάθε επίπεδο της μεταβλητής X) από τις σχέσεις:

$$odds(X = 1) = \frac{\pi_{1|i=1}}{\pi_{2|i=1}} = \frac{\pi_{1|i=1}}{1 - \pi_{1|i=1}} = \frac{\pi_{11}}{\pi_{12}} \quad \text{και} \quad odds(X = 2) = \frac{\pi_{1|i=2}}{\pi_{2|i=2}} = \frac{\pi_{1|i=2}}{1 - \pi_{1|i=2}} = \frac{\pi_{21}}{\pi_{22}}.$$

Οι παραπάνω ποσότητες μπορούν να εκτιμηθούν ως εξής

$$\widehat{odds}(X = 1) = \frac{p_{1|i=1}}{p_{2|i=1}} = \frac{n_{11}/n_{1\bullet}}{n_{12}/n_{1\bullet}} = \frac{n_{11}}{n_{12}} \quad \text{και} \quad \widehat{odds}(X = 2) = \frac{p_{1|i=2}}{p_{2|i=2}} = \frac{n_{21}/n_{2\bullet}}{n_{22}/n_{2\bullet}} = \frac{n_{21}}{n_{22}}.$$

Γυρίζοντας στις κύριες μεταβλητές της Ιατρικής έρευνας, μας ενδιαφέρουν οι σχετικές πιθανότητες εμφάνισης της νόσου όταν έχουμε έκθεση ή όχι στον παράγοντα κινδύνου. Συνεπώς μας ενδιαφέρουν οι σχετικές πιθανότητες

$$odds_E = \frac{P(A|E)}{1 - P(A|E)} = \frac{\pi_E}{1 - \pi_E} \quad \text{και} \quad odds_{\bar{E}} = \frac{P(A|\bar{E})}{1 - P(A|\bar{E})} = \frac{\pi_{\bar{E}}}{1 - \pi_{\bar{E}}}$$

οι οποίες εκτιμούνται από τους τύπους

$$\widehat{odds}_E = \frac{p_E}{1 - p_E} \quad \text{και} \quad \widehat{odds}_{\bar{E}} = \frac{p_{\bar{E}}}{1 - p_{\bar{E}}}.$$

3.2.4.2 Σύγκριση Σχετικών πιθανοτήτων με τη χρήση του Odds Ratio.

Στην παραπάνω παράγραφο ορίσαμε τις σχετικές πιθανότητες για να μπορέσουμε να προχωρήσουμε στον ορισμό του «λόγου σχετικών πιθανοτήτων» ή “odds ratio” στην Αγγλική γλώσσα. Όπως φαίνεται και από την ονομασία του δεν είναι τίποτα άλλο από μια σύγκριση odds για διαφορετικές τιμές ή επίπεδα μια δεύτερης μεταβλητής X . Έτσι λοιπόν ο λόγος σχετικών πιθανοτήτων του $X = 1$ έναντι του $X = 2$ δίνεται από τον τύπο

$$OR = \frac{odds(X = 1)}{odds(X = 2)} = \frac{\pi_{1|i=1}/\pi_{2|i=1}}{\pi_{1|i=2}/\pi_{2|i=2}} = \frac{\pi_{11}/\pi_{12}}{\pi_{21}/\pi_{22}} = \frac{\pi_{11}\pi_{22}}{\pi_{21}\pi_{12}}. \quad (3.2)$$

Από τα παραπάνω βλέπουμε ότι δεν είναι τίποτα άλλο από το χιαστό λόγο των πιθανοτήτων που εμφανίζονται σε ένα 2×2 πίνακα συνάφειας. Για το λόγο αυτό πολλές φορές ονομάζεται και «λόγος σταυρωτού ή χιαστού πολλαπλάσιου» (“crossproduct ratio”). Στην Ιατρική Στατιστική χρησιμοποιούνται και οι όροι: «προσεγγιστικός σχετικός κίνδυνος» ή «σχετικός λόγος» ή απλά «σχετικός κίνδυνος» (για λεπτομέρειες βλ. Τριχόπουλος, Τζώνου και Κατσουγιάννη, 2000, σελ. 70–75) οι οποίοι όμως δεν αποδίδουν πλήρως την έννοια του odds ratio. Τα τελευταία χρόνια η ίδια σχολή χρησιμοποιεί τον όρο «σχετικός λόγος συμπληρωματικών πιθανοτήτων» ο οποίος είναι μακροσκελής και δεν είναι τόσο εύχρηστος και γενικός όσο ο όρος «λόγος σχετικών πιθανοτήτων».

Στους 2×2 πίνακες συνάφειας ο λόγος σχετικών πιθανοτήτων (ΛΣΠ) εκτιμάται από τον εκτιμητή

$$\widehat{OR} = \frac{p_{11}p_{22}}{p_{21}p_{12}} = \frac{n_{11}n_{22}}{n_{21}n_{12}}.$$

δηλαδή από το λόγο των χιαστών πολλαπλάσιων του πίνακα συχνότητων. Αν θέλουμε να ορίσουμε το ΛΣΠ ως συνάρτηση των πιθανοτήτων εμφάνισης της νόσου τότε $OR = \frac{\pi_E}{1-\pi_E} \times \frac{1-\pi_{\bar{E}}}{\pi_{\bar{E}}}$ και ενώ ο αντίστοιχος εκτιμητής είναι $\widehat{OR} = \frac{p_E}{1-p_E} \times \frac{1-p_{\bar{E}}}{p_{\bar{E}}}$.

Ο ΛΣΠ χρησιμοποιείται ευρέως κυρίως στην Ιατρική. Η δημοτικότητα του ΛΣΠ οφείλεται στη σχετικά εύκολη ερμηνεία του, στο γεγονός ότι μπορεί να υπολογιστεί και σε προοπτικές και σε αναδρομικές μελέτες, στο ότι αναφέρεται σε συγκρίσεις σχετικών πιθανοτήτων (odds) και στο ότι προκύπτει άμεσα από τις παραμέτρους των μοντέλων λογιστικής παλινδρόμησης που θα εξετάσουμε παρακάτω.

Η σύνδεση των ΛΣΠ (OR) με την έννοια της ανεξαρτησίας σε πίνακες 2×2 είναι άμεση αφού

$$\text{«ανεξαρτησία»} \Leftrightarrow OR = 1.$$

για το λόγο αυτό ο έλεγχος (3.1) μπορεί να γραφτεί και ως

$$H_0 : OR = 1 \text{ έναντι της εναλλακτικής } H_1 : OR \neq 1.$$

Εναλλακτικά, αν $OR > 1$ τότε το πρώτο επίπεδο της κατηγορικής μεταβλητής (δηλαδή η πρώτη γραμμή του πίνακα συνάφειας) είναι πιο πιθανή από το δεύτερο επίπεδο. Για παράδειγμα αν $OR = 4$ τότε το πρώτο επίπεδο της μεταβλητής X (πρώτη γραμμή του πίνακα, $X = 1$) έχει σχετική πιθανότητα τετραπλάσια της σχετικής πιθανότητας του δεύτερου επιπέδου της μεταβλητής X (δεύτερης γραμμής, δηλ. $X = 2$). Αν $OR = 1/4$ τότε η δεύτερη γραμμή του πίνακα (δηλαδή για $X = 2$) έχει σχετική πιθανότητα τετραπλάσια της σχετικής πιθανότητας της πρώτης γραμμής (δηλαδή για $X = 1$). Από τα παραπάνω είναι προφανές ο ΛΣΠ αλλάζει (και η αντίστοιχη ερμηνεία του) αν αλλάξουμε τη σειρά των επιπέδων της μιας από τις δύο μεταβλητές. Αν όμως αλλάξουμε τη σειρά και στις δύο μεταβλητές και δημιουργήσουμε τον αντίστοιχο πίνακα ο ΛΣΠ

δεν μεταβάλλεται. Έτσι μπορούμε να ερμηνεύσουμε τον ΛΣΠ με τους ακόλουθους ισοδύναμους τρόπους:

- «η σχετική πιθανότητα του $Y = 1$ όταν $X = 1$ είναι ίση με OR φορές την αντίστοιχη συμπληρωματική πιθανότητα όταν $X = 2$ »
- «η σχετική πιθανότητα του $Y = 2$ όταν $X = 2$ είναι ίση με OR φορές την αντίστοιχη συμπληρωματική πιθανότητα όταν $X = 1$ ».

Επιστρέφοντας στις συνηθισμένες μεταβλητές των Ιατρικών μελετών (νόσος και παράγοντας κινδύνου) τότε η ερμηνεία μπορεί να γραφτεί με κάποιον από τους ακόλουθους εναλλακτικού τρόπους:

1. Αν $OR = a$ τότε η σχετική πιθανότητα της νόσου ($Y = 1$) όταν ένα άτομο εκτεθεί στον κίνδυνο ($X = 1$) είναι ίση με a φορές την ίδια σχετική πιθανότητα όταν δεν εκτεθεί στον κίνδυνο ($X = 2$). Η ισοδύναμα: η σχετική πιθανότητα μη εμφάνισης της νόσου ($Y = 2$) όταν ένα άτομο εκτεθεί στον κίνδυνο ($X = 2$) είναι ίση με a φορές την ίδια σχετική πιθανότητα όταν εκτεθεί στον κίνδυνο ($X = 1$).
2. Αν $OR = a > 1$ τότε η σχετική πιθανότητα της νόσου ($Y = 1$) όταν ένα άτομο εκτεθεί στον κίνδυνο ($X = 1$) είναι $a - 1$ (ή $\{a - 1\}100\%$) φορές μεγαλύτερη της ίδιας σχετικής πιθανότητας όταν δεν εκτεθεί στον κίνδυνο ($X = 2$).
3. Αν $OR = a < 1$ τότε η σχετική πιθανότητα της νόσου ($Y = 1$) όταν ένα άτομο εκτεθεί στον κίνδυνο ($X = 1$) είναι $1 - a$ (ή $\{1 - a\}100\%$) φορές μικρότερη της ίδιας σχετικής πιθανότητας όταν δεν εκτεθεί στον κίνδυνο ($X = 2$).

Ο ΛΣΠ είναι στενά συνδεδεμένος με το σχετικό κίνδυνο στην Ιατρική διότι, κάτω από προ-ϋποθέσεις, αποτελεί προσεγγιστική εκτίμηση του στις μελέτες μαρτύρων-ασθενών. Γενικά όταν $OR > 1$ τότε ο κίνδυνος αυξάνει με την παρουσία του παράγοντα X (άρα ο X παράγοντας ή μεταβλητή είναι επιβαρυντικός για τη νόσο ή αλλιώς παράγοντας κινδύνου) ενώ αν $OR < 1$ τότε ο κίνδυνος μειώνεται με την παρουσία του X (άρα ο X παράγοντας ή μεταβλητή είναι προστατευτικός για τη νόσο).

Στην στατιστική πολλές φορές χρησιμοποιούμε και το λογάριθμο του ΛΣΠ. Ο λόγος είναι ότι μπορούμε να εξετάσουμε την κατανομή του πιο εύκολα. Επιπλέον ο λογάριθμος είναι αυτός που επίσης χρησιμοποιείται στα μοντέλα λογιστικής παλινδρόμησης. Στην περίπτωση αυτή η υπόθεση της ανεξαρτησίας (3.1) μπορεί να γραφεί ως

$$H_0 : \log OR = 0 \quad \text{έναντι της εναλλακτικής} \quad H_1 : \log OR \neq 0.$$

Ο παραπάνω ορισμός του ΛΣΠ βασίζεται στη νόσο ως μεταβλητή απόκρισης. Πιο συγκεκριμένα, λαμβάνει υπόψη της ως επιτυχία την εμφάνιση της νόσου, και μετα συγκρίνει τις σχετικές πιθανότητες της νόσου μεταξύ των ατόμων που έχουν ή όχι εκτεθεί σε ένα παράγοντα κινδύνου. Ο λόγος των σχετικών πιθανοτήτων εμφάνισης μιας νόσου για άτομα διαφορετικής έκθεσης σε ένα παράγοντα ονομάζεται και λόγος σχετικής πιθανότητας μιας νόσου (disease odds ratio, Rosner, 1994, σελ. 365).

Εάν λάβουμε ως επιτυχία την έκθεση στον κίνδυνο και υπολογίσουμε το ΛΣΠ ως το λόγο των σχετικών πιθανοτήτων έκθεσης στον κίνδυνο για τους ασθενείς και μάρτυρες τότε υπολογίζουμε τον ΛΣΠ έκθεσης σε ένα παράγοντα (exposure odds ratio, Rosner, 1994, σελ. 366). Στην ουσία στους 2×2 πίνακες οι δύο ΛΣΠ συμπίπτουν και είναι σαν να βλέπουμε το ίδιο αντικείμενο από διαφορετική οπτική γωνία. Συνέπεια του παραπάνω αποτελέσματος είναι και το γεγονός ότι αν αλλάξουμε τη διάταξη των μεταβλητών σε ένα 2×2 πίνακα (δηλαδή ποια μεταβλητή καθορίζει τις γραμμές και ποια μεταβλητή τις στήλες του πίνακα) τότε ο ΛΣΠ θα παραμείνει αμετάβλητος.

3.2.5 Εκτίμηση του σχετικού κινδύνου και του ΛΣΠ με τη Χρήση SPSS.

Στο SPSS μπορούμε να υπολογίσουμε το σχετικό κίνδυνο και το λόγο σχετικών πιθανοτήτων για πίνακες 2×2 . Αυτό μπορεί να γίνει στο μενού

Analyze>Descriptive Statistics>Crosstabs.

Αφού εισάγουμε τις μεταβλητές που επιθυμούμε στις στήλες και στις γραμμές, επιλέγουμε το δεύτερο σε σειρά κουτί επιλογών που εμφανίζεται στο κάτω μέρος της οθόνης με τίτλο **Statistics** και τσεκάρουμε την εκτύπωση των δεικτών κινδύνου με ένδειξη **Risk**. Προσοχή οι δείκτες κινδύνου (σχετικός κίνδυνος και λόγος σχετικών πιθανοτήτων) υπολογίζονται μόνο για 2×2 πίνακες.

3.3 Σχέση Σχετικού Κινδύνου και ΛΣΠ (Odds Ratio)

Ένας από τους λόγους που ο ΛΣΠ έγινε δημοφιλής στην Ιατρική στατιστική (ή Βιοστατιστική) είναι και το γεγονός ότι μπορεί να υπολογιστεί και στους δύο τύπους μελετών που χρησιμοποιούνται στην ιατρική δηλαδή και στις προοπτικές (μελέτες κοορτής) και στις αναδρομικές (μελέτες μαρτύρων-ασθενών) σε αντίθεση με το σχετικό κίνδυνο που μπορεί να υπολογιστεί μόνο σε προόπτικες μελέτες. Επιπλέον, ο ΛΣΠ είναι ασυμπτωτικά ίσος με το σχετικό κίνδυνο όταν η πιθανότητα εμφάνισης της νόσου είναι μικρή δηλαδή στις σπάνιες αρρώστιες. Όπως είδαμε

και παραπάνω

$$\begin{aligned} OR &= \frac{Odds_E}{Odds_{\bar{E}}} = \frac{Odds(X=1)}{Odds(X=2)} = \frac{\pi_{j=1|1}/\pi_{j=2|1}}{\pi_{j=1|2}/\pi_{j=2|2}} = \frac{\pi_{j=1|1}}{\pi_{j=1|2}} \times \frac{\pi_{j=2|2}}{\pi_{j=2|1}} \\ &= \frac{RR(Y=1)}{RR(Y=2)}. \end{aligned}$$

Αν το γεγονός που εξετάζουμε έχει πιθανότητα εμφάνισης πολύ μικρή (για παράδειγμα μια σπάνια ασθένεια) τότε $\pi_{j=2|1} \approx 1$ και $\pi_{j=2|2} \approx 1$ άρα και $RR(Y=2) \approx 1$ καταλήγοντας στο αποτέλεσμα $OR \approx RR(Y=1)$. Σύμφωνα με τον Rosner (1994, σελ. 368) για να έχουμε ικανοποιητική προσέγγιση του σχετικού κινδύνου από το λόγο σχετικών πιθανοτήτων θα πρέπει ο επιπολασμός (δηλ. πιθανότητα) της νόσου να είναι μικρότερη του 0.10.

Από τα παραπάνω μπορούμε να κατανοήσουμε και το λόγο για τον οποίο ο σχετικός κίνδυνος δεν μπορεί να υπολογιστεί σε αναδρομικές μελέτες. Για να υπολογίσουμε το σχετικό κίνδυνο χρειάζεται να γνωρίζουμε (ή να μπορούμε να εκτιμήσουμε) την πιθανότητα εμφάνισης της νόσου $\pi_{j=1|1}$ και $\pi_{j=1|2}$. Κάτι τέτοιο δεν είναι δυνατόν στις μελέτες μαρτύρων-ασθενών όπου το μέγεθος τόσο των ασθενών όσο και των μαρτύρων είναι προκαθορισμένο λόγω του σχεδιασμού της μελέτης άρα δεν μπορούν να θεωρηθούν ως τυχαίες μεταβλητές. Αντίθετα στις προοπτικές μελέτες, όπου λαμβάνεται δείγμα αντιπροσωπευτικό του υπό μελέτη πληθυσμού, αυτές οι ποσότητες μπορούν να εκτιμηθούν άρα να υπολογιστεί και ο σχετικός κίνδυνος. Από την άλλη μεριά, και ο ΛΣΠ, που δίνεται από τον τύπο 3.2, φαίνεται να προαπαιτεί την εκτίμηση αυτών των πιθανοτήτων. Αναλύοντας όμως την εξίσωση (3.2) βλέπουμε ότι

$$\begin{aligned} OR &= \frac{\pi_{1|i=1}/\pi_{2|i=1}}{\pi_{1|i=2}/\pi_{2|i=2}} = \frac{\pi_{11}\pi_{22}}{\pi_{21}\pi_{12}} = \frac{\pi_{i=1|1}\pi_{i=2|2}}{\pi_{i=1|2}\pi_{i=2|1}} \Leftrightarrow \\ OR &= \frac{\pi_{i=1|1}\pi_{i=2|2}}{\pi_{i=1|2}\pi_{i=2|1}}. \end{aligned} \quad (3.3)$$

Συνεπώς ο ΛΣΠ ορίζεται επαρκώς αν αντί για τις δεσμευμένες πιθανότητες $\pi_{j=1|i}$ για $i = 1, 2$ γνωρίζουμε τις πιθανότητες $\pi_{i|j=1}$ για $i = 1, 2$ οι οποίες μπορούν να εκτιμηθούν από μια μελέτη μαρτύρων-ασθενών. Τα παραπάνω αποδεικνύουν ότι ο ΛΣΠ μπορεί να εκτιμηθεί χωρίς πρόβλημα ανεξαρτήτως ποια από τις δύο μεταβλητές (X ή Y) θα είναι τυχαία. Συνεπώς μπορεί να χρησιμοποιηθεί και να υπολογιστεί και στις αναδρομικές και στις προοπτικές μελέτες.

3.4 Υπολογισμός του Σχετικού Κινδύνου στις Μελέτες Μαρτύρων-Ασθενών

Επιστρέφοντας στο σχετικό κίνδυνο έχουμε:

$$RR = \frac{P(A|E)}{P(A|\bar{E})} = \frac{\pi_E}{\pi_{\bar{E}}}.$$

Όμως στις μελέτες μαρτύρων-ασθενών η κατανομή των ασθενών (A) και μαρτύρων \bar{A} είναι προκαθορισμένη και συνεπώς οι παραπάνω πιθανότητες δεν μπορούν να εκτιμηθούν. Για το λόγο αυτό θα χρησιμοποιήσουμε το θεώρημα του Bayes για να εκφράσουμε το σχετικό κίνδυνο ως πιθανότητες ή κατανομή του παράγοντα κινδύνου X (δηλαδή των ενδεχομένων E και \bar{E}). Συνεπώς

$$\begin{aligned} RR &= \frac{P(A|E)}{P(A|\bar{E})} = \frac{P(A, E)/P(E)}{P(A, \bar{E})/P(\bar{E})} \\ &= \frac{P(E|A)P(A)/P(E)}{P(\bar{E}|A)P(A)/P(\bar{E})} = \frac{P(E|A)}{P(\bar{E}|A)} \times \frac{P(\bar{E})}{P(E)}. \end{aligned}$$

Στην παραπάνω ισότητα, η πιθανότητα έκθεσης και μη έκθεσης αναφέρεται στο γενικό πληθυσμό. Σε μία μελέτη μαρτύρων ασθενών, λόγω της δειγματοληψίας, αυτή πιθανότητα δεν εκτιμάται σωστά. Άρα με τη χρήση του θεωρήματος της ολικής πιθανότητας μπορούμε να γράψουμε

$$\begin{aligned} RR(A) &= \frac{P(E|A)}{1 - P(E|A)} \times \frac{P(\bar{E}|A)P(A) + P(\bar{E}|\bar{A})P(\bar{A})}{P(E|A)P(A) + P(E|\bar{A})P(\bar{A})} \\ &= \frac{P(E|A)}{1 - P(E|A)} \times \frac{[1 - P(E|A)]P(A) + [1 - P(E|\bar{A})][1 - P(A)]}{P(E|A)P(A) + P(E|\bar{A})[1 - P(A)]}. \end{aligned}$$

Στην παραπάνω ισότητα πρέπει να γνωρίζουμε τις πιθανότητες έκθεσης στον παράγοντα κινδύνου δεδομένου ότι κάποιος έχει ή όχι τη νόσο δηλαδή είναι ασθενής ή μάρτυρας - πιθανότητες $P(E|A)$ και $P(E|\bar{A})$. Επιπλέον πρέπει να γνωρίζουμε και τον επιπολασμό της νόσου $P(A)$ ο οποίος όμως δεν μπορεί να εκτιμηθεί από μια μελέτη μαρτύρων-ασθενών αλλά από μία συγχρονική ή προοπτική μελέτη με τυχαίο δείγμα από το γενικό πληθυσμό.

3.5 Παραδείγματα

3.5.1 Μυοκαρδιακή Ανεπάρκεια και Αντισύλληψη (Μελέτη Μαρτύρων-Ασθενών).

Παράδειγμα 3.1 Τα δεδομένα του 2×2 Πίνακα 3.1 προέρχονται από την ερευνητική δουλειά των Mann *et. al* (1975, Brit. J. Med.). Στη έρευνα αυτή 58 γυναίκες κάτω των 45 χρονών με μυοκαρδιακή ανεπάρκεια εξετάσθηκαν ως προς τη χρήση ή όχι αντισυλληπτικού χαπιού. Το δείγμα προερχόταν από δύο νοσοκομεία της Αγγλίας και της Ουαλίας την περίοδο 1968-1972. Επίσης επιλέχτηκε περίπου τριπλάσιος αριθμός μαρτύρων από τα ίδια νοσοκομεία που δεν είχαν την νόσο της μυοκαρδιακής ανεπάρκειας. Τα δεδομένα αυτά δίνονται επίσης στις σελ. 11-13 από τον Agresti (1990).

X: Αντισυλ.Χάπι	Y: Μυοκαρδιακή Ανεπάρκεια		Περ.Κατ. X
	1: Ναι	2: Όχι	
1: Ναι	23	34	57
2: Όχι	35	132	167
Περ.Κατ. Y	58	166	224

Πίνακας 3.1: Δεδομένα Παραδείγματος 3.1: Μυοκαρδιακή Ανεπάρκεια και Αντισυλληπτικό Χάπι (Mann *et. al* , 1975).

3.5.1.1 Εισαγωγή των Δεδομένων στο SPSS.

Για να εισάγουμε τον παραπάνω πίνακα στο SPSS ή σε άλλο στατιστικό πακέτο συνήθως χρειαζόμαστε 3 στήλες. Μία με τη μεταβλητή X, μία με την μεταβλητή Y και μία με τις αντίστοιχες συχνότητες. Το πλήθος των γραμμών θα είναι ίσος με το συνδυασμό των πιθανών τιμών των δύο μεταβλητών. Συνεπώς εδώ θα έχουμε 4 γραμμές (2×2). Άρα τα παραπάνω δεδομένα θα περαστούν στο SPSS με την ακόλουθη μορφή.

contrac	myocard	counts
1	1	23
1	2	34
2	1	35
2	2	132

Μετά πρέπει να ορίσουμε τη μεταβλητή counts ως βάρος στο μενού

Data>Weight Cases

Η ανάλυση για τους πίνακες διπλής εισόδου γίνεται από το μενού

Analyze>Descriptive Stats>Crosstabs

τοποθετώντας στο row(s) και στο column(s) τις μεταβλητές contrac και myocard αντίστοιχα.

Σε μερικά πακέτα όπως το Splus ή Stat-exact μπορούμε να εισάγουμε τον παραπάνω πίνακα κατευθείαν όπως δίνονται στον Πίνακα 3.1.

3.5.1.2 Από κοινού και δεσμευμένες Συναρτήσεις Πιθανότητας.

Η από κοινού συνάρτηση κατανομής για το παραπάνω παράδειγμα δίνεται από τον ακόλουθο πίνακα

X: Αντισυλ.Χάπι	Y: Μυοκαρδιακή Ανεπάρκεια		Περ.Κατ. X
	1: Ναι	2: Όχι	
1: Ναι	$p_{11} = \frac{23}{224} = 0.103$	$p_{12} = \frac{34}{224} = 0.152$	$p_{1\bullet} = \frac{23+34}{224} = 0.255$
2: Όχι	$p_{21} = \frac{35}{224} = 0.156$	$p_{22} = \frac{132}{224} = 0.589$	$p_{2\bullet} = \frac{35+132}{224} = 0.745$
Περ.Κατ. Y	$p_{\bullet 1} = \frac{24+35}{224} = 0.259$	$p_{\bullet 2} = \frac{34+132}{224} = 0.741$	

Από το SPSS μπορούμε να υπολογίσουμε τον παραπάνω πίνακα αν στην επιλογή **cells** του μενού **Crosstabs** επιλέξουμε την εμφάνιση μόνο των **Percentages:Total**.

Στο συγκεκριμένο παράδειγμά μας, εφόσον η μεταβλητή απόκρισης είναι η μυοκαρδιακή ανεπάρκεια, μας ενδιαφέρει να συγκρίνουμε την κατανομή της νόσου μεταξύ των ατόμων που χρησιμοποίησαν αντισύλληψη με την αντίστοιχη κατανομή στην ομάδα των ατόμων που δεν χρησιμοποίησαν αντισύλληψη. Η σύγκριση αυτή θα μας δώσει και την επίδραση της χρήσης αντισυλληπτικού χαπιού στην πιθανότητα εμφάνισης της νόσου. Η παραπάνω σύγκριση σημαίνει ότι θα συγκρίνουμε τις δεσμευμένες κατανομές $P(Y|X = 1)$ και $P(Y|X = 2)$ (για τις ομάδες που χρησιμοποιούν ή όχι αντισυλληπτικό χάπι αντίστοιχα). Οι κατανομές αυτές μπορούν εύκολα να εξαχθούν στο SPSS από το υπομενού **Cells** του μενού **Crosstabs** επιλέγοντας τα ποσοστά ανά γραμμή (**Percentages:Row**). Έτσι οι δεσμευμένες κατανομές της νόσου δίνονται από τον ακόλουθο πίνακα:

X: Αντισυλ.Χάπι	Y: Μυοκαρδιακή Ανεπάρκεια	
	1: Ναι	2: Όχι
1: Ναι	$p_{1 i=1} = \frac{23}{57} = 0.404$	$p_{2 i=1} = \frac{34}{57} = 0.596$
2: Όχι	$p_{1 i=2} = \frac{35}{167} = 0.210$	$p_{2 i=2} = \frac{132}{167} = 0.790$
Περ.Κατ. Y	$p_{\bullet 1} = \frac{24+35}{224} = 0.259$	$p_{\bullet 2} = \frac{34+132}{224} = 0.741$

Πρακτικά, στον παραπάνω πίνακα θέλουμε να συγκρίνουμε τις πιθανότητες (ή τα ποσοστά) εμφάνισης της νόσου στις δύο ομάδες δηλαδή το 0.404 με το 0.210 (40.4% με 21%). Αυτό είναι ισοδύναμο με το έλεγχο της ισότητας των πιθανοτήτων (ή ποσοστών) της μη εμφάνισης της νόσου (δηλαδή του 0.596 με το 0.79).

Στο παραπάνω πρόβλημα όμως η κατανομή της μεταβλητής απόκρισης (μυοκαρδιακής ανεπάρκειας) είναι σταθερή (προκαθορισμένη) από το σχεδιασμό του πειράματος - μελέτης. Για το λόγο αυτό, η εξέταση των παραπάνω κατανομών δεν έχει νόημα. Αυτό που μπορούμε να ελέγξουμε, χωρίς όμως να μπορεί να μας μεταφέρει όλη τη επιθυμητή πληροφορία, είναι οι κατανομές $P(X|Y)$ εφόσον το X είναι όντως τυχαίο. Αυτές οι μελέτες λέγονται αναδρομικές και στη συγκεκριμένη περίπτωση μελέτη μαρτύρων - ασθενών (ή περιπτώσεων). Ο πίνακας που δίνει τις δεσμευμένες κατανομές ως προς τη X μεταβλητή (χρήση αντισυλληπτικού χαπιού) δίνεται στο SPSS από το **Cells** του μενού **Crosstabs** επιλέγοντας τα ποσοστά ανά στήλη

(Percentages:Column). Έτσι οι δεσμευμένες κατανομές της χρήσης του χαπιού δίνονται από τον ακόλουθο πίνακα :

X: Αντισυλ.Χάπι	Y: Μυοκαρδιακή Ανεπάρκεια		Περ.Κατ.
	1: Ναι	2: Όχι	X
1: Ναι	$p_{1 j=1} = \frac{23}{58} = 0.397$	$p_{1 j=2} = \frac{34}{166} = 0.205$	$p_{1\bullet} = \frac{23+34}{224} = 0.255$
2: Όχι	$p_{2 j=1} = \frac{35}{58} = 0.603$	$p_{2 j=2} = \frac{132}{166} = 0.795$	$p_{1\bullet} = \frac{35+132}{224} = 0.745$

Προσοχή εδώ πρέπει να δοθεί στη ερμηνεία των παραπάνω αποτελεσμάτων. Έτσι εδώ μπορούμε να συγκρίνουμε το ποσοστό των ασθενών που χρησιμοποιούν αντισυλληπτικό χάπι (39.7%) με το αντίστοιχο ποσοστό των μαρτύρων (20.5%). Με χρήση του θεωρήματος του Bayes μπορούμε να υπολογίσουμε και τις δεσμευμένες κατανομές $P(Y|X)$ που μας ενδιαφέρουν αλλά χρειάζεται να ξέρουμε το ποσοστό της ασθένειας στον πληθυσμό (δηλαδή τον επιπολασμό).

3.5.1.3 Εκτίμηση ΛΣΠ και Συνάρτησης Κινδύνου.

Το επόμενο βήμα είναι να υπολογίσουμε το σχετικό κίνδυνο. Τα αποτελέσματα (όπως δίνονται από το SPSS) είναι τα ακόλουθα

Από τα παραπάνω βλέπουμε ότι

1. $OR = 2.551$:

(α') Η σχετική πιθανότητα εμφάνισης της μυοκαρδιακής ανεπάρκειας ($Y = 1$) σε μια γυναίκα που χρησιμοποιεί αντισυλληπτικό χάπι ($X = 1$) είναι ίση με 2.55 φορές

την αντίστοιχη πιθανότητα σε μια γυναίκα που δε χρησιμοποιεί αντισυλληπτικό χάπι ($X = 2$).

(β') Αν αλλάξουμε τη σειρά των κατηγοριών στις δύο μεταβλητές τότε δεν αλλάζει ο ΛΣΠ. Σε αυτή την περίπτωση η ερμηνεία μπορεί να γίνει ως εξής: η σχετική πιθανότητα μη εμφάνισης της μυοκαρδιακής ανεπάρκειας ($Y = 2$) σε μια γυναίκα που δεν χρησιμοποιεί αντισυλληπτικό χάπι ($X = 2$) είναι ίση με 2.55 φορές την αντίστοιχη πιθανότητα μιας γυναίκας που χρησιμοποιεί αντισυλληπτικό χάπι ($X = 1$).

(γ') Πιο απλά θα μπορούσαμε να πούμε ότι η χρήση του αντισυλληπτικού χαπιού από τις γυναίκες αυξάνει τη (σχετική) πιθανότητα (ή τον κίνδυνο εμφάνισης) μυοκαρδιακής ανεπάρκειας.

(δ') Οι παραπάνω ερμηνείες καταλήγουν στο ίδιο συμπέρασμα: ότι η χρήση του αντισυλληπτικού χαπιού αποτελεί παράγοντα κινδύνου για τη μυοκαρδιακή ανεπάρκεια.

2. $RR(myocard = 1) = \frac{23/57}{35/167} = \frac{0.404}{0.210} = 1.936$: δηλαδή η πιθανότητα μυοκαρδιακής ανεπάρκειας για μια γυναίκα που χρησιμοποιεί αντισυλληπτικό χάπι είναι 93% υψηλότερη από την αντίστοιχη πιθανότητα μιας γυναίκας που δε χρησιμοποιεί τέτοιο είδος αντισύλληψης.

3. $RR(myocard = 2) = \frac{34/57}{132/167} = \frac{0.596}{0.790} = 0.759$: δηλαδή η πιθανότητα να μην εμφανίσει τη νόσο της μυοκαρδιακής ανεπάρκειας μια γυναίκα που χρησιμοποιεί αντισυλληπτικό χάπι είναι 24% μικρότερη της αντίστοιχης πιθανότητας μιας γυναίκας που δε χρησιμοποιεί τέτοιου είδους αντισύλληψη.

4. ΠΡΟΣΟΧΗ: Παρόλο που ερμηνεύσαμε παραπάνω το σχετικό κίνδυνο σε μία μελέτη μαρτύρων-ασθενών αυτό δεν είναι σωστό παρά μόνο στις προοπτικές μελέτες. Και αυτό γιατί η κατανομή της μεταβλητής Y δηλαδή της νόσου έχει προκαθοριστεί από τον ερευνητή και ως εκ-τούτου δεν είναι τυχαία και συνεπώς δεν μπορεί να εκτιμηθεί παρά μόνο με τη χρήση του θεωρήματος του Bayes και αν γνωρίζουμε πια είναι η πιθανότητα της μυοκαρδιακής ανεπάρκειας στο γενικό πληθυσμό των γυναικών. Υπενθυμίζουμε επίσης ότι εάν αυτή η πιθανότητα είναι μικρή τότε μπορούμε να θεωρήσουμε το ΛΣΠ ως καλή προσέγγιση του σχετικού κινδύνου.

3.5.2 Επίδραση Καπνίσματος στην Επιβίωση σε Ασθενείς με ένα επεισόδιο Καρδιακής ανακοπής (Προοπτική Μελέτη)

Σε αυτή την ενότητα θα χρησιμοποιήσουμε το Παράδειγμα των Daly *et. al* (1991, σελ. 185) για να δούμε πως θα συγκρίνουμε τα μέτρα κινδύνου σε μια προοπτική μελέτη.

Παράδειγμα 3.2 Σε μια προοπτική μελέτη κατά την οποία εξετάσθηκαν 368 άνδρες καπνιστές ηλικίας κάτω των 60 ετών οι οποίοι έπαθαν μια καρδιακή ανακοπή και επιβίωσαν. Μετά από 2 έτη εξετάσθηκαν πόσοι από αυτούς είχαν επιβιώσει και τους χωρίσαμε ανάλογα εάν είχαν κόψει το τσιγάρο ή όχι. Έτσι εδώ μας ενδιαφέρει να εξετάσουμε αν το σταμάτημα του καπνίσματος (X) είχε ευνοϊκή επίδραση στην επιβίωση μετά από δύο έτη (Y). Τα δεδομένα δίνονται στον 2 × 2 Πίνακα που ακολουθεί

X: Συνέχισαν το κάπνισμα ;	Y: Επιβίωση σε 2 χρόνια		Περ.Κατ. X
	1: Πεθαμένος	2: Ζωντανός	
1: Ναι	19 (12.3%)	135 (87.7%)	154
2: Όχι	15 (7.0%)	199 (93.0%)	214
Περ.Κατ. Y	34 (9.2%)	334 (90.8%)	368

Πίνακας 3.2: Δεδομένα Παραδείγματος 3.2: Κάπνισμα και Επιβίωση σε Ασθενείς Καρδιακής Ανεπάρκειας (Daly *et. al* , 1983).

Υπολογίζοντας τον αποδοτέο κίνδυνο βρίσκουμε ότι είναι ίσος με 5.3%. Συνεπώς οι ασθενείς που συνέχισαν να καπνίζουν μετά το καρδιακό επεισόδιο παρουσιάζουν αυξημένη πιθανότητα θανάτου, σε σχέση με αυτούς που διέκοψαν το κάπνισμα, κατά 5.3 ποσοστιαίες μονάδες. Αυτή η διαφορά αποδίδεται στο σταμάτημα του καπνίσματος. Το πλεονέκτημα του αποδοτέου κινδύνου είναι ότι δίνει τη διαφορά σε ποσοστιαίες μονάδες κάτι που δε φαίνεται στο σχετικό κίνδυνο που θα δούμε παρακάτω.

Ο σχετικός κίνδυνος εδώ είναι ίσος με $RR = 0.123/0.07 = 1.757$. Αυτό σημαίνει ότι ο κίνδυνος θανάτου για τους ασθενείς που συνέχισαν το κάπνισμα είναι ίσος με 1.76 φορές (ή 76% μεγαλύτερος από) τον ίδιο κίνδυνο των ασθενών που σταμάτησαν το κάπνισμα.

Ο ποσοστιαίος αποδοτέος κίνδυνος είναι ίσος με $0.053/0.123 = 43.1\%$. Αυτό σημαίνει ότι το 43% του κινδύνου θανάτου που διατρέχει ένας ασθενής που συνέχισε το κάπνισμα μπορεί να αποδοθεί στο γεγονός ότι δε σταμάτησε το κάπνισμα. Συνεπώς, το 43% των θανάτων (2 χρόνια μετά το πρώτο καρδιακό επεισόδιο) των ασθενών που συνεχίσαν να καπνίζουν θα μπορούσε να είχε αποφευχθεί αν είχαμε μπορέσει να τους πείσουμε να κόψουν το κάπνισμα (δηλαδή περίπου 8 άτομα, $19 \times 0.431 = 8.1$).

Τέλος εδώ ο ΛΣΠ είναι ίσος με $OR = (19 \times 199)/(15 \times 135) = 1.867$. Δηλαδή η σχετική πιθανότητα θανάτου για τους ασθενείς που συνέχισαν το κάπνισμα είναι 86% μεγαλύτερη από την πιθανότητα θανάτου των ασθενών που σταμάτησαν το κάπνισμα. Βλέπουμε εδώ, ότι παρόλο

που η πιθανότητα θανάτου είναι λίγο μεγαλύτερη του 10% για τη μία ομάδα, ο ΛΣΠ πλησιάζει αρκετά το σχετικό κίνδυνο.

3.5.3 Παράδειγμα 3: Άσκηση για το Σπίτι.

Παράδειγμα 3.3 Πριν μερικά χρόνια, υπήρχαν υποψίες στην Ιατρική βιβλιογραφία ότι ο καρκίνος του στήθους στις γυναίκες σχετίζεται με την ηλικία της πρώτης γέννας. Η υπόθεση αυτή αναφέρει ότι ο κίνδυνος καρκίνου τους στήθους αυξάνεται όσο μεγαλώνει η ηλικία που γυναίκα γεννάει το πρώτο της παιδί. Άρα, με βάση τη θεωρία αυτή, ένας σημαντικός παράγοντας κινδύνου για τον καρκίνο του στήθους είναι η ηλικία της γυναίκας στην πρώτη γέννα. Αυτή η θεωρία εξηγεί και γιατί ο καρκίνος του στήθους είναι σημαντικά αυξημένος σε ομάδες με αυξημένο κοινωνικό - οικονομικό και εκπαιδευτικό επίπεδο αφού ζευγάρια που ανήκουν σε αυτές τις ομάδες γεννούν συνειδητά σε μεγάλες ηλικίες. Το 1970 έγινε μια διεθνής πολυκεντρική μελέτη (MacKahon *et. al* , 1970) με τη συμμετοχή των ΗΠΑ, Ελλάδα, Γιουγκοσλαβίας, Βραζιλίας, Ταϊβάν, και Ιαπωνίας. Περιπτώσεις καρκινοπαθών επιλέχθηκαν από συγκεκριμένα νοσοκομεία στις παραπάνω χώρες. Οι μάρτυρες επιλέχθηκαν από τα ίδια νοσοκομεία έτσι ώστε να είναι συγκρίσιμης ηλικίας και να μην έχουν καρκίνο μαστού ή άλλα σχετικά νοσήματα. Όλες γυναίκες χωρίστηκαν σε δύο ομάδες κινδύνου ανάλογα αν η ηλικία της πρώτης γέννας ήταν μικρότερη ή μεγαλύτερη των 30 ετών ($E : X \geq 30$ και $\bar{E} : X \leq 29$). Λεπτομέρειες σχετικές με το παράδειγμα μπορείτε να βρείτε και στο βιβλίο του Rosner (1994 σελ. 346). Τα δεδομένα δίνονται στον 2×2 Πίνακα που ακολουθεί

	Ηλικία στην πρώτη γέννα		Περ.Κατ. X
	(1) ≥ 30	(2) ≤ 29	
1: Ασθενής	683	2537	3220
2: Μάρτυρας	1498	8747	10245
Περ.Κατ. Y	2181	11284	13465

Πίνακας 3.3: Δεδομένα Παραδείγματος 3.3: Καρκίνος του στήθους και ηλικία πρώτης γέννας (MacKahon *et. al* , 1970).

Να υπολογίσετε τα κατάλληλα μέτρα κινδύνου. Ποια ταιριάζουν και γιατί;

3.6 Συμπερασματολογία και Έλεγχοι Υπόθεσης για 2×2 Πίνακες Συνάφειας

3.6.1 Ισότητα Δύο Διωνυμικών Ποσοστών

3.6.1.1 Αποδιδόμενος κίνδυνος

Γενική Προσέγγιση: Ας υποθέσουμε ότι έχουμε δύο ομάδες με διαφορετική έκθεση στον κίνδυνο E και \bar{E} και ότι η μεταβλητή απόκρισης Y είναι επίσης δίτιμη: έχει ή όχι τη νόσο (A και \bar{A}). Σε αυτή την περίπτωση μας ενδιαφέρει να ελέγξουμε κατά πόσο η πιθανότητα εμφάνισης της νόσου είναι ίδια στις δύο ομάδες. Συνεπώς

$$\left. \begin{array}{l} H_0 : P(A|E) = P(A|\bar{E}) \\ \text{έναντι της εναλλακτικής} \\ H_1 : P(A|E) \neq P(A|\bar{E}) \end{array} \right\} \Leftrightarrow \left. \begin{array}{l} H_0 : P(A|E) - P(A|\bar{E}) = 0 \\ H_1 : P(A|E) - P(A|\bar{E}) \neq 0 \end{array} \right\} \Leftrightarrow \left. \begin{array}{l} H_0 : \pi_E = \pi_{\bar{E}} \\ \text{έναντι της εναλλακτικής} \\ H_1 : \pi_E \neq \pi_{\bar{E}} \end{array} \right\} \Leftrightarrow \left. \begin{array}{l} H_0 : \pi_E - \pi_{\bar{E}} = 0 \\ H_1 : \pi_E - \pi_{\bar{E}} \neq 0 \end{array} \right\} \Leftrightarrow \begin{array}{l} H_0 : AR = 0 \\ H_1 : AR \neq 0 \end{array} .$$

Συνεπώς μπορούμε να υποθέσουμε ότι ο αριθμός των ατόμων που έχουν τη νόσο σε κάθε ομάδα έκθεσης ακολουθεί διωνυμική κατανομή οπότε

$$Y|E \sim \text{Bin}(\pi_E, n_E) \text{ και } Y|\bar{E} \sim \text{Bin}(\pi_{\bar{E}}, n_{\bar{E}}).$$

Από τα παραπάνω συνεπάγεται ότι μπορούμε να χρησιμοποιήσουμε τη διωνυμική κατανομή για να ελέγξουμε κατά πόσο ο αποδιδόμενος κίνδυνος είναι μηδέν ή όχι.

Επιπλέον μπορούμε να χρησιμοποιήσουμε το κεντρικό οριακό θεώρημα και να φτιάξουμε ένα z -test. Αν οι αναμενόμενες τιμές $n_E \pi_E$, $n_E(1 - \pi_E)$, $n_{\bar{E}} \pi_{\bar{E}}$ και $n_{\bar{E}}(1 - \pi_{\bar{E}})$ είναι μεγαλύτερες του πέντε τότε μπορούμε να υποθέσουμε ότι

$$p_E \sim N\left(\pi_E, \frac{\pi_E(1 - \pi_E)}{n_E}\right) \text{ και } p_{\bar{E}} \sim N\left(\pi_{\bar{E}}, \frac{\pi_{\bar{E}}(1 - \pi_{\bar{E}})}{n_{\bar{E}}}\right)$$

όπου p_E και $p_{\bar{E}}$ είναι οι δειγματικές αναλογίες εμφάνισης της νόσου για τις ομάδες με έκθεση ή όχι στον παράγοντα κινδύνου. Αν τώρα πάρουμε τις διαφορές θα έχουμε

$$\widehat{AR} = p_E - p_{\bar{E}} \sim N\left(\pi_E - \pi_{\bar{E}}, \frac{\pi_E(1 - \pi_E)}{n_E} + \frac{\pi_{\bar{E}}(1 - \pi_{\bar{E}})}{n_{\bar{E}}}\right).$$

Άρα μπορούμε να χρησιμοποιήσουμε για τον παραπάνω έλεγχο τη συνάρτηση ελέγχου:

$$z = \frac{\widehat{AR} - AR}{\sqrt{\frac{\pi_E(1 - \pi_E)}{n_E} + \frac{\pi_{\bar{E}}(1 - \pi_{\bar{E}})}{n_{\bar{E}}}}} \sim N(0, 1).$$

Αν ισχύει η H_0 τότε $AR = \pi_E - \pi_{\bar{E}} = 0$ και $\pi_E = \pi_{\bar{E}} = \pi$ συνεπώς η συνάρτηση ελέγχου γίνεται

$$z = \frac{\widehat{AR}}{\sqrt{\pi(1-\pi)\left(\frac{1}{n_E} + \frac{1}{n_{\bar{E}}}\right)}} \sim N(0, 1).$$

Αντικαθιστούμε το κοινό π με την αντίστοιχη εκτίμηση του $p = (n_E p_E + n_{\bar{E}} p_{\bar{E}})/(n_E + n_{\bar{E}})$ οπότε

$$z = \frac{\widehat{AR}}{\sqrt{p(1-p)\left(\frac{1}{n_E} + \frac{1}{n_{\bar{E}}}\right)}} = \frac{p_E - p_{\bar{E}}}{\sqrt{p(1-p)\left(\frac{1}{n_E} + \frac{1}{n_{\bar{E}}}\right)}} \sim N(0, 1).$$

Εφόσον ο αποδιδόμενος κίνδυνος μπορεί να εκτιμηθεί μόνο στις προοπτικές μελέτες, μόνο εκεί μπορούμε να εφαρμόσουμε την παραπάνω προσέγγιση. Στις αναδρομικές μελέτες, η παραπάνω σύγκριση μπορεί να γίνει ανάποδα για να συγκρίνουμε την έκθεση στον κίνδυνο για τους μάρτυρες και τους ασθενείς, δηλαδή να ελέγξουμε την υπόθεση

$$H_0 : P(E|A) = P(E|\bar{A}) \text{ έναντι της εναλλακτικής } H_1 : P(E|A) \neq P(E|\bar{A}).$$

Από τα παραπάνω επίσης προκύπτει ότι μπορούμε να φτιάξουμε $100(1 - \alpha)\%$ διαστήματα εμπιστοσύνης για τον αποδιδόμενο κίνδυνο το οποίο θα δίνεται από τον τύπο

$$\widehat{AR} \pm z_{1-\alpha/2} \sqrt{\frac{p_E(1-p_E)}{n_E} + \frac{p_{\bar{E}}(1-p_{\bar{E}})}{n_{\bar{E}}}}.$$

Σύμφωνα με τον Rosner (1994, σελ. 363) η παραπάνω προσέγγιση είναι ικανοποιητική μόνο αν $n_E p_E(1-p_E) \geq 5$ και $n_{\bar{E}} p_{\bar{E}}(1-p_{\bar{E}}) \geq 5$.

Αποδιδόμενος κίνδυνος σε πίνακες 2×2 : Στους πίνακες 2×2 όπου η σειρά των ενδεχομένων για τις δύο μεταβλητές είναι (E, \bar{E}) και (A, \bar{A}) ο αποδιδόμενος κίνδυνος δίνεται από τον τύπο

$$AR = \pi_E - \pi_{\bar{E}} = \pi_{1|1} - \pi_{1|2}$$

και εκτιμάται

$$\begin{aligned} \widehat{AR} &= p_E - p_{\bar{E}} = p_{1|1} - p_{1|2} = \frac{n_{11}}{n_{1\bullet}} - \frac{n_{21}}{n_{2\bullet}} = \frac{n_{11}}{n_{11} + n_{12}} - \frac{n_{21}}{n_{21} + n_{22}} \\ &= \frac{n_{11}n_{22} + n_{11}n_{22} - n_{21}n_{11} - n_{21}n_{12}}{(n_{11} + n_{12})(n_{21} + n_{22})} = \frac{n_{11}n_{22} - n_{21}n_{12}}{(n_{11} + n_{12})(n_{21} + n_{22})} \\ &= \frac{n_{11}n_{22} - n_{21}n_{12}}{n_{1\bullet}n_{2\bullet}} \end{aligned}$$

Το κοινό ποσοστό εκτιμάται ίσο με

$$p = \frac{n_{\bullet 1}}{n} = \frac{n_{11} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}}.$$

ΣΥΝΑΡΤΗΣΗ ΕΛΕΓΧΟΥ AR

$H_0 : \pi_E = \pi_{\bar{E}}$ έναντι της εναλλακτικής $H_1 : \pi_E \neq \pi_{\bar{E}} \Leftrightarrow$

$H_0 : AR = 0$ έναντι της εναλλακτικής $H_1 : AR \neq 0$

$$z_{AR} = \frac{p_E - p_{\bar{E}}}{\sqrt{p(1-p) \left(\frac{1}{n_E} + \frac{1}{n_{\bar{E}}} \right)}}$$

$$p = \frac{n_E p_E + n_{\bar{E}} p_{\bar{E}}}{n_E + n_{\bar{E}}}$$

Απορρίπτουμε την H_0 όταν $|z_{AR}| < Z_{\alpha/2}$.

ΔΙΑΣΤΗΜΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

Από τα παραπάνω επίσης προκύπτει ότι μπορούμε να φτιάξουμε $100(1 - \alpha)\%$ διαστήματα εμπιστοσύνης για τον αποδιδόμενο κίνδυνο το οποίο θα δίνεται από τον τύπο

$$\widehat{AR} \pm z_{1-\alpha/2} \sqrt{\frac{p_E(1-p_E)}{n_E} + \frac{p_{\bar{E}}(1-p_{\bar{E}})}{n_{\bar{E}}}}$$

Πίνακας 3.4: Συνοπτικός Πίνακας Διαστήματος εμπιστοσύνης και ελέγχου υπόθεσης Αποδιδόμενου κινδύνου.

Επιπλέον το τυπικό σφάλμα του αποδιδόμενου κινδύνου, εάν η H_0 ισχύει, θα είναι ίση με

$$\begin{aligned} se(\widehat{AR}) &= \sqrt{p(1-p) \left(\frac{1}{n_{1\bullet}} + \frac{1}{n_{2\bullet}} \right)} = \sqrt{\frac{n_{\bullet 1} n_{\bullet 2}}{n^2} \left(\frac{1}{n_{1\bullet}} + \frac{1}{n_{2\bullet}} \right)} \\ &= \sqrt{\frac{n_{\bullet 1} n_{\bullet 2}}{n^2} \left(\frac{n_{1\bullet} + n_{2\bullet}}{n_{1\bullet} n_{2\bullet}} \right)} = \sqrt{\frac{n_{\bullet 1} n_{\bullet 2}}{n n_{1\bullet} n_{2\bullet}}}. \end{aligned}$$

Συνεπώς η συνάρτηση ελέγχου δίνεται από τον τύπο

$$\begin{aligned} Z &= \frac{\frac{n_{11}}{n_{11}+n_{12}} - \frac{n_{21}}{n_{21}+n_{22}}}{\sqrt{\frac{n_{\bullet 1} n_{\bullet 2}}{n n_{1\bullet} n_{2\bullet}}}} = \frac{\frac{n_{11}n_{22} - n_{21}n_{12}}{n_{1\bullet}n_{2\bullet}}}{\sqrt{\frac{n_{\bullet 1} n_{\bullet 2}}{n n_{1\bullet} n_{2\bullet}}}} \\ &= \frac{(n_{11}n_{22} - n_{21}n_{12})\sqrt{n}}{\sqrt{n_{\bullet 1}n_{\bullet 2}n_{1\bullet}n_{2\bullet}}}. \end{aligned}$$

Οι παραπάνω τύποι δε θα αλλάξουν αν χρησιμοποιήσουμε αντίστροφα (όπως είναι και ο συνηθισμένος τρόπος) τα ενδεχόμενα των δύο μεταβλητών δηλαδή έχουμε $X = (\bar{E}, E)$ και $Y = (\bar{A}, A)$.

3.6.1.2 Σχετικός κίνδυνος.

Γενική Προσέγγιση: Όπως είδαμε ο σχετικός κίνδυνος ορίζεται ως ο λόγος των πιθανοτήτων εμφάνισης κινδύνου για τις δύο ομάδες διαφορετικής έκθεσης στον κίνδυνο δηλαδή $RR = \pi_E/\pi_{\bar{E}}$. Σε αυτή την ενότητα θα ασχοληθούμε με την κατανομή δειγματοληψίας του δειγματικού σχετικού κινδύνου $\widehat{RR} = p_E/p_{\bar{E}}$ την οποία μπορούμε να τη χρησιμοποιήσουμε για την κατασκευή διαστημάτων εμπιστοσύνης και ελέγχων υποθέσεων. Για να είμαστε πιο ακριβείς θα ασχοληθούμε με την κατανομή δειγματοληψίας του λογαρίθμου του σχετικού κινδύνου (γενικά όταν έχουμε λόγους εκτιμητριών θα δουλεύουμε με λογαρίθμους). Έτσι έχουμε

$$\log \widehat{RR} = \log p_E - \log p_{\bar{E}}$$

γνωρίζοντας ότι

$$n_E p_E \sim \text{Bin}(\pi_E, n_E)$$

$$n_{\bar{E}} p_{\bar{E}} \sim \text{Bin}(\pi_{\bar{E}}, n_{\bar{E}}),$$

όπου $\text{Bin}(p, n)$ είναι η διωνυμική κατανομή με πιθανότητα επιτυχίας p και αριθμό επαναλήψεων n . Από τα παραπάνω μπορούμε να υπολογίσουμε τη μέση τιμή και τη διακύμανση των εκτιμητριών p_E και $p_{\bar{E}}$ οι οποίες δίνονται από τις ποσότητες

$$E(p_E) = \pi_E \text{ και } \text{Var}(p_E) = \frac{\pi_E(1 - \pi_E)}{n_E}$$

$$E(p_{\bar{E}}) = \pi_{\bar{E}} \text{ και } \text{Var}(p_{\bar{E}}) = \frac{\pi_{\bar{E}}(1 - \pi_{\bar{E}})}{n_{\bar{E}}}.$$

Για αρκετά μεγάλο δείγμα (πιο συγκεκριμένα για $n_E p_E \geq 5$ και $n_E(1 - p_E) \geq 5$) μπορούμε να θεωρήσουμε ότι η κατανομή του p_E προσεγγίζεται ικανοποιητικά από την κανονική κατανομή με μέσο και διακύμανση που δίνονται από την παραπάνω εξίσωση. Ακριβώς το ίδιο ισχύει και για την ομάδα μη εκτεθειμένων δηλαδή πρέπει να ισχύει $n_{\bar{E}} p_{\bar{E}} \geq 5$ και $n_{\bar{E}}(1 - p_{\bar{E}}) \geq 5$. Έτσι έχουμε

$$p_E \sim N\left(\pi_E, \frac{\pi_E(1 - \pi_E)}{n_E}\right) \text{ και } p_{\bar{E}} \sim N\left(\pi_{\bar{E}}, \frac{\pi_{\bar{E}}(1 - \pi_{\bar{E}})}{n_{\bar{E}}}\right).$$

Για να προχωρήσουμε τώρα θα χρειαστούμε τη σειρά Taylor η οποία μας λέει ότι κάθε συνάρτηση $h(x)$ μπορεί να γραφτεί ως

$$h(x) = \sum_{k=0}^{\infty} h^{(k)}(a) \frac{(x-a)^k}{k!} = h(a) + h'(a)(x-a) + h''(a) \frac{(x-a)^2}{2} + h'''(a) \frac{(x-a)^3}{6} + \dots$$

όπου $h^{(k)}(x)$ είναι η παράγωγος k τάξης της συνάρτησης $h(x)$. Έστω λοιπόν ότι X είναι τυχαία μεταβλητή και $a = E(X) = \mu$ τότε μπορούμε να χρησιμοποιήσουμε την παραπάνω έκφραση

για να υπολογίσουμε ασυμπτωτικά τη μέση τιμή και τη διακύμανση μιας συνάρτησης της X . Συνεπώς παίρνοντας την προσέγγιση που βασίζεται στους τρεις πρώτους όρους έχουμε

$$\begin{aligned} E(h(X)) &= \sum_{k=0}^{\infty} h^{(k)}(\mu) \frac{E(X - \mu)^k}{k!} \\ &\approx h(\mu) + h'(\mu)E(X - \mu) + h''(\mu) \frac{E(X - \mu)^2}{2} \\ &\approx h(\mu) + h''(\mu) \frac{V(X)}{2}. \end{aligned} \quad (3.4)$$

Όμοια μπορούμε να υπολογίσουμε τη διακύμανση βασιζόμενοι στους δύο πρώτους όρους της σειράς. Έτσι έχουμε

$$\begin{aligned} V(h(X)) &= V\left(\sum_{k=0}^{\infty} h^{(k)}(\mu) \frac{(X - \mu)^k}{k!}\right) \\ &\approx V(h(\mu) + h'(\mu)(X - \mu)) \\ &\approx \{h'(\mu)\}^2 V(X). \end{aligned} \quad (3.5)$$

Γυρνώντας στην περίπτωση του σχετικού κινδύνου, θέτοντας $X = p_E$, $h(X) = \log(p_E)$, $E(X) = \pi_E$ και $V(X) = \pi_E(1 - \pi_E)/n_E$ στον τύπο (3.4) έχουμε

$$E(\log p_E) \approx \log \pi_E - \frac{1}{\pi_E^2} \frac{\pi_E(1 - \pi_E)}{2n_E} = \log \pi_E - \frac{1}{2n_E} \frac{(1 - \pi_E)}{\pi_E}$$

το οποίο για μεγάλο n_E γίνεται ίσο με $\log \pi_E$. Συνεπώς το $\log p_E$ είναι ασυμπτωτικά αμερόληπτος εκτιμητής του $\log \pi_E$. Όμοια από τον τύπο (3.5) προκύπτει ότι

$$V(\log p_E) \approx \frac{1}{\pi_E^2} \frac{\pi_E(1 - \pi_E)}{n_E} = \frac{1}{n_E} \frac{(1 - \pi_E)}{\pi_E}.$$

Χρησιμοποιώντας τους δύο πρώτους όρους της σειρά Taylor μπορούμε να γράψουμε

$$\log p_E = \log \pi_E + (p_E - \pi_E)(\log \pi_E)' + O$$

όπου

$$O = \sum_{k=2}^{\infty} (\log \pi_E)^{(k)} \frac{(p_E - \pi_E)^k}{k!} = \sum_{k=2}^{\infty} (-1)^{k-1} k! \pi_E^{-k} \frac{(p_E - \pi_E)^k}{k!} = \sum_{k=2}^{\infty} (-1)^{k-1} \left(\frac{p_E - \pi_E}{\pi_E}\right)^k.$$

Συνεπώς μπορούμε να γράψουμε

$$\sqrt{n}(\log p_E - \log \pi_E) = \sqrt{n_E} \frac{p_E - \pi_E}{\pi_E} + \sqrt{n_E} O$$

Για $n \rightarrow \infty$ προκύπτει ότι

$$\sqrt{n_E}(\log p_E - \log \pi_E) \rightarrow \sqrt{n_E} \frac{p_E - \pi_E}{\pi_E} \stackrel{\text{ασυμπτ.}}{\sim} N\left(0, \frac{1 - \pi_E}{\pi_E}\right)$$

και συνεπώς

$$\log p_E \stackrel{\text{ασυμπτ.}}{\sim} N\left(\log \pi_E, \frac{1}{n_E} \frac{(1 - \pi_E)}{\pi_E}\right).$$

Με ακριβώς τα ίδια επιχειρήματα βρίσκουμε και την ασυμπτωτική κατανομή του $\log p_{\bar{E}}$ η οποία δίνεται ως

$$\log p_{\bar{E}} \stackrel{\text{ασυμπτ.}}{\sim} N\left(\log \pi_{\bar{E}}, \frac{1}{n_{\bar{E}}} \frac{(1 - \pi_{\bar{E}})}{\pi_{\bar{E}}}\right).$$

Από τις παραπάνω εξισώσεις προκύπτει και η κατανομή δειγματοληψίας του λογαρίθμου (υποθέτοντας ανεξαρτησία των δύο ομάδων έκθεσης στον κίνδυνο) η οποία θα δίνεται από τον τύπο

$$\log \widehat{RR} = \log p_e - \log p_{\bar{E}} \stackrel{\text{ασυμπτ.}}{\sim} N\left(\log RR, \frac{1}{n_E} \frac{(1 - \pi_E)}{\pi_E} + \frac{1}{n_{\bar{E}}} \frac{(1 - \pi_{\bar{E}})}{\pi_{\bar{E}}}\right).$$

Στην πράξη εκτιμούμε το τυπικό σφάλμα χρησιμοποιώντας τις δειγματικές αναλογίες p_E και $p_{\bar{E}}$ από τον τύπο

$$\begin{aligned} V(\log \widehat{RR}) &= \frac{1}{n_E} \frac{(1 - p_E)}{p_E} + \frac{1}{n_{\bar{E}}} \frac{(1 - p_{\bar{E}})}{p_{\bar{E}}} \\ &= \frac{1}{n_E} \frac{n_E - r_E}{r_E} + \frac{1}{n_{\bar{E}}} \frac{(n_{\bar{E}} - r_{\bar{E}})}{r_{\bar{E}}} \\ &= \frac{1}{r_E} - \frac{1}{n_E} + \frac{1}{r_{\bar{E}}} - \frac{1}{n_{\bar{E}}}. \end{aligned}$$

όπου r_E και $r_{\bar{E}}$ είναι ο παρατηρούμενος αριθμός των ασθενών στις ομάδες με ή χωρίς έκθεση στον κίνδυνο αντίστοιχα. Ένα $100(1 - \alpha)\%$ διάστημα εμπιστοσύνης για το λογάριθμο του σχετικού κινδύνου θα δίνεται από τις τιμές $\log \widehat{RR} \pm z_{1-\alpha/2} \sqrt{V(\log \widehat{RR})}$ δηλαδή

$$\log \widehat{RR} \pm z_{1-\alpha/2} \sqrt{\frac{1}{r_E} - \frac{1}{n_E} + \frac{1}{r_{\bar{E}}} - \frac{1}{n_{\bar{E}}}}$$

όπου z_q είναι το $100q$ ποσοστημόριο της τυποποιημένης κανονικής κατανομής. Αντίστοιχα για το σχετικό κίνδυνο το αντίστοιχο διάστημα εμπιστοσύνης προκύπτει από τις τιμές

$$e^{\log \widehat{RR} \pm z_{1-\alpha/2} \sqrt{\frac{1}{r_E} - \frac{1}{n_E} + \frac{1}{r_{\bar{E}}} - \frac{1}{n_{\bar{E}}}}}$$

όπου $e \approx 2.71$ είναι η βάση των φυσικού λογαρίθμου. Σύμφωνα με τον Rosner (1994, σελ. 364) η παραπάνω προσέγγιση είναι ικανοποιητική μόνο αν $n_E p_E (1 - p_E) \geq 5$ και $n_{\bar{E}} p_{\bar{E}} (1 - p_{\bar{E}}) \geq 5$.

Αν θέλουμε να ελέγξουμε την υπόθεση της ανεξαρτησίας μεταξύ νόσου και παράγοντα κινδύνου τότε μπορούμε να κάνουμε τον έλεγχο:

$$H_0 : RR = 1 \text{ έναντι της εναλλακτικής } H_1 : RR \neq 1.$$

Χρησιμοποιούμε τη συνάρτηση ελέγχου

$$Z = \frac{\log \widehat{RR}}{\sqrt{\frac{1}{r_E} - \frac{1}{n_E} + \frac{1}{r_{\bar{E}}} - \frac{1}{n_{\bar{E}}}}} \sim N(0, 1).$$

Αν $|Z| < Z_{1-a/2}$ τότε δεν απορρίπτουμε την υπόθεση της ανεξαρτησίας σε επίπεδο σημαντικότητας $100a\%$ αλλιώς απορρίπτουμε την H_0 .

Επιπλέον πολλές φορές μας ενδιαφέρει και αν ένας παράγοντας είναι επιβαρυντικός ή προστατευτικός. Σε αυτές τις περιπτώσεις χρησιμοποιούμε την ίδια μηδενική υπόθεση αλλά διαφορετική εναλλακτική: $H_1 : RR > 1$ (επιβαρυντικός παράγοντας) ή $H_1 : RR < 1$ (προστατευτικός παράγοντας). Στην ουσία εδώ κάνουμε ελέγχους μίας ουράς (one sided tests) όπου απορρίπτουμε την H_0 στην περίπτωση όπου $Z < Z_a$ όταν ελέγχουμε αν ο παράγοντας είναι προστατευτικός ($H_1 : RR < 1$) και στην περίπτωση όπου $Z > Z_{1-a}$ όταν ελέγχουμε αν ο παράγοντας είναι επιβαρυντικός ($H_1 : RR > 1$).

Σχετικός κίνδυνος σε 2×2 Πίνακες: Σε 2×2 πίνακες το τυπικό σφάλμα του $\log \widehat{RR}$ είναι ίσο με

$$se(\log \widehat{RR}) = \frac{1}{n_{11}} - \frac{1}{n_{1.}} + \frac{1}{n_{21}} - \frac{1}{n_{2.}}.$$

3.6.1.3 Λόγος Σχετικών Πιθανοτήτων.

Γενική Προσέγγιση: Ο λόγος σχετικών πιθανοτήτων (ΛΣΠ ή Odds Ratio ή OR) δίνεται από τον τύπο

$$OR = \frac{Odds_E}{Odds_{\bar{E}}} = \frac{\pi_E / (1 - \pi_E)}{\pi_{\bar{E}} / (1 - \pi_{\bar{E}})} = \frac{\pi_E (1 - \pi_{\bar{E}})}{\pi_{\bar{E}} (1 - \pi_E)}$$

και εκτιμάται από το

$$\widehat{OR} = \frac{p_E / (1 - p_E)}{p_{\bar{E}} / (1 - p_{\bar{E}})} = \frac{p_E (1 - p_{\bar{E}})}{p_{\bar{E}} (1 - p_E)}.$$

Όπως και στις προηγούμενες περιπτώσεις δουλεύουμε με το λογάριθμο δηλαδή με

$$\log \widehat{OR} = \log \frac{p_E}{1 - p_E} - \log \frac{p_{\bar{E}}}{1 - p_{\bar{E}}}.$$

Με βάση την διωνυμική κατανομή βρίσκουμε ότι

$$E(p_E) = \pi_E \text{ και } V(p_E) = \pi_E (1 - \pi_E) / n_E.$$

Από τον ασυμπτωτικό τύπο (3.4) έχουμε

$$\begin{aligned} E \left(\log \frac{p_E}{1 - p_E} \right) &\approx \log \frac{\pi_E}{1 - \pi_E} + \left(-\pi_E^{-2} + (1 - \pi_E)^{-2} \right) \frac{\pi_E (1 - \pi_E)}{n_E} \\ &\approx \log Odds_E + \frac{1}{n_E} \left(Odds_E - \frac{1}{Odds_E} \right) \end{aligned}$$

ΔΙΑΣΤΗΜΑ ΕΜΠΙΣΤΟΣΥΝΗΣ RR

Ένα $100(1 - \alpha)\%$ διάστημα εμπιστοσύνης για το λογάριθμο του σχετικού κινδύνου θα δίνεται από τις τιμές $\log \widehat{RR} \pm z_{1-\alpha/2} \sqrt{V(\log \widehat{RR})}$ δηλαδή

$$\log \widehat{RR} \pm z_{1-\alpha/2} \sqrt{\frac{1}{r_E} - \frac{1}{n_E} + \frac{1}{r_{\bar{E}}} - \frac{1}{n_{\bar{E}}}}$$

όπου z_q είναι το $100q$ ποσοστημόριο της τυποποιημένης κανονικής κατανομής. Αντίστοιχα για το σχετικό κίνδυνο το αντίστοιχο διάστημα εμπιστοσύνης προκύπτει από τις τιμές

$$e^{\log \widehat{RR} \pm z_{1-\alpha/2} \sqrt{\frac{1}{r_E} - \frac{1}{n_E} + \frac{1}{r_{\bar{E}}} - \frac{1}{n_{\bar{E}}}}}$$

όπου $e \approx 2.71$ είναι η βάση των φυσικού λογαρίθμου.

Η παραπάνω προσέγγιση είναι ικανοποιητική μόνο αν $n_E p_E (1 - p_E) \geq 5$ και $n_{\bar{E}} p_{\bar{E}} (1 - p_{\bar{E}}) \geq 5$.

ΣΥΝΑΡΤΗΣΗ ΕΛΕΓΧΟΥ RR

Αν θέλουμε να ελέγξουμε την υπόθεση της ανεξαρτησίας μεταξύ νόσου και παράγοντα κινδύνου τότε μπορούμε να κάνουμε τον έλεγχο:

$$H_0 : \log RR = 0 \text{ έναντι της εναλλακτικής } H_1 : \log RR \neq 0.$$

Χρησιμοποιούμε τη συνάρτηση ελέγχου

$$Z = \frac{\log \widehat{RR}}{\sqrt{\frac{1}{r_E} - \frac{1}{n_E} + \frac{1}{r_{\bar{E}}} - \frac{1}{n_{\bar{E}}}}} \sim N(0, 1).$$

Αν $|Z| < Z_{1-\alpha/2}$ τότε δεν απορρίπτουμε την υπόθεση της ανεξαρτησίας σε επίπεδο σημαντικότητας $100\alpha\%$ αλλιώς απορρίπτουμε την H_0 .

Πίνακας 3.5: Συνοπτικός Πίνακας Διαστήματος εμπιστοσύνης και ελέγχου υπόθεσης Σχετικού κινδύνου.

εφόσον $h(x) = \log x - \log(1-x)$, $h'(x) = x^{-1} + (1-x)^{-1} = 1/[x(1-x)]$ και $h''(x) = -x^{-2} + (1-x)^{-2}$. Επιπλέον για n_E μεγάλο έχουμε $E\left(\log \frac{p_E}{1-p_E}\right) \approx \log Odds_E$. Όσον αφορά τη διακύμανση έχουμε

$$\begin{aligned} V\left(\log \frac{p_E}{1-p_E}\right) &\approx \{h'(\mu)\}^2 V(p_E) \\ &\approx \left(\frac{1}{\pi_E(1-\pi_E)}\right)^2 \frac{\pi_E(1-\pi_E)}{n_E} \\ &\approx \frac{1}{n_E} \frac{1}{\pi_E(1-\pi_E)} \\ &\approx \left(n_E \frac{r_E}{n_E} \frac{n_E - r_E}{n_E}\right)^{-1} = \left(\frac{r_E(n_E - r_E)}{n_E}\right)^{-1} \\ &\approx \left(\frac{n_E}{r_E(n_E - r_E)}\right) = \left(\frac{n_E - r_E + r_E}{r_E(n_E - r_E)}\right) \\ &\approx \frac{1}{n_E} + \frac{1}{n_E - r_E}. \end{aligned}$$

Ακολουθώντας την ίδια λογική

$$\begin{aligned} E\left(\log \frac{p_{\bar{E}}}{1-p_{\bar{E}}}\right) &\approx \log Odds_{\bar{E}} \text{ για } n_{\bar{E}} \text{ μεγάλο,} \\ V\left(\log \frac{p_{\bar{E}}}{1-p_{\bar{E}}}\right) &= \frac{1}{r_{\bar{E}}} + \frac{1}{n_{\bar{E}} - r_{\bar{E}}}. \end{aligned}$$

Συνεπώς

$$\begin{aligned} E(\log \widehat{OR}) &= \log \frac{\pi_E}{1-\pi_E} - \log \frac{\pi_{\bar{E}}}{1-\pi_{\bar{E}}} = \log Odds_E - \log Odds_{\bar{E}} = \log OR \\ V(\log \widehat{OR}) &= \frac{1}{r_E} + \frac{1}{n_E - r_E} + \frac{1}{r_{\bar{E}}} + \frac{1}{n_{\bar{E}} - r_{\bar{E}}}. \end{aligned}$$

άρα για $(n_E, n_{\bar{E}})$ αρκετά μεγάλα έχουμε (χρησιμοποιώντας τα ίδια επιχειρήματα όπως στο PP)

$$\log \widehat{OR} \overset{\text{ασυμπτ.}}{\approx} N\left(\log OR, \frac{1}{r_E} + \frac{1}{n_E - r_E} + \frac{1}{r_{\bar{E}}} + \frac{1}{n_{\bar{E}} - r_{\bar{E}}}\right).$$

Το παραπάνω αποτέλεσμα μπορεί να προκύψει άμεσα από την κατανομή δειγματοληψίας του σχετικού κίνδυνου αν γράψουμε το ΛΣΠ ως

$$\log \widehat{OR} = \log \frac{\pi_E}{\pi_{\bar{E}}} - \log \frac{1-\pi_E}{1-\pi_{\bar{E}}} = \log \widehat{RR} - \log \widehat{RR}'.$$

Στην προηγούμενη ενότητα υπολογίσαμε την κατανομή του $\log \widehat{RR}$ ενώ η κατανομή του $\log \widehat{RR}' = \log \frac{1-\pi_E}{1-\pi_{\bar{E}}}$ προκύπτει με τον ίδιο τρόπο αφού είναι ο σχετικός κίνδυνος του συμπληρωματικού ενδεχομένου \bar{A} (ενώ ο RR εξετάζει το A).

ΔΙΑΣΤΗΜΑ ΕΜΠΙΣΤΟΣΥΝΗΣ OR

Ένα $100(1 - \alpha)\%$ διάστημα εμπιστοσύνης για το λογάριθμο του ΛΣΠ θα δίνεται από τις τιμές $\log \widehat{OR} \pm z_{1-\alpha/2} \sqrt{V(\log \widehat{OR})}$ δηλαδή

$$\log \widehat{OR} \pm z_{1-\alpha/2} \sqrt{\frac{1}{r_E} + \frac{1}{n_E - r_E} + \frac{1}{r_{\bar{E}}} + \frac{1}{n_{\bar{E}} - r_{\bar{E}}}}$$

όπου z_q είναι το $100q$ ποσοστημόριο της τυποποιημένης κανονικής κατανομής.

Αντίστοιχα για το ΛΣΠ το αντίστοιχο διάστημα εμπιστοσύνης προκύπτει από τις τιμές

$$e^{\log \widehat{OR} \pm z_{1-\alpha/2} \sqrt{\frac{1}{r_E} + \frac{1}{n_E - r_E} + \frac{1}{r_{\bar{E}}} + \frac{1}{n_{\bar{E}} - r_{\bar{E}}}}}$$

όπου $e \approx 2.71$ είναι η βάση των φυσικού λογαρίθμου.

Η παραπάνω προσέγγιση είναι ικανοποιητική μόνο αν $n_E p_E (1 - p_E) \geq 5$ και $n_{\bar{E}} p_{\bar{E}} (1 - p_{\bar{E}}) \geq 5$.

ΣΥΝΑΡΤΗΣΗ ΕΛΕΓΧΟΥ OR

Αν θέλουμε να ελέγξουμε την υπόθεση της ανεξαρτησίας μεταξύ νόσου και παράγοντα κινδύνου τότε μπορούμε να κάνουμε τον έλεγχο:

$$H_0 : \log OR = 0 \text{ έναντι της εναλλακτικής } H_1 : \log OR \neq 0.$$

Χρησιμοποιούμε τη συνάρτηση ελέγχου

$$Z = \frac{\log \widehat{OR}}{\sqrt{\frac{1}{r_E} + \frac{1}{n_E - r_E} + \frac{1}{r_{\bar{E}}} + \frac{1}{n_{\bar{E}} - r_{\bar{E}}}}} \sim N(0, 1).$$

Αν $|Z| < Z_{1-\alpha/2}$ τότε δεν απορρίπτουμε την υπόθεση της ανεξαρτησίας σε επίπεδο σημαντικότητας $100\alpha\%$ αλλιώς απορρίπτουμε την H_0 .

Πίνακας 3.6: Συνοπτικός Πίνακας Διαστήματος εμπιστοσύνης και ελέγχου υπόθεσης λόγου σχετικών πιθανοτήτων.

ΛΣΠ σε 2×2 Πίνακες: Σε 2×2 πίνακες ο παραπάνω τύπος απλοποιείται σε

$$\log \widehat{OR} \overset{\text{ασυμπτ.}}{\sim} N \left(\log OR, \frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}} \right).$$

Στην περίπτωση που ένα κελί είναι ίσο με μηδέν τότε ο ΛΣΠ θα είναι ίσος με μηδέν ή θα απειρίζεται. Αντίστοιχα δεν μπορούμε να υπολογίσουμε το τυπικό σφάλμα το οποίο επίσης απειρίζεται. Στην περίπτωση αυτή έχει επικρατήσει να προσθέτουμε την ποσότητα $\delta = 0.5$ σε όλα τα κελιά και να ξαναυπολογίζουμε τον εκτιμητή και τη διακύμανση του δηλαδή

$$\widehat{OR}_{cor} = \frac{(n_{11} + \delta)(n_{22} + \delta)}{(n_{21} + \delta)(n_{12} + \delta)}$$

$$Var(\log \widehat{OR}_{cor}) = \frac{1}{n_{11} + \delta} + \frac{1}{n_{12} + \delta} + \frac{1}{n_{21} + \delta} + \frac{1}{n_{22} + \delta}.$$

Όταν όλα τα n_{ij} είναι σχετικά μεγάλα τότε ο παραπάνω εκτιμητής διαφέρει ελάχιστα του αρχικού εκτιμητή.

3.6.1.4 Παράδειγμα 3.2 (συνέχεια)

Από τον πίνακα 3.2 του παραδείγματος 3.2 μπορούμε να υπολογίσουμε τα ακόλουθα μέτρα κινδύνου:

α) Αποδιδόμενος κίνδυνος: $\widehat{AR} = p_E - p_{\bar{E}} = 0.123 - 0.070 = 0.053$. Επιθυμούμε να δούμε αν αυτή η διαφορά θνησιμότητας είναι στατιστικά σημαντική, συνεπώς θέλουμε να ελέγξουμε την υπόθεση

$$H_0 : AR = 0 \text{ έναντι της εναλλακτικής } H_1 : AR \neq 0.$$

Αν η μηδενική υπόθεση ισχύει έχουμε

$$\begin{aligned} V(\widehat{AR}|H_0) &= p(1-p) \left(\frac{1}{n_E} + \frac{1}{n_{\bar{E}}} \right) \\ &= \frac{34}{368} \frac{334}{368} \left(\frac{1}{154} + \frac{1}{214} \right) \\ &= 0.092 \times 0.9080.1117 = 0.000936 \Leftrightarrow \\ se(\widehat{AR}|H_0) &= \sqrt{0.000936} = 0.0306. \end{aligned}$$

Άρα $z = 0.0503/0.0306 = 1.7413 < z_{0.975} = 1.96$ άρα δεν απορρίπτουμε την H_0 συνεπώς δεν υπάρχει στατιστική διαφορά της θνησιμότητας ανάλογα αν ο ασθενής συνεχίσει το κάπνισμα (προσοχή η διαφορά είναι στατιστικά σημαντική για $\alpha = 0.10$). Σημείωση: ισχύει η προϋπόθεση $n_E p_E (1 - p_E) = 154 \times 0.123(1 - 0.123) = 16.6 > 5$ και $n_{\bar{E}} p_{\bar{E}} (1 - p_{\bar{E}}) = 214 \times 0.07(1 - 0.07) = 23.96 > 5$.

Για να βρούμε το 95% διάστημα εμπιστοσύνης υπολογίζουμε πρώτα το τυπικό σφάλμα

$$\begin{aligned} se(\widehat{AR}) &= \sqrt{\frac{p_E(1-p_E)}{n_E} + \frac{p_{\bar{E}}(1-p_{\bar{E}})}{n_{\bar{E}}}} \\ &= \sqrt{\frac{0.123(1-0.123)}{154} + \frac{0.07(1-0.07)}{214}} = \sqrt{0.000700461 + 0.0003042056} = 0.0317 \end{aligned}$$

συνεπώς το 95% ΔΕ δίνεται από τον τύπο $0.053 \pm 1.96 \times 0.0317 = (-0.009, 0.115)$.

β) Σχετικός κίνδυνος:

$$\begin{aligned} \widehat{RR} &= \frac{p_E}{p_{\bar{E}}} = 0.123/0.070 = 1.76 \\ \log \widehat{RR} &= \log 1.76 = 0.5654 \\ se(\log \widehat{RR}) &= \sqrt{\frac{1}{19} - \frac{1}{154} + \frac{1}{15} - \frac{1}{214}} = \sqrt{0.108132} = 0.3288. \end{aligned}$$

Το 95% ΔΕ για το λογάριθμο του σχετικού κινδύνου δίνεται από τον τύπο $0.5654 \pm 1.96 \times 0.3288 = (-0.0791, 1.2099)$ ενώ για το σχετικό κίνδυνο είναι ίσο με $(e^{-0.0791}, e^{1.2099}) = (0.924, 3.353)$. Μπορούμε να ελέγξουμε την υπόθεση $H_0 : RR = 1$ έναντι της εναλλακτικής $H_1 : RR \neq 1$ με την ελεγχουσυνάρτηση $Z_{RR} = 0.5654/0.3288 = 1.72 < z_{0.975}$ άρα δεν απορρίπτουμε την H_0 για $\alpha = 5\%$ και συνεπώς τα ποσοστά θνησιμότητας δεν αλλάζουν για τα άτομα που συνέχισαν να καπνίζουν σε σχέση με τα άτομα που σταματήσαν το κάπνισμα.

γ) Λόγος σχετικών πιθανοτήτων:

$$\begin{aligned} \widehat{OR} &= \frac{19 \times 199}{15 \times 135} = 1.867 \\ \log \widehat{OR} &= \log 1.867 = 0.6244 \\ se(\log \widehat{OR}) &= \sqrt{\frac{1}{19} + \frac{1}{135} + \frac{1}{15} + \frac{1}{199}} = \sqrt{0.1317} = 0.363. \end{aligned}$$

Το 95% ΔΕ για το λογάριθμο του ΛΣΠ δίνεται από τον τύπο $0.6244 \pm 1.96 \times 0.363 = (-0.0867, 1.3358)$ ενώ για το σχετικό κίνδυνο είναι ίσο με $(e^{-0.0867}, e^{1.3358}) = (0.9167, 3.8030)$. Μπορούμε να ελέγξουμε την υπόθεση $H_0 : OR = 1$ έναντι της εναλλακτικής $H_1 : OR \neq 1$ με την ελεγχουσυνάρτηση $Z_{OR} = 0.6244/0.363 = 1.72 < z_{0.975}$ άρα δεν απορρίπτουμε την H_0 για $\alpha = 5\%$ και συνεπώς η σχετική πιθανότητα θανάτου δεν διαφοροποιείται σημαντικά για τα άτομα που συνέχισαν να καπνίζουν σε σχέση με τα άτομα που σταματήσαν το κάπνισμα.

3.6.2 Έλεγχος Ανεξαρτησίας χ^2 του του Pearson.

Η προηγούμενη ενότητα αναφέρεται στον έλεγχο της ισότητας δύο ποσοστών δηλαδή στο έλεγχο της σχέσης δύο δίτιμων κατηγορικών μεταβλητών. Όταν θέλουμε να ελέγξουμε γενικότερα την ισότητα της πιθανότητας εμφάνισης της νόσου σε πολλές διαφορετικές ομάδες ή γενικότερα μεταξύ δύο κατηγορικών μεταβλητών με πολλά επίπεδα τότε χρησιμοποιούμε τον έλεγχο της ανεξαρτησίας του Pearson.

Ας υποθέσουμε λοιπόν ότι έχουμε X και Y κατηγορικές μεταβλητές με I και J επίπεδα. Επιθυμούμε να ελέγξουμε την υπόθεση H_0 : «Ανεξαρτησία μεταξύ X και Y » έναντι της εναλλακτικής H_1 : « X και Y είναι εξαρτημένες μεταβλητές». Τότε, όπως είδαμε και στην ενότητα 3.2.1.3 η ανεξαρτησία συνεπάγεται ότι

$$\pi_{ij} = \pi_{i\bullet} \pi_{\bullet j}.$$

Οι αναμενόμενες τιμές των κελιών i, j κάτω από την υπόθεση της ανεξαρτησίας θα είναι ίσες με

$$E(N_{ij}|H_0) = e_{ij} = n\pi_{i\bullet} \pi_{\bullet j}$$

οι οποίες θα εκτιμηθούν στο δείγμα από τις ποσότητες

$$e_{ij} = n \frac{n_{i\bullet}}{n} \frac{n_{\bullet j}}{n} = \frac{n_{i\bullet} n_{\bullet j}}{n}.$$

Αν υποθέσουμε ότι ο αριθμός των παρατηρήσεων σε κάθε κελί ακολουθεί Poisson κατανομή δηλαδή

$$N_{ij} \sim \text{Poisson}(\varepsilon_{ij})$$

τότε

$$\frac{N_{ij} - \varepsilon_{ij}}{\sqrt{\varepsilon_{ij}}} \underset{\text{ασυμ.}}{\sim} N(0, 1).$$

Συνεπώς

$$\left(\frac{N_{ij} - \varepsilon_{ij}}{\sqrt{\varepsilon_{ij}}} \right)^2 \underset{\text{ασυμ.}}{\sim} \chi_1^2$$

και το άθροισμα τους ακολουθεί τη χ^2 κατανομή με $(I - 1)(J - 1)$ βαθμούς ελευθερίας

$$\sum_{i=1}^I \sum_{j=1}^J \frac{(N_{ij} - \varepsilon_{ij})^2}{\varepsilon_{ij}} \underset{\text{ασυμ.}}{\sim} \chi_{(I-1)(J-1)}^2.$$

Πρακτικά στο δείγμα χρησιμοποιούμε τα αντίστοιχα δειγματικά μέτρα δηλαδή την παρατηρούμενη συχνότητα n_{ij} και την εκτιμώμενη από το δείγμα αναμενόμενη τιμή e_{ij} κάθε κελιού (i, j) .

Συνεπώς υπολογίζουμε

$$\chi_{obs}^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(N_{ij} - \varepsilon_{ij})^2}{\varepsilon_{ij}}. \quad (3.6)$$

Απορρίπτουμε την υπόθεση της ανεξαρτησίας όταν $\chi_{obs}^2 > \chi_{(I-1)(J-1), 1-\alpha}^2$. Η παραπάνω προσέγγιση είναι ικανοποιητική για $e_{ij} \geq 5$.

Όσον αφορά τους 2×2 πίνακες συνάφειας, η συνάρτηση ελέγχου μπορεί να απλοποιηθεί στην ακόλουθη ποσότητα (βλ. Armitage και Berry, 1994, σελ. 135)

$$\chi_{obs}^2 = \frac{n(n_{11}n_{22} - n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}} = z^2. \quad (3.7)$$

Ο έλεγχος ανεξαρτησίας του Pearson στους 2×2 πίνακες είναι ισοδύναμος με τον έλεγχο για τον αποδιδόμενο κίνδυνο που περιγράψαμε στην ενότητα 3.6.1.1.

Στους πίνακες 2×2 για λόγους καλύτερης προσέγγισης χρησιμοποιείται το χ^2 τεστ με τη διόρθωση του Yates το οποίο δίνεται από τον τύπο:

$$\chi_{Yates}^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(|n_{ij} - e_{ij}| - 1/2)^2}{e_{ij}} \sim \chi_1^2. \quad (3.8)$$

Παράδειγμα 3.2 (συνέχεια II): Στο παράδειγμά μας

$$\chi_{obs}^2 = \frac{368(19 \times 199 - 15 \times 135)^2}{34 \times 334 \times 154 \times 214} = 3.03 (= 1.74^2) < \chi_{0.95}^2 = 3.841$$

άρα δεν απορρίπτουμε την υπόθεση της ανεξαρτησίας.

Εναλλακτικά μπορούμε να χρησιμοποιήσουμε τον τύπο (3.6) οπότε

$$\begin{aligned} e_{11} &= 34 \times 154/368 = 14.23 \\ e_{12} &= 139.77 \\ e_{21} &= 19.77 \\ e_{22} &= 194.2 \\ \chi_{obs}^2 &= \frac{(19 - 14.23)^2}{14.23} + \frac{(135 - 139.77)^2}{139.77} + \frac{(15 - 19.77)^2}{19.77} + \frac{(199 - 194.23)^2}{194.23} \\ &= 3.03 \Leftrightarrow p - value = 0.082. \end{aligned}$$

Τέλος αν χρησιμοποιήσουμε το χ^2 με τη διόρθωση του Yates έχουμε

$$\begin{aligned} \chi_{Yates}^2 &= \frac{(|19 - 14.23| - 1/2)^2}{14.23} + \frac{(|135 - 139.77| - 1/2)^2}{139.77} + \frac{(|15 - 19.77| - 1/2)^2}{19.77} + \frac{(|199 - 194.23| - 1/2)^2}{194.23} \\ &= 2.43 \Leftrightarrow p - value = 0.119. \end{aligned}$$

Συνεπώς δεν απορρίπτουμε την υπόθεση της ανεξαρτησίας για επίπεδο σημαντικότητας 5%.

3.6.3 Έλεγχος Μεγίστης Πιθανοφάνειας.

Το 1935 ο Wilks απέδειξε

$$G^2 = -2 \log \frac{L_0}{L_1} \sim \chi_{d_1 - d_0}^2$$

όπου L_k είναι η πιθανοφάνεια και d_k είναι ο αριθμός των εκτιμώμενων παραμέτρων κάτω από από την υπόθεση H_k , για $k = 0, 1$. Όταν έχουμε $I \times J$ πίνακες οι υποθέσεις είναι

H_0 : «X,Y ανεξάρτητες μεταβλητές» έναντι της H_1 : «X,Y εξαρτημένες μεταβλητές».

Όταν η H_1 ισχύει, πρέπει να εκτιμήσουμε όλες τις από κοινού πιθανότητες π_{ij} που είναι $d_1 = IJ - 1$ (αφαιρούμε μία επειδή ισχύει ότι $\sum_{i=1}^I \sum_{j=1}^J \pi_{ij} = 1$ άρα εκτιμούμε μία λιγότερη αφού την τελευταία την υπολογίζουμε απλά ως συνάρτηση των υπολοίπων). Όταν η H_0 ισχύει, πρέπει να εκτιμήσουμε τις περιθώριες κατανομές $\pi_{i\bullet}$ και $\pi_{\bullet j}$ δηλαδή πρέπει να εκτιμήσουμε $d_0 = (I - 1) + (J - 1) = I + J - 2$ παραμέτρους. Συνεπώς $d_1 - d_0 = IJ - 1 - I - J + 2 = IJ - I - J + 1 = I(J - 1) - (J - 1) = (I - 1)(J - 1)$.

Επιπλέον, υποθέτοντας πολυωνυμική κατανομή έχουμε

$$\begin{aligned} L_0 &= \prod_{i=1}^I \prod_{j=1}^J (\pi_{i\bullet} \pi_{\bullet j})^{N_{ij}} \Leftrightarrow \log L_0 = \sum_{i=1}^I \sum_{j=1}^J N_{ij} \log \pi_{i\bullet} + \sum_{i=1}^I \sum_{j=1}^J N_{ij} \log \pi_{\bullet j} \\ L_1 &= \prod_{i=1}^I \prod_{j=1}^J \pi_{ij}^{N_{ij}} \Leftrightarrow \log L_1 = \sum_{i=1}^I \sum_{j=1}^J N_{ij} \log \pi_{ij}. \end{aligned}$$

άρα

$$\begin{aligned} G^2 &= -2(\log L_0 - \log L_1) = -2 \sum_{i=1}^I \sum_{j=1}^J N_{ij} \log \frac{\pi_{i\bullet} \pi_{\bullet j}}{\pi_{ij}} \\ &= -2 \sum_{i=1}^I \sum_{j=1}^J N_{ij} \log \frac{\varepsilon_{ij}/n}{\pi_{ij}} = -2 \sum_{i=1}^I \sum_{j=1}^J N_{ij} \log \frac{\varepsilon_{ij}}{n\pi_{ij}} \Rightarrow \\ G_{obs}^2 &= -2 \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log \frac{e_{ij}}{n_{ij}}. \end{aligned}$$

Όπως και στον έλεγχο του Pearson απορρίπτουμε την υπόθεση της ανεξαρτησίας αν $G_{obs}^2 > \chi_{(I-1)(J-1), 1-\alpha}^2$. Προϋπόθεση για να ισχύει ασυμπτωτικά ότι $G^2 \sim \chi_{(I-1)(J-1)}^2$ είναι $n/(IJ) \geq 5$ η οποία δεν είναι τόσο αυστηρή όσο οι προϋποθέσεις του ελέγχου ανεξαρτησίας του Pearson.

Από τα παραπάνω βλέπουμε ότι και το G^2 είναι ένα μέτρο απόκλισης μεταξύ των παρατηρούμενων (n_{ij}) και των αναμενόμενων (e_{ij}) συχνοτήτων. Το αποτέλεσμα είναι ανάλογο αν υποθέσουμε ότι η κατανομή δειγματοληψίας είναι Poisson ή διωνυμική. Όταν η μηδενική υπόθεση της ανεξαρτησίας είναι αληθής τότε και η χ^2 δοκιμασία του Pearson και ο έλεγχος πιθανοφάνειας θα καταλήγουν στο ίδιο συμπέρασμα.

Παράδειγμα 3.2 (συνέχεια III): Στο παράδειγμά μας

$$\begin{aligned} G_{obs}^2 &= -2(19 \log 14.2319 + 135 \log 139.77135 + 15 \log 19.7715 + 199 \log 194.23199) \\ &= 2.982695 \Leftrightarrow p - value = 0.0842. \end{aligned}$$

Συνεπώς δεν απορρίπτουμε την υπόθεση της ανεξαρτησίας για επίπεδο σημαντικότητας 5%.

3.6.4 Ακριβής Έλεγχος Ανεξαρτησίας του Fisher.

Οι παραπάνω έλεγχοι είναι ασυμπτωτικοί που σημαίνει ότι για να τους εφαρμόσουμε θα πρέπει το δείγμα πρέπει να είναι ικανοποιητικά μεγάλο και ως εκ τούτου πρέπει να ισχύουν κάποιες προϋποθέσεις. Όταν λέμε ότι ένας έλεγχος είναι ακριβής (exact) εννοούμε ότι δε βασίζεται σε ασυμπτωτικά αποτελέσματα και συνεπώς δεν προαπαιτεί την ισχύ προϋποθέσεων όσον αφορά το μέγεθος και τη σύνθεση του δείγματος. Συνήθως οι ασυμπτωτικοί έλεγχοι χρησιμοποιούνται συχνότερα διότι βασίζονται στην κανονική και στις παράγωγες της κατανομές (όπως για παράδειγμα τη χ^2). Συνεπώς οι κριτικές τιμές και τα p – value και μπορούν να υπολογιστούν πολύ πιο εύκολα από αυτές τις κατανομές.

Εδώ θα δούμε τον ακριβή έλεγχο του Fisher (Fisher's exact test) ο οποίος βασίζεται στο αποτέλεσμα ότι αν η υπόθεση της ανεξαρτησίας ισχύει και κρατήσουμε σταθερές τις περιθώριες συχνότητες τότε το N_{11} θα ακολουθεί υπεργεωμετρική κατανομή. Έτσι μπορούμε να υπολογίσουμε όλους τους πιθανούς πίνακες που αντιστοιχούν στις συγκεκριμένες περιθώριες συχνότητες και τις αντίστοιχες πιθανότητες με βάση την υπεργεωμετρική κατανομή. Από αυτές μπορούμε να υπολογίσουμε το p – value ως άθροισμα των πιθανοτήτων που αντιστοιχούν σε πίνακες με χειρότερη ή ίση τιμή ελεγχουσυνάρτησης από αυτή που έχουμε παρατηρήσει.

Παράδειγμα 3.4 Έστω ότι έχουμε παρατηρήσει τον πίνακα

X	Y		Περ.Κατ. X
	A	\bar{A}	
E	3	1	4
\bar{E}	1	3	4
Περ.Κατ. Y	4	4	8

και μας ενδιαφέρει να ελέγξουμε την υπόθεση $H_0 : OR = 1$ έναντι της $H_1 : OR > 1$. Φυσικά δε μπορούμε να χρησιμοποιήσουμε τις παραπάνω ασυμπτωτικές διαδικασίες αλλά θα πρέπει να χρησιμοποιήσουμε την ακριβή δοκιμασία του Fisher. Εφόσον οι περιθώριες συχνότητες είναι σταθερές οι πιθανές τιμές του n_{11} είναι 0, 1, 2, 3, 4. Άρα πιθανοί πίνακες είναι

X	Y		Περ.Κατ. X	Πίνακας Α
	A	\bar{A}		
E	3	1	4	$n_{11} = 3$
\bar{E}	1	3	4	
Περ.Κατ. Y	4	4	8	OR = 9

X	Y		Περ.Κατ. X	Πίνακας Β
	A	\bar{A}		
E	3+1=4	1-1=0	4	$n_{11} = 4$
\bar{E}	1-1=0	3+1=4	4	
Περ.Κατ. Y	4	4	8	OR = ∞ ή $OR_{cor} = 81$

X	Y		Περ.Κατ. X	Πίνακας Γ
	A	\bar{A}		
E	3-1=2	1+1=2	4	$n_{11} = 2$
\bar{E}	1+1=2	3-1=2	4	
Περ.Κατ. Y	4	4	8	OR = 1

X	Y		Περ.Κατ. X	Πίνακας Δ
	A	\bar{A}		
E	3-2=1	1+2=3	4	$n_{11} = 1$
\bar{E}	1+2=3	3-2=1	4	
Περ.Κατ. Y	4	4	8	OR = 1/9 = 0.111

X	Y		Περ.Κατ. X	Πίνακας Ε
	A	\bar{A}		
E	3-3=0	1+3=4	4	$n_{11} = 0$
\bar{E}	1+3=4	3-3=0	4	
Περ.Κατ. Y	4	4	8	OR = 0 ή $OR_{cor} = 1/81 = 0.012$

Για να υπολογίσουμε το p – value θα βρούμε τις πιθανότητες των πινάκων με OR μεγαλύτερο ή ίσο του 9 που παρατηρήσαμε στο δείγμα. Συνεπώς

$$\begin{aligned}
 p - \text{value} &= P(A) + P(B) = \frac{\binom{4}{3} \binom{4}{1}}{\binom{8}{4}} + \frac{\binom{4}{4} \binom{4}{0}}{\binom{8}{4}} \\
 &= P(A) + P(B) = \frac{4 \times 4}{5 \times 6 \times 7 \times 8 / (2 \times 3 \times 4)} + \frac{1}{5 \times 6 \times 7 \times 8 / (2 \times 3 \times 4)} \\
 &= \frac{16}{5 \times 2 \times 7} + \frac{1}{5 \times 2 \times 7} = \frac{17}{70} = 0.24286 .
 \end{aligned}$$

συνεπώς δεν απορρίπτουμε την H_0 . Αν κάνουμε τον έλεγχο δύο ουρών τότε

$$p - \text{value} = P(\text{Πίνακες με } OR \geq 9) + P(\text{Πίνακες με } OR \leq 1/9)$$

$$\begin{aligned}
&= P(A) + P(B) + P(\Delta) + P(E) = 1 - P(\Gamma) = 1 - \frac{\binom{4}{2} \binom{4}{2}}{\binom{8}{4}} \\
&= 1 - \frac{36}{70} = 1 - 0.514 = 0.486
\end{aligned}$$

Δεν απορρίπτουμε την H_0 ούτε στον έλεγχο των δύο ουρών.

Γενικά στους 2×2 πίνακες η πιθανότητα ενός πίνακα δίνεται και από τον τύπο

$$P(n_{11}, n_{12}, n_{21}, n_{22} | n_{\bullet 1}, n_{\bullet 2}, n_{1 \bullet}, n_{2 \bullet}) = \frac{n_{\bullet 1}! n_{\bullet 2}! n_{1 \bullet}! n_{2 \bullet}!}{n! n_{11}! n_{12}! n_{21}! n_{22}!},$$

για λεπτομέρειες παραπέμπουμε στο βιβλίο του Rosner (1994, σελ. 371).

Η διαδικασία του ελέγχου του Fisher μπορεί να περιγραφεί περιληπτικά από τα ακόλουθα βήματα :

1. Κατασκευάζουμε όλους τους πιθανούς πίνακες που αντιστοιχούν στις περιθώριες συχνότητες που έχουμε παρατηρήσει.
2. Υπολογίζουμε την συνάρτηση ελέγχου για κάθε πίνακα.
3. Υπολογίζουμε την πιθανότητα εμφάνισης για τους πίνακες με χειρότερη (πιο απομακρυσμένη) ή ίση τιμή ελεγχουσυνάρτησης από αυτή που παρατηρήσαμε στο δείγμα μας.
4. Υπολογίζουμε το p-value

(α') Αν $H_1 : \pi_E \neq \pi_{\bar{E}}$ ή $H_1 : OR \neq 1$ τότε $p\text{-value} = 2 \min\{P(N_{11} \leq n_{11}), P(N_{11} \geq n_{11})\}$.

(β') Αν $H_1 : \pi_E > \pi_{\bar{E}}$ ή $H_1 : OR > 1$ τότε $p\text{-value} = P(N_{11} \leq n_{11})$.

(γ') Αν $H_1 : \pi_E < \pi_{\bar{E}}$ ή $H_1 : OR < 1$ τότε $p\text{-value} = P(N_{11} \geq n_{11})$.

Οι ακριβείς έλεγχοι ισχύουν πάντα αλλά τους εφαρμόζουμε μόνο αν $e_{ij} < 5$ όποτε δε μπορούμε να χρησιμοποιήσουμε τους αντίστοιχους ασυμπτωτικούς ελέγχους.

3.6.5 Μόντε Κάρλο Έλεγχοι Ανεξαρτησίας.

Σε αρκετές περιπτώσεις, όταν χρειάζεται να υπολογίσουμε τα ακριβή τεστ, ο αριθμός των πιθανών πινάκων είναι μεγάλος. Στην περίπτωση αυτή μπορούμε, αντί να υπολογίσουμε όλους

τους πίνακες, να προσομοιώσουμε ένα ορισμένο αριθμό από αυτούς (δηλαδή να κατασκευάσουμε ένα τυχαίο δείγμα) έτσι ώστε να εκτιμήσουμε το πραγματικό $p - value$. Το μέγεθος του δείγματος των πινάκων που προσομοιώνουμε είναι συνήθως μεγάλο (π.χ. 10000 πίνακες) έτσι ώστε να εκτιμήσουμε με ακρίβεια το απαιτούμενο $p - value$. Ως αποτέλεσμα παίρνουμε ένα διάστημα εμπιστοσύνης (συνήθως 99%) για το $p - value$. Για να απορρίψουμε την H_0 πρέπει όλο το διάστημα εμπιστοσύνης να είναι μικρότερο του α που χρησιμοποιούμε για τον έλεγχο της μηδενικής υπόθεσης που μας ενδιαφέρει. Αν το α εμπεριέχεται στο ΔΕ του $p - value$ τότε δεν μπορούμε να αποφασίσουμε για την H_0 ενώ η H_0 δεν απορρίπτεται αν το α είναι μικρότερο από όλο το ΔΕ του $p - value$.

3.7 Διαγνωστικοί έλεγχοι

Ως διαγνωστικούς ελέγχους (diagnostic tests) ή προσυμπτωτικούς ελέγχους (screening tests) ονομάζουμε διάφορους Ιατρικούς ελέγχους που γίνονται με σκοπό τον εντοπισμό και την πρόληψη ασθενειών. Τέτοιοι έλεγχοι είναι οι σειρά των ετήσιων Ιατρικών check up, οι αιματολογικές εξετάσεις, το τεστ Παπανικολάου και διάφορες ακτινογραφίες. Σκοπός τους είναι ο εντοπισμός ασθενειών που δεν μπορούν να φανούν με άλλο τρόπο ή ο εντοπισμός της νόσου σε πρώιμο στάδιο (πριν την εμφάνιση συμπτωμάτων) ώστε να αντιμετωπιστεί εγκαίρως και αποτελεσματικά. Χαρακτηριστικό των διαγνωστικών ελέγχων είναι ότι αν κάποιος εμφανίσει θετικό αποτέλεσμα τότε είναι πιθανό να έχει την υπό εξέταση νόσο και συνεπώς υποβάλλεται σε επιπλέον (και πιο μεγάλης ακρίβειας) εξετάσεις ή κατευθείαν σε θεραπεία (αν το διαγνωστικό τεστ είναι πολύ ακριβές).

Στις περιπτώσεις αυτές είναι πολύ σημαντικό το θεώρημα του Bayes το οποίο δίνεται από τον τύπο

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_j P(A|B_j)P(B_j)}$$

με $\sum_j P(B_j) = 1$ και $A_i \cap A_j = \emptyset$ για κάθε $i \neq j$.

3.7.1 Ευαισθησία και Ειδικότητα.

Σε όλη την παρακάτω ενότητα θα συμβολίζουμε με A και \bar{A} το ενδεχόμενο παρουσίας ή απουσίας μιας ασθένειας και με T^+ και T^- τη θετική ή αρνητική διάγνωση ενός τεστ ως προς την παρουσία της ίδιας ασθένειας.

Ορισμός 3.1 Η θετική προγνωστική τιμή (positive predictive value) ενός διαγνωστικού τεστ είναι η πιθανότητα κάποιος να είναι ασθενής όταν ο διαγνωστικός έλεγχος έχει βρεθεί θετικός

δηλαδή

$$PV^+ = P(A | T^+).$$

Ορισμός 3.2 Η αρνητική προγνωστική τιμή (*negative predictive value*) ενός διαγνωστικού τεστ είναι η πιθανότητα κάποιος να μην είναι ασθενής όταν ο διαγνωστικός έλεγχος έχει βρεθεί αρνητικός δηλαδή

$$PV^- = P(\bar{A} | T^-).$$

Παράδειγμα 3.5 Η πιθανότητα εμφάνισης μίας νόσου είναι ένα άτομο στα 10000 όταν ένας έλεγχος βρεθεί αρνητικός και ένα άτομο στα 100 όταν βρεθεί ο ίδιος έλεγχος θετικός. Τότε η θετική προγνωστική τιμή είναι ίση με $PV^+ = 1/100 = 0.01$ ενώ αρνητική προγνωστική τιμή είναι ίση με $PV^- = 1 - 1/10000 = 0.9999$.

Για να έχουμε ένα καλό διαγνωστικό έλεγχο θα πρέπει και οι δύο προγνωστικές τιμές να είναι υψηλές και κοντά στο ένα. Σε πολλές περιπτώσεις μόνο η μια από τις προγνωστικές τιμές είναι υψηλή (όπως στο παράδειγμα παραπάνω). Σε αυτή την περίπτωση ο διαγνωστικός έλεγχος χρησιμοποιείται για να αποκλείσει κάποιες περιπτώσεις. Έτσι στο παραπάνω παράδειγμα, αν κάποιος βγει αρνητικός τότε είμαστε σίγουροι ότι δεν έχει τη νόσο (αποκλείσαμε δηλαδή αυτή την περίπτωση). Αντίθετα αν βγει θετικός τότε δε σημαίνει ότι έχει τη νόσο αφού το PV^+ είναι πολύ χαμηλό και ο γιατρός οφείλει να παραπέμψει τον ασθενή για επιπλέον εξετάσεις με μεγαλύτερη διαγνωστική ακρίβεια (δηλαδή υψηλότερο PV^+).

Πολλές φορές, στη θέση του διαγνωστικού ελέγχου μπορεί να είναι και ένα απλό σύμπτωμα μιας ασθένειας.

Ορισμός 3.3 Ευαισθησία (*sensitivity*) ενός διαγνωστικού τεστ (ή συμπτώματος) είναι η πιθανότητα το τεστ να είναι θετικό (ή να εμφανίζεται το σύμπτωμα) δεδομένου ότι κάποιος έχει τη νόσο που εξετάζουμε δηλαδή

$$\text{sensitivity} = a = P(T^+ | A).$$

Ορισμός 3.4 Ειδικότητα (*specificity*) ενός διαγνωστικού τεστ (ή συμπτώματος) είναι η πιθανότητα το τεστ να είναι αρνητικό (ή να μην εμφανίζεται το σύμπτωμα) δεδομένου ότι κάποιος δεν έχει τη νόσο που εξετάζουμε δηλαδή

$$\text{specificity} = \beta = P(T^- | \bar{A}).$$

Ορισμός 3.5 Ψευδώς αρνητική περίπτωση (*false negative case*) ονομάζεται ένα άτομο που έχει τη νόσο αλλά το διαγνωστικό τεστ είναι αρνητικό.

Ορισμός 3.6 Ψευδώς θετική περίπτωση (*false positive case*) ονομάζεται ένα άτομο που δεν έχει τη νόσο αλλά το διαγνωστικό τεστ είναι θετικό.

Σημείωση ότι η ευαισθησία και η ειδικότητα είναι αντίστοιχες με τα σφάλματα τύπου I και II στους στατιστικούς ελέγχους υποθέσεων.

Σε μία μελέτη συνήθως παρατηρούμε την ειδικότητα και την ευαισθησία και όχι τις προγνωστικές τιμές που μας ενδιαφέρουν άμεσα. Και αυτές οι δύο τιμές πρέπει να είναι υψηλές και κοντά στο ένα για να έχουμε ένα καλό διαγνωστικό τεστ.

Συνήθως η ειδικότητα και η ευαισθησία υπολογίζονται από μελέτες μαρτύρων-ασθενών. Σε αυτές παίρνουμε ένα προκαθορισμένο αριθμό ασθενών και μαρτύρων, τους κάνουμε τον διαγνωστικό έλεγχο που επιθυμούμε και εκτιμούμε άμεσα τις παραπάνω ποσότητες από τις πιθανότητες $p_{1|1}$ και $p_{2|2}$. Ας δούμε όμως τι γίνεται με τις προγνωστικές τιμές που μας ενδιαφέρουν άμεσα:

$$\begin{aligned} PV^+ &= P(A | T^+) = \frac{P(A)P(T^+|A)}{P(A)P(T^+|A) + P(\bar{A})P(T^+|\bar{A})} \\ &= \frac{P(A) \textit{sensitivity}}{P(A) \textit{sensitivity} + \{1 - P(A)\}\{1 - P(T^+|\bar{A})\}} \\ &= \frac{\Pi_A \textit{sensitivity}}{\Pi_A \textit{sensitivity} + (1 - \Pi_A)(1 - \textit{specificity})} . \end{aligned} \quad (3.9)$$

όπου $\Pi_A = P(A)$ δηλαδή ο επιπολασμός της νόσου. Με παρόμοιο τρόπο βρίσκουμε ότι

$$PV^- = \frac{(1 - \Pi_A) \textit{specificity}}{(1 - \Pi_A) \textit{specificity} + \Pi_A(1 - \textit{sensitivity})} . \quad (3.10)$$

Συνεπώς μπορούμε να υπολογίσουμε τις προγνωστικές τιμές και σε μελέτες μαρτύρων-ασθενών όπου είναι εύκολη η εκτίμηση της ευαισθησίας και της ειδικότητας αλλά θα πρέπει να γνωρίζουμε τον επιπολασμό της νόσου που εξετάζουμε.

Παράδειγμα 3.6 Έστω μια μελέτη όπου λάβαμε 100 υπερτασικούς και 100 άτομα με κανονική πίεση. Από τους υπερτασικούς οι 84 είχαν ένα διαγνωστικό τεστ θετικό ενώ από τους συμμετέχοντες με κανονική πίεση βρέθηκαν 23 άτομα με τεστ θετικό. Η αναλογία υπερτασικών στο σύνολο του πληθυσμού είναι 1:4.

Από τα παραπάνω έχουμε

$$\begin{aligned} \Pi_A &= 0.2 \\ \textit{sensitivity} &= P(T^+|A) = 84/100 = 0.84 \\ \textit{specificity} &= P(T^-|\bar{A}) = 1 - P(T^+|\bar{A}) = 1 - 23/100 = 0.77 . \end{aligned}$$

Από τα παραπάνω και με τη χρήση των τύπων (3.9) και (3.10) προκύπτει ότι

$$PV^+ = \frac{0.2 \times 0.84}{0.2 \times 0.84 + 0.8 \times 0.23} = 0.48,$$

$$PV^- = \frac{0.8 \times 0.77}{0.8 \times 0.77 + 0.2 \times 0.16} = 0.95 .$$

Συνεπώς το τεστ έχει μόνο αρνητική προγνωστική τιμή υψηλή και μπορεί να χρησιμοποιηθεί κυρίως για τον αποκλεισμό της συγκεκριμένης νόσου.

3.7.2 Ευαισθησία και ειδικότητα σε 2×2 πίνακες

Έστω ότι τα δεδομένα μιας διαγνωστικής μελέτης πινακοποιούνται με την ακόλουθη μορφή

Διαγνωστικό Τέστ	Νόσος		Περ.Κατ. X
	A	\bar{A}	
T^+	n_{11}	n_{12}	$n_{1\bullet}$
T^-	n_{21}	n_{22}	$n_{2\bullet}$
Περ.Κατ. Y	$n_{\bullet 1}$	$n_{\bullet 2}$	n

α) Μελέτες Κοορτής: Σε μελέτες κοορτής, οι προγνωστικές τιμές μπορούν να υπολογιστούν κατευθείαν. Συνεπώς έχουμε

$$\widehat{PV}^+ = \widehat{P}(A|T^+) = n_{11}/n_{1\bullet} = n_{11}/(n_{11} + n_{12})$$

$$\widehat{PV}^- = \widehat{P}(\bar{A}|T^-) = n_{22}/n_{2\bullet} = n_{22}/(n_{21} + n_{22})$$

$$\widehat{\Pi}_A = \widehat{P}(A) = n_{\bullet 1}/n = (n_{11} + n_{21})/n .$$

Αν η κοορτή είναι ένα αντιπροσωπευτικό δείγμα του πληθυσμού που εξετάζουμε τότε μπορούμε να υπολογίσουμε και την ειδικότητα και την ευαισθησία:

$$\widehat{Sensitivity} = \widehat{P}(T^+|A) = n_{11}/n_{\bullet 1} = n_{11}/(n_{11} + n_{21})$$

$$\widehat{Specificity} = \widehat{P}(T^-|\bar{A}) = n_{22}/n_{\bullet 2} = n_{22}/(n_{12} + n_{22}) .$$

β) Μελέτες Μαρτύρων-Ασθενών: Στις μελέτες μαρτύρων-ασθενών υπολογίζουμε με τον ίδιο τρόπο όπως παραπάνω την ευαισθησία και την ειδικότητα. Επιπλέον πρέπει να γνωρίζουμε τον επιπολασμό από προηγούμενες μελέτες και μετά να χρησιμοποιήσουμε τους τύπους (3.9) και (3.10) για εκτιμήσουμε τις προγνωστικές τιμές. Έτσι θα έχουμε:

$$\widehat{PV}^+ = \frac{\widehat{\Pi}_A n_{11}/(n_{11} + n_{21})}{\widehat{\Pi}_A n_{11}/(n_{11} + n_{21}) + (1 - \widehat{\Pi}_A)n_{12}/(n_{12} + n_{22})}$$

$$\widehat{PV}^- = \frac{(1 - \widehat{\Pi}_A) n_{22}/(n_{12} + n_{22})}{(1 - \widehat{\Pi}_A) n_{22}/(n_{12} + n_{22}) + \widehat{\Pi}_A n_{21}/(n_{11} + n_{21})} .$$

3.7.3 Καμπύλες ROC

Η παραπάνω διαδικασία περιγράφει τη διάγνωση μίας νόσου με τη χρήση ενός διαγνωστικού ελέγχου του οποίου το αποτέλεσμα είναι δίτιμο (θετικό ή αρνητικό τεστ). Συνήθως όμως οι μετρήσεις των διαγνωστικών ελέγχων γίνονται με ποσοτικές - συνεχείς μεταβλητές. Σε αυτή την περίπτωση μας ενδιαφέρει το σημείο διαχωρισμού (cut-off point) πέρα από το οποίο θεωρούμε ότι το τεστ είναι θετικό (δηλαδή το σημείο πέρα από το οποίο υπάρχει αυξημένη πιθανότητα εμφάνισης της νόσου). Δηλαδή, αν συμβολίσουμε με T την μέτρηση του διαγνωστικού τεστ, ψάχνουμε για μια τιμή t για την οποία για $T \leq t$ θα θεωρούμε ότι η νόσος δεν είναι παρούσα (αρνητικό τεστ) ενώ για $T > t$ θα θεωρούμε ότι η νόσος είναι παρούσα (τεστ θετικό).

Στη επιλογή του σημείου διαχωρισμού μας βοηθάνε οι καμπύλες ROC (Receiver Operating Characteristic curves) οι οποίες απεικονίζουν με μια καμπύλη τους συνδυασμούς της αναλογίας των ψευδών θετικών περιπτώσεων ($1 - specificity$) και της ευαισθησίας (στους άξονες X και Y αντίστοιχα) για όλες τις τιμές που παρατηρούμε στο δείγμα. Τιμές κατάλληλες ως σημεία διαχωρισμού είναι τα σημεία της καμπύλης που βρίσκονται κοντά στην άνω αριστερή γωνία του διαγράμματος (και έχουν χαμηλό αριθμό ψευδών θετικών περιπτώσεων και υψηλή ευαισθησία).

Συνεπώς οι καμπύλες ROC είναι η διαγραμματική απεικόνιση των χαρακτηριστικών ενός ποσοτικού διαγνωστικού τεστ και μας βοηθάνε να εξετάσουμε την απόδοση του τεστ για διαφορετικά σημεία προγνωστικού ελέγχου.

Σημαντική ποσότητα στις καμπύλες ROC είναι η τιμή AUC (area under curve) η οποία μας δίνει την περιοχή κάτω από την καμπύλη δηλαδή το εμβαδόν που περικλείεται μεταξύ της καμπύλης ROC και του άξονα των X . Η AUC μετράει την πιθανότητα η τιμή του τεστ ενός ασθενή (T^A) να είναι μεγαλύτερη από την τιμή του τεστ ενός ατόμου χωρίς την υπό εξέταση νόσο ($T^{\bar{A}}$) δηλαδή

$$AUC = P(T^A > T^{\bar{A}})$$

και εκτιμάται από

$$w = \frac{1}{n_A n_{\bar{A}}} \sum_{i=1}^{n_A} \sum_{j=1}^{n_{\bar{A}}} S(T_i^A, T_j^{\bar{A}})$$

όπου n_A και $n_{\bar{A}}$ ο αριθμός των ατόμων με ή χωρίς τη νόσο αντίστοιχα, T_i^A είναι η τιμή του διαγνωστικού τεστ για το i άτομο της ομάδας των ασθενών, $T_j^{\bar{A}}$ είναι η τιμή του διαγνωστικού τεστ για το j άτομο της ομάδας των μη ασθενών και $S(T_i^A, T_j^{\bar{A}})$ παίρνει την τιμή ένα (1) αν $T_i^A > T_j^{\bar{A}}$, την τιμή $1/2$ αν $T_i^A = T_j^{\bar{A}}$ και την τιμή μηδέν (0) αν $T_i^A < T_j^{\bar{A}}$.

Όταν ο διαγνωστικός έλεγχος συνδέεται αρνητικά με τη νόσο (δηλαδή μικρές τιμές του τεστ υποδεικνύουν μεγάλη πιθανότητα εμφάνισης της νόσου) τότε υπολογίσουμε την ποσότητα $w' = 1 - w$ ή απλά μετασχηματίζουμε το τεστ έτσι ώστε να συνδέεται θετικά με τη νόσο (για

παράδειγμα πολλαπλασιάζουμε με το μείον ένα).

Εδώ ενδιαφέρον έχει ο έλεγχος της υπόθεσης $H_0 : AUC = 0.5$ έναντι της εναλλακτικής $H_1 : AUC > 0.5$. Η τιμή $AUC = 0.5$ αντιστοιχεί σε ένα τεστ που μαντεύει τυχαία (όπως αν χρησιμοποιούσαμε ένα νόμισμα για να μαντέψουμε ποιος έχει τη νόσο) και ουσιαστικά δεν έχει καμία προγνωστική ικανότητα.

Για παραδείγματα βλέπε αντίστοιχες διαφάνειες (αρχείο «05_ROC EXAMPLE 1.pdf» στο δικτυακό τόπο eclass του μαθήματος).

Παράδειγμα 3.7 Παράδειγμα WAIS

Παράδειγμα 3.8 Παράδειγμα hivassay

3.7.4 Υπολογισμός Επιπολασμού από Διαγνωστικές Μελέτες.

Οι διαγνωστικές μελέτες μπορούν να χρησιμοποιηθούν και για τον υπολογισμό του επιπολασμού. Έτσι έχουμε

$$\begin{aligned} P(T^+) &= P(T^+, A) + P(T^+, \bar{A}) \\ &= P(T^+|A)P(A) + P(T^+|\bar{A})P(\bar{A}) \\ &= \Pi_A \textit{sensitivity} + (1 - \Pi_A)(1 - \textit{specificity}) \Leftrightarrow \\ \Pi_A &= \frac{P(T^+) - (1 - \textit{specificity})}{\textit{sensitivity} + \textit{specificity} - 1} . \end{aligned}$$

Προσοχή, όταν ο επιπολασμός είναι πολύ μικρός (δηλαδή έχουμε σπάνια ασθένεια) ο εκτιμητής αυτός δεν είναι αξιόπιστος και μπορεί να δώσει αρνητική τιμή.

Παράδειγμα 3.9 Έστω ότι έχουμε τα ακόλουθα δεδομένα από μια διαγνωστική μελέτη μαρτύρων ασθενών

Διαγνωστικό Τέστ	Νόσος		Περ.Κατ. X
	A	\bar{A}	
T^+	90	180	270
T^-	10	720	730
Περ.Κατ. Y	100	900	1000

Από τον παραπάνω πίνακα έχουμε

$$\begin{aligned} \widehat{\textit{Sensitivity}} &= 90/100 = 0.9 \\ \widehat{\textit{Specificity}} &= 720/900 = 0.8 \\ \widehat{P}(T^+) &= 270/1000 = 0.27 \\ \widehat{\Pi}_A &= \frac{0.27 - (1 - 0.8)}{0.9 + 0.8 - 1} = 0.07/0.7 = 0.1 . \end{aligned}$$

Συνεπώς 10% του συνολικού πληθυσμού έχει τη νόσο που εξετάζει η παραπάνω μελέτη.

Κεφάλαιο 4

Κλινικές Δοκιμές (Clinical Trials)

4.1 Εισαγωγή και Γενικές Έννοιες

Τι είναι οι Κλινικές Δοκιμές

Ένας πρώτος ορισμός αναφέρει ότι, κλινικές δοκιμές (clinical trials) είναι πειραματικές – παρεμβατικές μελέτες που συγκρίνουν την αποτελεσματικότητα διαφορετικών θεραπευτικών μέτρων σε ασθενείς (βλ. Armitage & Berry, 1994, σελ. 189).

Ένας δεύτερος ορισμός αναφέρει ότι, οι κλινικές δοκιμές είναι συγκριτικές μελέτες στις οποίες οι επιστήμονες παρεμβαίνουν στη φυσική πορεία της νόσου δίνοντας φάρμακα ή θεραπευτικές αγωγές σε τουλάχιστον μία από τις ομάδες μελέτης και μετά εξετάζουν την επίδραση των αγωγών στη νόσο (βλ. Perreira-Maxwell, 1998, AZ of Medical Statistics, σελ. 11).

Επιπλέον οι κλινικές δοκιμές μπορούν να οριστούν ως πολύ προσεκτικά σχεδιασμένες προοπτικές μελέτες παρέμβασης, που στόχο έχουν τον έλεγχο της αποτελεσματικότητας φαρμάκων (drugs) ή θεραπευτικών αγωγών (treatments).

Ποια είναι τα βασικά στοιχεία μιας Κλινικής Δοκιμής

Πρώτο βασικό στοιχείο σε μια κλινική δοκιμή είναι η **Θεραπευτική Αγωγή (treatment)** ή αλλιώς η **Τεχνική Παρέμβασης (intervention technique)**. Η θεραπευτική αγωγή είναι ο τρόπος παρέμβασης στον ασθενή, με σκοπό τη βελτίωση της κατάστασης του. Σε κάθε κλινική δοκιμή υπάρχει τουλάχιστον μια θεραπευτική αγωγή.

Δεύτερο βασικό στοιχείο είναι η **Ομάδα Ελέγχου ή Ψευδοφαρμάκου ή Αδρανούς Θεραπείας (control/placebo groups)**. Ως ψευδοφάρμακο ή αδρανής θεραπεία (placebo), ορίζεται η αγωγή κατά την οποία δε γίνεται ουσιαστική θεραπεία. Μια τέτοια ομάδα αναφοράς, δεν

λαμβάνει ουσιαστικά καμία θεραπεία, όμως στην πράξη παίρνει ένα εικονικό φάρμακο, για ψυχολογικούς κυρίως λόγους. Το εικονικό ψεύτικο ή αδρανές αυτό φάρμακο /χάπι ονομάζεται ψευδοφάρμακο. Το εικονικό αυτό ψευδοφάρμακο δίνεται στον ασθενή, έτσι ώστε αυτός να νομίζει ότι παίρνει παρόμοια θεραπεία με τις υπόλοιπες ομάδες με σκοπό να αποφεύγονται ψυχολογικές επιδράσεις. Εδώ θα πρέπει να προσθέσουμε ότι ενώ ως ομάδα ελέγχου ή αδρανούς θεραπείας, ονομάζεται η ομάδα ασθενών που δε λαμβάνει θεραπεία, συνήθως επειδή είναι αντιδεοντολογικό να μη δίνεται καθόλου θεραπεία, ως **ομάδα ελέγχου ή ομάδα αναφοράς (reference treatment group)**, ονομάζεται η ομάδα που παίρνει τη συνηθισμένη θεραπεία που επικρατεί στην ιατρική βιβλιογραφία.

Ένα τρίτο βασικό στοιχείο των κλινικών δοκιμών είναι ο όρος της **Μονής ή Διπλής Τύφλωσης/άγνοιας (Single or double blind)**. **Μονή Τύφλωση/άγνοια** ονομάζεται η κατάσταση κατά την οποία, για ψυχολογικούς λόγους και αποφυγή μεροληπιών συνήθως οι ασθενείς δε γνωρίζουν την αγωγή που λαμβάνουν (single blinded). Επιπλέον **Διπλή Τύφλωση/άγνοια** είναι η περίπτωση που και οι Γιατροί ή οι νοσοκόμοι που καταγράφουν τα δεδομένα δεν γνωρίζουν την αγωγή που λαμβάνει ο κάθε ασθενής (double blinded).

Ένα σημαντικό σημείο των κλινικών δοκιμών στο οποίο πρέπει να σταθούμε είναι το γεγονός ότι μια σωστά σχεδιασμένη και εκτελεσμένη κλινική δοκιμή είναι ένα ισχυρό πειραματικό εργαλείο/τεχνική για τη μελέτη και εκτίμηση της αποτελεσματικότητας μιας αγωγής.

Τελευταίο βασικό στοιχείο των κλινικών δοκιμών είναι η **Τυχαιοποίηση(Randomization/Random allocation)**. **Τυχαιοποίηση** ονομάζεται η τυχαία κατανομή αγωγών στα άτομα που συμμετέχουν σε μια κλινική δοκιμή. Επιπλέον έχουμε και την Τυχαιοποιημένη Ελεγχόμενη Δοκιμή (Randomized Controlled Trial), που συνήθως χρησιμοποιείται σε δοκιμές φάσης III ή σε κλινικές δοκιμές φάσεων I - IV.

4.2 Σύντομη Ιστορική Αναδρομή.

Η ιστορία των κλινικών δοκιμών αρχίζει από τον 18ο αιώνα όπου σε μια μελέτη του Lind για τους λόγους που προκαλείται το σκορβούτο σε ναύτες, ο Lind εξέτασε 6 θεραπείες σε 12 ναύτες του πλοίου "Salisbury". Δύο από τους 12 ασθενείς που έφαγαν πορτοκάλια και λεμόνια συνήλθαν σε 6 ημέρες. Στη συνέχεια το 1926 ο Fisher εισάγει την έννοια της τυχαιοποίησης, ενώ ο Amberson το 1931 (1931, *Am. Rev. Tuberc.*), διεξάγει την πρώτη τυχαιοποιημένη μελέτη, σύμφωνα με την οποία 24 ασθενείς χωρίστηκαν σε δύο ομάδες με τη χρήση ενός νομίσματος, ενώ εμφανίζεται και για πρώτη φορά από τον ίδιο η έννοια της τύφλωσης/άγνοιας θεραπείας, όπου ως ψευδοφάρμακο χρησιμοποιήθηκε νερό σε σύριγγα. Το 1938 εισάγεται ο όρος placebo (Dielh et al. 1938 JAMA), ενώ το 1948, γίνεται για πρώτη φορά αναφορά στην χρήση τυχαίων

αριθμών.

4.3 Φάσεις Κλινικών Δοκιμών

4.3.1 Μελέτες Φάσης I

Στην φάση I μελέτης και ανάπτυξης μίας θεραπείας ή φαρμάκου μας ενδιαφέρει να δούμε πόσο μπορεί να ανεχτεί ο ανθρώπινος οργανισμός το φάρμακο ή την αγωγή. Οι μελέτες αυτές γίνονται σε μικρό αριθμό ανθρώπων εφόσον έχουν προηγηθεί εργαστηριακές δοκιμές σε ζώα. Οι ασθενείς που συμμετέχουν στις μελέτες αυτής της φάσης έχουν συνήθως δοκιμάσει ανεπιτυχώς όλες τις υπόλοιπες θεραπείες (και συνεπώς είναι εθελοντές). Σκοπός της πρώτης φάσης είναι να εκτιμήσουμε το μέγιστο ύψος της δόσης του φαρμάκου που μπορεί να γίνει αποδεκτό από τον οργανισμό χωρίς σημαντικές παρενέργειες οι οποίες συνήθως μετριούνται από το βαθμό τοξικότητας (toxicity level). Αυτό το όριο ονομάζεται **μέγιστα αποδεκτή δόση - ΜΑΔ** (maximum tolerated dose - MTD) και για τον καθορισμό της, προκλινικές μελέτες σε μεγάλα ζώα βοηθούν (λίγο) στον καθορισμό της ΜΑΔ.

ΠΑΡΑΔΕΙΓΜΑ (από μελέτες καρκίνου):

1. Ένας μικρός αριθμός ασθενών (συνήθως 3) εισάγεται ακολουθιακά στη μελέτη και τους δίνεται μια μικρή δόση φαρμάκου .
 2. Αν σε κανένα ασθενή δεν παρατηρηθεί υψηλός βαθμός τοξικότητας τότε αυξάνουμε τη δόση.
 3. Αν έχουμε ένα ασθενή να ξεπερνά το όριο τοξικότητας τότε εισάγουμε άλλους 3 στη μελέτη στην ίδια δόση.
 4. Αν κανένας από τους νέους ασθενείς δεν ξεπεράσει το όριο τοξικότητας τότε περνάμε στο επόμενο επίπεδο της δόσης.
 5. Αν παρατηρήσουμε τουλάχιστον 2 ασθενείς που ξεπερνούν το όριο τοξικότητας τότε σταματάμε τη μελέτη και θέτουμε ΜΑΔ το αμέσως προηγούμενο επίπεδο δόσης.
- Ως ΜΑΔ ορίζεται η ποσότητα φαρμάκου στο οποίο το 1/3 των ασθενών ξεπερνά το όριο τοξικότητας

Το παραπάνω παράδειγμα είναι ενδεικτικό και φυσικά υπάρχουν παραλλαγές των παραπάνω μελετών .

Κριτήρια Επιλογής Ασθενών

Τα κριτήρια επιλογής των ασθενών σε μελέτες Φάσης I είναι τα ακόλουθα :

1. Να έχουν δοκιμάσει ανεπιτυχώς όλες τις γνωστές θεραπείες.
2. Να έχουν την κατάλληλη ηλικία σύμφωνα με προκαθορισμένα ιατρικά πρότυπα.
3. Αποκλείονται έγκυες γυναίκες.
4. Ο αναμενόμενος χρόνος ζωής να είναι μεγαλύτερος από 12 εβδομάδες έτσι ώστε να προλάβει να έχει επίδραση το φάρμακο.
5. Να υπάρχει κανονική λειτουργία οργάνων.
6. Να μην λαμβάνουν άλλη παράλληλη θεραπεία ή λήψη φαρμάκων.
7. Να υπάρχει έγγραφη συγκατάθεση του αρρώστου.

Στο σημείο αυτό πρέπει να αναφέρουμε ότι πριν ξεκινήσουν οι μελέτες Φάσης I, ο ασθενής πληροφορείται ότι σπάνια παρατηρείται βελτίωση της ιατρικής κατάστασης σε τέτοιες μελέτες. Για το λόγο αυτό ο ασθενής έχει το δικαίωμα να αποσυρθεί από την μελέτη όποια στιγμή επιθυμεί.

Ακόμη θα πρέπει να διευκρινίσουμε και κάποιους άλλους όρους που θα μας είναι χρήσιμοι στη συνέχεια.

- Ως **«Αρχική Δόση»** συνήθως καθορίζεται η δόση ενός σκευάσματος, που είναι συνήθως ίση με το 1/3 της ελάχιστης τοξικής δόσης (toxic low dose, TLD) έτσι όπως καθορίζεται σε προκλινικές μελέτες σε μεγάλα ζώα.
- Η TLD είναι η μικρότερη δόση που προκαλεί αιματολογική, χημική ή μορφολογική αλλαγή στον οργανισμό ενός ζώου ενώ θα πρέπει να αναφέρουμε ότι διπλασιασμός της δόσης αυτής δεν προκαλεί θάνατο, ενώ το πρόγραμμα χορήγησης γίνεται καθημερινά για 3-4 εβδομάδες.
- Ακόμη το συνηθισμένο σχήμα αύξησης δόσης για 2-12 δόσεις είναι το ακόλουθο: $\delta \rightarrow 2\delta \rightarrow 3.3\delta \rightarrow 5\delta \rightarrow 7\delta \rightarrow 9\delta \rightarrow 12\delta \rightarrow 16\delta$
- Ως **προτεινόμενη δόση για θεραπευτική αγωγή** καθορίζεται η δόση που προκαλεί μέτρια τοξικότητα που μπορεί να εξαλειφθεί.
- Ως **ελάχιστη τοξική δόση** ορίζεται η δόση στην οποία ένας ή παραπάνω από τους 3 ασθενείς δείχνουν συνεπή και εύκολα αφαιρούμενη τοξικότητα.

4.3.2 Μελέτες Φάσης II

Αφού καθορίσαμε τη ΜΑΔ, επόμενος στόχος είναι ο καθορισμός της βιολογικής επίδρασης της αγωγής και τον προσδιορισμό εμφάνισης παρενεργειών. Οι μελέτες Φάσης II, είναι μελέτες ασφάλειας και αποτελεσματικότητας των φαρμάκων/θεραπειών. Βασίζονται πολύ στη μελέτη φάσης I και εξετάζουν την αποτελεσματικότητα του φαρμάκου στο επίπεδο της ΜΑΔ της φάσης I. Έτσι, αν η ΜΑΔ είναι χαμηλή τότε η επίδραση του φαρμάκου είναι χαμηλή ή αργή, ενώ αν η ΜΑΔ είναι υψηλή, τότε παρουσιάζονται πολλές παρενέργειες. Για το λόγο αυτό ο σωστός καθορισμός της ΜΑΔ στη μελέτη φάσης I, είναι πολύ σημαντικός για την συνέχεια των δοκιμών.

ΠΑΡΑΔΕΙΓΜΑ: Μελέτες 2 σταδίων (2-stage trials).

Στο 2ο στάδιο του πειράματος προσπαθούμε να εντοπίσουμε μελέτες χωρίς ή με χαμηλή επίδραση θεραπείας, έτσι λέμε ότι αν $\theta \geq 20\%$ τότε προχωράμε, αλλιώς αν $\theta < 20\%$ τότε σταματάμε τη μελέτη. Συνήθως στις μελέτες Φάσης II, στο πρώτο στάδιο έχουμε συνήθως 14 ασθενείς. Στο 2ο στάδιο, έχουμε 10-20 ασθενείς επιπλέον των ασθενών του 1ου σταδίου και στο 2ο στάδιο πλέον ελέγχουμε αν το φάρμακο έχει αρκετά μεγάλη επίδραση. Έτσι ώστε να αντικαταστήσει ένα προυπάρχον σκεύασμα.

Πιθανά μέτρα δράσης των φαρμάκων

Στις διάφορες φάσεις των κλινικών δοκιμών υπάρχουν κάποια μέτρα σύμφωνα με τα οποία ελέγχουμε την δράση των φαρμάκων. Ένα πρώτο μέτρο είναι η αποτελεσματικότητα που μας δείχνει το επί τις % ποσοστό θετικής απόκρισης στο φάρμακο. Ένα άλλο μέτρο είναι οι παρενέργειες που όπως είδαμε παραπάνω εξαρτώνται από τον βαθμό τοξικότητας. Αν τα αποτελέσματα είναι θετικά τότε προχωρούμε στις μελέτες φάσης III, όπου ελέγχουμε ακόμα πιο αναλυτικά την αποτελεσματικότητα του φαρμάκου. Βέβαια θα πρέπει να αναφέρουμε ότι μια ψευδώς αρνητική μελέτη φάσης II μπορεί να είναι καταστρεπτική καθώς αποτρέπει τη χρήση ενός αποτελεσματικού φαρμάκου.

Ένα ακόμη μέτρο είναι ο στατιστικός έλεγχος στο δεύτερο στάδιο της υπόθεσης:

$$H_0 : \theta > \theta_0 \text{ (μη αποτελεσματικό φάρμακο)}$$

έναντι της εναλλακτικής

$$H_1 : \theta \leq \theta_0 \text{ (αποτελεσματικό φάρμακο)}.$$

Φυσικά μας ενδιαφέρει να απορρίψουμε την H_0 , δηλαδή την αναποτελεσματικότητα του φαρμάκου. Στο έλεγχο αυτό, πολύ σημαντικό ρόλο παίζει το μέγεθος του δείγματος, μιας και μεγάλο δείγμα κοστίζει, ενώ μικρό δείγμα έχει πολύ μικρή ισχύ. Λύση στο πρόβλημα αυτό δίνουν οι μελέτες πολλών σταδίων (ακολουθιακή έγκριση της αποτελεσματικότητας ενός

φαρμάκου και σταδιακή εισαγωγή ασθενών), όπου αυξάνουν την ισχύ του ελέγχου στατιστικής σημαντικότητας ενώ το κόστος παραμένει σε λογικά όρια (μειώνει το κόστος στην περίπτωση μη αποτελεσματικών φαρμάκων).

4.3.3 Μελέτες Φάσης III.

Στην ουσία εδώ ελέγχεται πραγματικά η αποτελεσματικότητα του φαρμάκου/θεραπείας. Αποτελεί επέκταση και συνέχεια της φάσης II. Συνήθως οι Βιοστατιστικοί διαχειρίζονται δεδομένα φάσης III. Όταν λέμε απλά κλινική δοκιμή αναφερόμαστε στις μελέτες φάσης III.

4.3.4 Μελέτες Φάσης IV.

Όταν το φάρμακο βρεθεί αποτελεσματικό με βάση τη μελέτη φάσης III τότε τα αποτελέσματα ελέγχονται από ένα κρατικό οργανισμό και παρέχεται η άδεια κυκλοφορίας του φαρμάκου. Οι μελέτες φάσης IV αφορούν την παρακολούθηση (screening) του φαρμάκου μετά την κυκλοφορία. Εδώ σκοπός είναι ο εντοπισμός παρενεργειών ή προβλημάτων που δεν μπορούσαν να εντοπιστούν στη μελέτη φάσης III, για παράδειγμα τι γίνεται μετά από χρόνια χρήση του φαρμάκου. Επίσης εδώ περιλαμβάνονται και μελέτες έρευνας αγοράς.

4.4 Τύποι Κλινικών Δοκιμών

Τρεις είναι οι βασικότεροι τύποι κλινικών δοκιμών. Ο πρώτος τύπος είναι η **κλινική ή θεραπευτική δοκιμή** στην οποία συμμετέχουν ασθενείς και οι αγωγές είναι θεραπευτικές. Στόχος αυτού του τύπου δοκιμών είναι η θεραπεία της νόσου, ενώ το μέγεθος δείγματος είναι 20 ως 100 άτομα.

Ο δεύτερος τύπος είναι οι **κύριες μελέτες πρόληψης (primary prevention studies)** στις οποίες συμμετέχουν υγιείς, οι αγωγές περιλαμβάνουν μέτρα πρόληψης – προφύλαξης και στόχος είναι η αποφυγή εμφάνισης της νόσου σε υγιείς.

Ο τρίτος τύπος κλινικών δοκιμών είναι οι **δευτερεύουσες μελέτες πρόληψης (secondary prevention studies)**. Στις δευτερεύουσες μελέτες πρόληψης συμμετέχουν ασθενείς (συνήθως με ένα επεισόδιο της νόσου), οι αγωγές περιλαμβάνουν μέτρα πρόληψης – προφύλαξης και στόχος είναι η αποφυγή εμφάνισης 2ου επεισοδίου της νόσου ή χειροτέρευση της κατάστασης. Ένα χαρακτηριστικό παράδειγμα αυτού του τύπου κλινικών δοκιμών είναι η χημειοθεραπεία σε καρκινοπαθείς.

Εδώ θα πρέπει να αναφέρουμε ότι στις Μελέτες Πρόληψης, απαιτείται μεγάλο μέγεθος

δείγματος. Πιο συγκεκριμένα στις Κύριες μελέτες πρόληψης απαιτείται δείγμα 10000 ατόμων ενώ στις δευτερεύουσες απαιτείται δείγμα 1000 ατόμων.

4.5 Τυχαιοποίηση

Τυχαιοποίηση (Randomization - random allocation), ονομάζουμε τη διαδικασία τυχαίας ανάθεσης αγωγών για κάθε άτομο. Στόχος της τυχαιοποίησης είναι η εξάλειψη μεροληψίας επιλογής (selection bias), δηλαδή η κατανομή των ασθενών να γίνει σύμφωνα με τις εκτιμήσεις ή προτιμήσεις του Γιατρού. Οι δύο ομάδες θα πρέπει να έχουν παρόμοια χαρακτηριστικά αλλά και παρόμοιο μέσο επίπεδο κατάστασης νόσου. Μερικές μεροληπτικές συγκρίσεις φαίνονται στο παρακάτω διάγραμμα.

Μερικές μεροληπτικές συγκρίσεις:

- | | | | |
|----|--|--------|---|
| 1. | Εγχειρισμένοι
(Βαριές περιπτώσεις) | έναντι | φαρμακευτικής αγωγής,
(ελαφριές περιπτώσεις) |
| 2. | Εθελοντές
(Βαριές περιπτώσεις)
(νέο φάρμακο) | έναντι | μη εθελοντών
(ελαφριές περιπτώσεις)
(παλιό φάρμακο) |

4.6 Λεπτομέρειες Μελετών Φάσης III

4.6.1 Προσχεδιασμός της μελέτης

Κάθε κλινική δοκιμή Φάσης III είναι σχεδιασμένη προσεκτικά και όλο το πλάνο της είναι προγραμματισμένο και γραμμένο σε ένα «**ΠΡΩΤΟΚΟΛΛΟ**». Η έγκριση για τη μελέτη γίνεται μέσω του πρωτοκόλλου. Το Πρωτόκολλο περιέχει πληροφορίες όσον αφορά στο ιατρικό υπόβαθρο (ασθένεια-θεραπευτικές αγωγές), την ανάθεση και την τυχαιοποίηση των αγωγών, τα κριτήρια επιλογής των ασθενών, τον αριθμό ασθενών (με προκαθορισμένη ισχύ και σημαντικότητα) και τη διάρκεια της μελέτης.

Το περίγραμμα ενός Πρωτοκόλλου σύμφωνα με τους Φριεδμαν *et al.* (1998), αποτελείται από τα ακόλουθα βασικά στοιχεία :

- Το Υπόβαθρο ή η Βιβλιογραφική ανασκόπηση Μελέτης (background).
- Τους στόχους της μελέτης (objectives).
- Το σχεδιασμό της Μελέτης (design).

- Την οργάνωση της Μελέτης (organization).
- Τα Παραρτήματα.

Αναλύοντας το κάθε ένα στοιχείο ξεχωριστά βλέπουμε ότι, στο **Υπόβαθρο/Βιβλιογραφική ανασκόπηση Μελέτης (background)** αναφέρονται οι προηγούμενες μελέτες-επιστημονικές εργασίες-διατριβές στο χώρο που εξετάζει η μελέτη. Ακόμη ορίζεται η νόσος, οι πιο πιθανοί παράγοντες κινδύνου, η εξέλιξη θεραπευτικών αγωγών και η επεξήγηση της καινούριας θεραπείας.

Στους **στόχους της μελέτης** (objectives) αναφέρονται :

- το κύριο ερώτημα και τη μεταβλητή απόκρισης,
- τα δευτερεύοντα ερωτήματα και τις μεταβλητές απόκρισης,
- τις υποθέσεις για τις υπό-ομάδες,
- τις δυσμενείς επιδράσεις (adverse effects)- παρενέργειες (side effects).

Στο **Σχεδιασμό Μελέτης** (design) περιγράφονται αρχικά ο υπό μελέτη Πληθυσμός (study or target population) και πιο συγκεκριμένα τα κριτήρια συνυπολογισμού στον πληθυσμού (inclusion criteria) και τα κριτήρια αποκλεισμού από τον πληθυσμό (exclusion criteria). Δεύτερον περιγράφεται το μέγεθος του δείγματος, οι προϋποθέσεις και οι εκτιμητές, ενώ γίνεται και εγγραφή των συμμετεχόντων στη μελέτη (enrolment of participants). Πιο συγκεκριμένα αναφέρεται η γραπτή συγκατάθεση των συμμετεχόντων (informed consent), η αποτίμηση της καταλληλότητας συμμετοχής (assessment of eligibility), η αρχική εξέταση (baseline examination) και η μέθοδος ανάθεσης αγωγών (τυχαιοποίηση). Τρίτον περιγράφονται οι μέθοδοι Παρέμβασης - αγωγές (Intervention) που είναι δύο

α) Περιγραφή αγωγής - Χρονοδιάγραμμα λήψης αγωγών (description and schedule)

β) Μέτρα συμμόρφωσης (measures of compliance)

Πέμπτον, γίνεται περιγραφή των επισκέψεων παρακολούθησης/καταγραφής και του χρονοδιάγραμματος (follow up description+schedule). Έκτον γίνεται καταγραφή των Μεταβλητών απόκρισης (ascertainment of response variables),

α) Εκπαίδευση (training), σύμφωνα με την οποία γίνονται κάποιες ερωτήσεις της μορφής: «Είχατε πόνο στο στήθος τους 3 τελευταίους μήνες;» ή «Δεν είχατε πόνο τους 3 τελευταίους μήνες, έτσι δεν είναι;».

β) Συλλογή δεδομένων (data collection)

γ) Ποιοτικός έλεγχος (quality monitoring/control). Πιο συγκεκριμένα

** Έλεγχος Ερωτηματολογίων, σύμφωνα με τον οποίο γίνεται σύγκριση για πλήρη συμπλήρωση και σύγκριση με τα προηγούμενα ερωτηματολόγια απο τον γρουπ συπερισσορ, αλλά και έλεγχος με ιατρικά δεδομένα.

** Έλεγχος Διαδικασιών, κατά τον οποίο γίνεται έλεγχος ακραίων βιολογικών τιμών, γίνεται έλεγχος συστηματικών μεροληψιών από τα άτομα που συλλέγουν τα δεδομένα, αλλά και έλεγχος εξοπλισμού. Εδώ θα πρέπει να αναφέρουμε η συχνότερη μεροληψία που εμφανίζεται είναι αυτή της μεροληψίας μέτρησης. Για την λύση της παίρνουμε μετρήσεις από όλους τους νοσοκόμους (data collectors).

** Έλεγχος Θεραπευτικών Αγωγών, όπου εδώ γίνεται έλεγχος ετικετών στα φάρμακα, έλεγχος στις δόσεις και έλεγχος αν οι ασθενείς παίρνουν σωστά τα φάρμακά τους.

Έβδομον, γίνεται περιγραφή της ανάλυσης των Δεδομένων

α) Προσωρινή ανάλυση (Interim analysis)

β) Τελική ανάλυση

Ενώ τελευταία αναφέρεται η ορολογία.

Στην **οργάνωση μελέτης** (organization), αναφέρονται οι ερευνητές που συμμετέχουν (participating investigators) στην μελέτη που είναι οι

α) Μονάδες στατιστικής ανάλυσης/ κέντρα διαχείρισης δεδομένων (statistical unit/data coordinating center)

β) Εργαστήρια και άλλες ειδικές μονάδες (labs and other special units)

γ) Κλινικές

αλλά γίνεται αναφορά και στην Διοίκηση της Μελέτης και πιο συγκεκριμένα στις Οργανωτικές επιτροπές και υπό-επιτροπές, στην Επιτροπή παρακολούθησης δεδομένων και την Οργάνωση χρηματοδότησης.

Τέλος στα Παραρτήματα γίνεται αναφορά στον ορισμό των κριτηρίων καταλληλότητας συμμετοχής στη μελέτη και στον ορισμό των μεταβλητών απόκρισης.

4.6.2 Ορισμός Ασθενών

Αν και γενικά ο ορισμός ενός ασθενή φαίνεται εύκολος στην πραγματικότητα δεν είναι. Ο ορισμός του ασθενή πρέπει να γίνει με ακρίβεια και λεπτομέρεια. Επίσης ο ορισμός των ασθενών περιλαμβάνει και τον καθορισμό μερικών υπο-ομάδων όπως για παράδειγμα αν θα αναφερθούμε μόνο σε γυναίκες ή άνδρες, μιας και θα πρέπει να περιορίσουμε τη μελέτη σε συγκεκριμένες ομάδες με γνώμονα την ηλικία ή το φύλο. Στο κεφάλαιο αυτό ορίζονται με σαφήνεια τα inclusion και exclusion κριτήρια του υπο-εξέταση πληθυσμού.

4.6.3 Ορισμός Θεραπευτικών αγωγών

Οι Θεραπευτικές Αγωγές περιλαμβάνουν το χρονοδιάγραμμα και το μέγεθος της δόσης, ακόμη δίνουν λύσεις στο πρόβλημα, τι γίνεται αν χαθεί μια δόση και ορίζουν το placebo - ψευδοφάρμακο /αδρανούς θεραπείας. Εδώ θα πρέπει να αναφέρουμε ότι έχουν παρατηρηθεί διάφορες αλλαγές στις μετρήσεις μεταβλητών απόκρισης στις ομάδες ασθενών που λαμβάνουν ψευδοφάρμακό/αδρανή θεραπεία σε σχέση με ασθενείς που δε λαμβάνουν καθόλου θεραπεία (*επίδραση ψευδοφάρμακου*), όπως για παράδειγμα στις μετρήσεις πίεσης, αιτία βέβαια για το φαινόμενο αυτό μπορεί να είναι και οι ψυχολογικοί λόγοι, που προκαλούν αλλαγές στην πίεση (*επίδραση Placebo*). Ακόμη γίνεται αναφορά στις **επιδράσεις Hawthorne** σύμφωνα με τις οποίες οι ασθενείς που λαμβάνουν την ίδια θεραπεία, έχουν καλύτερη θεραπευτική πορεία στα πλαίσια κλινικών δοκιμών σε σχέση με τους ασθενείς που λαμβάνουν την ίδια θεραπεία αλλά βρίσκονται εκτός κλινικών δοκιμών, εκτός μελέτης. Οι λόγοι που προκαλούν το παρακάτω φαινόμενο είναι οι ακόλουθοι

- Ψυχολογικοί
- Καλύτερη παρακολούθηση
- Γιατροί και ασθενείς πιο ενθουσιώδεις και αισιόδοξοι
- Γιατροί πιο πρόθυμοι να προσφέρουν βοήθεια
- Ασθενείς πιο πρόθυμοι να ακολουθήσουν τη θεραπεία

Εδώ θα πρέπει να αναφέρουμε ότι η δόση και η αγωγή μπορεί να αλλάξει ανάλογα με τις κατάσταση του ασθενή και τις παρενέργειες (ο τρόπος όμως πρέπει να αναφέρεται στο πρωτόκολλο).

4.6.4 Εκτίμηση και καθορισμός μεταβλητών απόκρισης

Κατά την εκτίμηση και τον καθορισμό των μεταβλητών απόκρισης, συνήθως μετράμε μία κύρια - κεντρική μεταβλητή και πολλές δευτερεύουσες, όπως:

- Συμπτώματα που αναφέρονται από ασθενείς
- Σημεία/συμπτώματα που καταγράφονται από Γιατρούς
- Βιοχημικές μετρήσεις ή μετρήσεις Ιατρικών ελέγχων
- Ερωτηματολόγια πόνου

Όμως **προσοχή** η μεροληψία μπορεί να προκύψει τόσο από τους ασθενείς όσο και από τους γιατρούς αν γνωρίζουν τις θεραπείες π.χ. η επίδραση του ψευδοφαρμάκου. Για το λόγο αυτό έχουμε:

- *Μονή τύφλωση/άγνοια ασθενών (single blinded)*: Οι ασθενείς δε γνωρίζουν τη θεραπεία που λαμβάνουν
- *Διπλή τύφλωση/άγνοια ασθενών-Ερευνητών (double blinded)*: Τόσο οι ασθενείς όσο και οι ερευνητές (Ιατροί/νοσοκόμοι) δε γνωρίζουν τη θεραπεία που λαμβάνει ο κάθε ασθενής

Η περίπτωση αυτή όμως δεν είναι πάντα εφικτή, καθώς μερικές θεραπείες έχουν συγκεκριμένες παρενέργειες από τις οποίες ο ερευνητής καταλαβαίνει και τη θεραπεία που λαμβάνει ο κάθε ασθενής.

4.6.5 Εισαγωγή στη Μελέτη και Ανάθεση Αγωγών

Οι ασθενείς που πληρούν τα κριτήρια του Πρωτοκόλλου επιλέγονται ως δείγμα, γίνονται δεκτοί επίσημα με την καταγραφή του ονόματός τους στον κατάλογο/ημερολόγιο της μελέτης και την ανάθεση τους σε ένα σειριακό κωδικό (π.χ. AA ή ID). Αμέσως μετά γίνεται η τυχαία ανάθεση τους σε μία θεραπεία. Η ανάθεση δεν είναι γνωστή/προκαθορισμένη πριν την εισαγωγή του ασθενή για την αποφυγή μεροληψίας επιλογής (selection bias). Συνήθως χρησιμοποιούνται:

- φάκελοι με σειριακούς κωδικούς όπου αναφέρεται η θεραπεία
- τηλεφωνικά κέντρα που αναθέτουν θεραπείες σύμφωνα με τυχαίους αριθμούς (χρήσιμο σε πολυκεντρικές μελέτες)
- στις πιο σύγχρονες μελέτες το διαδίκτυο

Αποκλίσεις από το Πρωτόκολλο

Αποκλίσεις από το Πρωτόκολλο αποτελούν οι αποκλεισμοί/αποσύρσεις ασθενών που μπορεί να προκύψουν λόγω κακής απόκρισης στη θεραπεία ή αποκάλυψης κάποιου λόγου αποκλεισμού του ασθενή ή από πολλές αποσύρσεις που εξασθενούν την αποτελεσματικότητα της τυχαιοποίησης και έτσι μπορεί να καταλήξουμε σε μεροληπτικές συγκρίσεις.

Οι αποκλίσεις από το πρωτόκολλο είναι ένα σημαντικό πρόβλημα για τις κλινικές δοκιμές μιας και η θεραπεία θα πρέπει να ακολουθείται ακριβώς όπως περιγράφεται στο πρωτόκολλο, ενώ οι ασθενείς που είναι επιρρεπείς σε αποκλίσεις από το πρωτόκολλο (π.χ. παιδιά και ηλικιωμένοι) θα πρέπει να αποκλείονται από τη μελέτη ή να παρακολουθούνται στενά (π.χ. από νοσοκόμες). Οι αποκλίσεις συχνά δεν αποφεύγονται λόγω της αλλαγής θεραπείας αλλά και εξαιτίας της εθελοντικής απόσυρσης ασθενών από τη μελέτη ή αλλαγή γιατρού/τόπου κατοικίας.

Πολλές φορές συμπεριλαμβάνουμε στη μελέτη ασθενείς με ελλιπή δεδομένα για να μη χάσουμε ισχύ και γιατί η συμπεριφορά των ασθενών αυτών μπορεί να είναι τυπική για συγκεκριμένες αγωγές. Ακόμη όταν έχουμε αγνοούμενες τιμές τότε :

- στην ανάλυση επιβίωσης, χειριζόμαστε τα δεδομένα ως λογοκριμένα
- σε άλλες περιπτώσεις απλά αφαιρούμε ή αντικαθιστούμε με διάμεσο ή μέσο ή αλλιώς χρησιμοποιούμε άλλες πιο προχωρημένες τεχνικές.

4.7 Άλλες λεπτομέρειες Κλινικών Δοκιμών

4.7.1 Μέτα-ανάλυση

Ως μέτα-ανάλυση ορίζονται οι στατιστικές τεχνικές που συνδυάζουν αποτελέσματα πολλών μελετών για να εξάγουν ένα ενιαίο συμπέρασμα. Υπάρχουν όμως διάφορα προβλήματα όπως οι διαφορετικοί σχεδιασμοί μελετών, οι διαφορετικοί πληθυσμοί, οι ασύμβατες αγωγές και τέλος η μεροληψία δημοσίευσης (publication bias).

Μερικές φορές στόχος δε είναι να αποδείξουμε ότι ένα φάρμακο είναι καλύτερο από ένα άλλο αλλά ότι είναι ισοδύναμο του (αρά ότι μπορεί να το υποκαταστήσει χωρίς ιδιαίτερα προβλήματα ή απώλεια αποτελεσματικότητας). Οι Μελέτες βιο-ισοδυναμίας εμφανίζονται κυρίως σε μελέτες φάσης I.

Έστω η διαφορά μεταξύ των δύο θεραπειών μετριέται με μια παράμετρο θ . Τότε μας ενδιαφέρει να ελέγχσουμε την υπόθεση

$$H_0 : \theta = 0 \text{ (ισοδυναμία).}$$

Στον παραπάνω έλεγχο πρέπει να ορίσουμε όρια $\theta_L < 0$ και $\theta_U > 0$ τα οποία ονομάζονται όρια ανεκτικότητας/ισοδυναμίας έτσι ώστε αν $\theta_L < \theta < \theta_U$ τότε να συμπεραίνουμε ισοδυναμία των δύο θεραπειών. Για τον έλεγχο αυτό λοιπόν έχουμε τις ακόλουθες δύο μεθόδους - προσεγγίσεις:

- **1η Μέθοδος – Προσέγγιση:**

1. $H_0^A : \theta > \theta_L$ έναντι $H_1^A : \theta \leq \theta_L$
2. $H_0^B : \theta < \theta_U$ έναντι $H_1^B : \theta \geq \theta_U$.
3. Αν δεν απορρίψουμε H_0^A και H_0^B τότε έχουμε βιο-ισοδυναμία.

- **2η Μέθοδος – Προσέγγιση:**

1. Βρίσκουμε το $(1-2\alpha)\%$ επίπεδο εμπιστοσύνης (θ_1, θ_2) .
2. Καταλήγουμε σε βιο-ισοδυναμία αν $\theta_L < (\theta_1, \theta_2) < \theta_U$ με α ίσο με το επίπεδο εμπιστοσύνης μονο-κατάληκτων ελέγχων της 1ης προσέγγισης.

4.8 Πειραματικοί Σχεδιασμοί

4.8.1 Απλή τυχαιοποίηση

Τυχαιοποίηση χωρίς περιορισμούς (unrestricted randomization), σύμφωνα με την οποία αναθέτουμε σε κάθε ασθενή τυχαία την αγωγή (ομάδα θεραπείας). Όμως υπάρχει ένα πρόβλημα, μπορεί να καταλήξουμε σε άνισο αριθμό ατόμων ανά ομάδα θεραπείας πράγμα που σημαίνει μειωμένη ισχύ ελέγχων.

Τυχαιοποίηση με περιορισμούς (blocked/ restricted randomization), όπου σε κάθε κ ασθενείς θέτουμε τον περιορισμό ότι ο αριθμός των ατόμων ανά ομάδα πρέπει να είναι ίσος με $\kappa/2$. Υπάρχει όμως ένα σοβαρό μειονέκτημα. Δεν μπορεί να είναι ολικώς άγνωστες οι αναθέσεις. Για την λύση του προβλήματος μπορεί να γίνει μεταβολή των block (δηλαδή του κ) ή επιλογή κ ασθενών και μετά συνολική ανάθεση.

ΠΑΡΑΔΕΙΓΜΑ

Έστω ότι έχουμε δύο ομάδες ασθενών A, B και για $\kappa=2$ οι αναθέσεις θα είναι 2: AB, BA, ενώ για $\kappa=4$ οι αναθέσεις θα είναι 6: AABB, ABAB, ABBA, BAAB, BABA, BBAA

ΑΝΑΛΥΣΗ

t-test, ANOVA, χ^2 test

4.8.2 Στρωματοποιημένη τυχαιοποίηση

Στην στρωματοποιημένη τυχαιοποίηση, συνήθως μας ενδιαφέρει να έχουμε τον έλεγχο όσον αφορά γνωστούς ή σημαντικούς συγχυτικούς παράγοντες. Έτσι όταν ελέγχουμε μια μελέτη ως προς τους συγχυτικούς παράγοντες σημαίνει ότι εξαλείφουμε την επίδραση τους (συνεπώς οι ομάδες θεραπείας δεν πρέπει να διαφέρουν ως προς τους συγχυτικούς παράγοντες). Ύστερα χωρίζουμε τους ασθενείς σε ομάδες ανάλογα με τους συγχυτικούς παράγοντες και γίνεται χωριστά η τυχαιοποίηση σε κάθε μια από τις ομάδες αυτές. Το πλεονέκτημα της στρωματοποιημένης τυχαιοποίησης είναι ότι εξασφαλίζει την ισόρροπη εκπροσώπηση όλων των ομάδων που μας ενδιαφέρουν να ελέγξουμε στο δείγμα

Η στρωματοποιημένη τυχαιοποίηση χρησιμοποιείται σε πολυκεντρικές μελέτες όπου τα πολλά στρώματα θα πρέπει να αποφεύγονται. Ισορροπία στο μέγεθος κάθε ομάδας επιτυγχάνεται μόνο με περιορισμένη τυχαιοποίηση μέσα σε κάθε στρώμα για το λόγο αυτό συνήθως γίνεται Πολυπαραγοντική ανάλυση διακύμανσης.

Τέλος παρακάτω φαίνονται επιγραμματικά οι διαφορές μεταξύ Τυχαίας Δειγματοληψίας και Τυχαιοποίησης με σκοπό την καλύτερη εμπέδωση των δύο αυτών εννοιών.

Τυχαία Δειγματοληψία

- Επιλέγουμε δείγματα αντιπροσωπευτικά του συνολικού πληθυσμού
- Εκτιμούμε χαρακτηριστικά του πληθυσμού

Τυχαιοποίηση

- χωρίζουμε το δείγμα σε ομάδες με βάση την τύχη έτσι ώστε οι δύο ομάδες να είναι ίδιες όσον αφορά τα χαρακτηριστικά τους

Ανάλυση : Multiway Analysis of Variance

Κεφάλαιο 5

Αδρανοποίηση των Επιδράσεων Συγχυτικών Παραγόντων

5.1 Συγχυτικοί Παράγοντες

Ορισμός 5.1 Συγχυτικός Παράγοντας (*Confounding factor*) είναι μία μεταβλητή που σχετίζεται και με την ασθένεια αλλιώς και με τον παράγοντα κινδύνου που εξετάζουμε. Οι συγχυτικοί παράγοντες πρέπει να λαμβάνονται υπόψη πριν να κοιτάξουμε τη σχέση ασθένειας - παράγοντα κινδύνου.

Η αδρανοποίηση (ή ο έλεγχος) ενός συγχυτικού παράγοντα μπορεί να γίνει είτε μέσω του σχεδιασμού του πειράματος με την εξομοίωση (*matching*) είτε με την τυποποίηση ή προτυποποίηση (*standardization*) των αποτελεσμάτων ως προς το συγχυτικό παράγοντα είτε με ειδική στατιστική ανάλυση. Η αδρανοποίηση των επιδράσεων ενός συγχυτικού παράγοντα ονομάζεται έλεγχος του συγχυτικού παράγοντα (*controlling for a confounding factor*) και τα αποτελέσματα διορθωμένα ως προς τον συγχυτικό παράγοντα (*confounder adjusted results*).

Ορισμός 5.2 Η μελέτη της σχέσης ασθένειας - παράγοντα κινδύνου για διαφορετικές ομάδες δεδομένων (συνήθως επίπεδα ενός ή περισσότερων συγχυτικών παραγόντων) ονομάζεται **στρωματοποίηση ή στρωματοποιημένη ανάλυση** (*stratification ή stratified analysis*).

Ορισμός 5.3 Θετικός Συγχυτικός Παράγοντας (*Positive confounder*) είναι ένας παράγοντας θετικά (ή αρνητικά) συσχετισμένος και με τη νόσο και με τον παράγοντα κινδύνου.

Ορισμός 5.4 Αρνητικός Συγχυτικός Παράγοντας (*Negative confounder*) είναι ένας παράγοντας που ικανοποιεί μία από τις ακόλουθες συνθήκες:

1. Είναι θετικά συσχετισμένος με τη νόσο και αρνητικά με τον παράγοντα κινδύνου.
2. Είναι αρνητικά συσχετισμένος με τη νόσο και θετικά με τον παράγοντα κινδύνου.

Ορισμός 5.5 Ένας συγχυτικός παράγοντας ονομάζεται **αιτιολογική δίοδος** (causal pathway) αν ο παράγοντας κινδύνου επηρεάζει το συγχυτικό παράγοντα και αυτός με τη σειρά του τη νόσο.

Παράδειγμα 5.1 Ας υποθέσουμε ότι θέλουμε να μελετήσουμε τη σχέση μεταξύ καρκίνου του πνεύμονα και υψηλής κατανάλωσης αλκοόλ (ορισμένης ως μεγαλύτερης ή ίσης των δύο ποτών ημερησίως). Τα αποτελέσματα μιας προοπτικής μελέτης δίνονται στον Πίνακα 5.1.

X: Καταν.Αλκοόλ	Y: Καρκίνος του Πνεύμονα		Περ.Κατ. X
	1: Ναι	2: Όχι	
1: ≥ 2 ποτών	33	1667	1700
2: < 2 ποτών	27	2273	2300
Περ.Κατ. Y	60	3940	4000

Πίνακας 5.1: Δεδομένα Καρκίνου του Πνεύμονα Παραδείγματος 5.1 (Rosner, 1994, σελ. 399–400).

Εφόσον πρόκειται για σπάνια ασθένεια μπορούμε να χρησιμοποιήσουμε το ΛΣΠ ως μέτρο σχετικού κινδύνου. Ο ΛΣΠ είναι ίσος με 1.667 στατιστικά σημαντικός σε επίπεδο σημαντικότητας $\alpha = 10\%$ (ακριδής χ^2 ελέγχος $p = 0.065$).

Στην παραπάνω ανάλυση - σχέση όμως δε λάβαμε υπόψη μας ένα πολύ ισχυρό συγχυτικό παράγοντα που δεν είναι άλλος από το κάπνισμα του ίδιου του ατόμου που εξετάζουμε. Έτσι λοιπόν η σωστή προσέγγιση είναι να κάνουμε στρωματοποιημένη ανάλυση ανάλογα με το αν

X: Καταν.Αλκοόλ	Z: Καπνιστές (1)			Z: Μη Καπνιστές (2)		
	Y: Καρκ. Πνεύμ.		Περ.Κατ. X	Y: Καρκ. Πνεύμ.		Περ.Κατ. X
	1: Ναι	2: Όχι		1: Ναι	2: Όχι	
1: ≥ 2 ποτών	24	776	800	9	891	900
2: < 2 ποτών	6	194	200	21	2079	3000
Περ.Κατ. Y	30	970	1000	30	2970	3000

Πίνακας 5.2: Στρωματοποιημένα ως προς το Κάπνισμα δεδομένα Καρκίνου του Πνεύμονα Παραδείγματος 5.1 (Rosner, 1994, σελ. 399–400).

κάποιος καπνίζει ή όχι. Έτσι έχουμε τον στρωματοποιημένο πίνακα 5.2. Αν υπολογίσουμε το ΛΣΠ στους δύο πίνακες θα τους βρούμε ίσους με ένα και συνεπώς δεν υπάρχει σχέση μεταξύ των δύο μεταβλητών που μας ενδιαφέρουν δηλαδή της υψηλής κατανάλωσης αλκοόλ και καρκίνου του πνεύμονα.

Παράδειγμα 5.2 Στο παράδειγμα αυτό θα χρησιμοποιήσουμε δεδομένα από την μελέτη των *Shapiro et. al (1979)* που δημοσιεύτηκε στο *Lancet*. Στη μελέτη αυτή εξετάσθηκε η σχέση αντιουβηλικτικού χαπιού και καρδιακής ανεπάρκειας υπό την ύπαρξη ενός ισχυρού συγχυτικού παράγοντα: της ηλικίας. Για το λόγο αυτό, η ηλικία ομαδοποιήθηκε σε πέντε ομάδες/τάξεις: 25-29, 30-34, 35-39, 40-44, 45-49. Τα δεδομένα δίνονται στον Πίνακα 5.3 που ακολουθεί.

Ηλικία	Χρήστης Αντιουλ.	Περίπτωσης (cases)	Μάρτυρες (controls)	ΛΣΠ (OR)	Χρήστες Αντιουλ.(%)	Ποσοστό Ασθενών
25-29	Ναι	4	62	7.2	23	2
	Όχι	2	224			
30-34	Ναι	9	33	8.9	9	5
	Όχι	12	390			
35-39	Ναι	4	26	1.5	8	9
	Όχι	33	330			
40-44	Ναι	6	9	3.7	3	16
	Όχι	65	362			
45-49	Ναι	6	5	3.9	3	25
	Όχι	93	301			
Σύνολο	Ναι	29	135	1.7	8	12
	Όχι	205	1607			

Πίνακας 5.3: Δεδομένα Καρδιακής Ανεπάρκειας Παραδείγματος 5.2.

Όπως είναι προφανές από τον Πίνακα 5.3, αν λάβουμε υπόψη μας μόνο τον αθροιστικό πίνακα αγνοώντας την ηλικία τότε βρίσκουμε $\text{ΛΣΠ}=1.7$. Αντίθετα, αν υπολογίσουμε τους ΛΣΠ για διαφορετικές ηλικιακές ομάδες ο κίνδυνος φαίνεται πολύ μεγαλύτερος κυρίως για τις μικρές ηλικίες.

Από τα παραπάνω προκύπτει ότι η ηλικία είναι αρνητικός συγχυτικός παράγοντας διότι

- Οι πιο ηλικιωμένες γυναίκες χρησιμοποιούν με μικρότερη συχνότητα αντιουβηλικτικά χάπια σε σχέση με τις νεότερες γυναίκες.

- Οι πιο ηλικιωμένες γυναίκες έχουν μεγαλύτερα ποσοστά εμφάνισης της νόσου σε σχέση με τις νεότερες γυναίκες.

5.1.1 Τρόποι Αδρανοποίησης των Επιδράσεων των Συγχυτικών Παραγόντων

Υπάρχουν διάφοροι τρόποι για να αδρανοποιήσουμε τη επίδραση των συγχυτικών παραγόντων. Μπορούν να χωριστούν σε δύο μεγάλες κατηγορίες

1. Αδρανοποίηση μέσω του σχεδιασμού της μελέτη (π.χ. με εξομοίωση) και της ανάλυσης εξαρτημένων δειγμάτων
2. Αδρανοποίηση μέσω ανάλυσης.

Όσον αφορά την πρώτη προσέγγιση οι δύο υπό σύγκριση ομάδες επιλέγονται έτσι ώστε να έχουν τα ίδια χαρακτηριστικά όσον αφορά τους συγχυτικούς παράγοντες και μετά εφαρμόζονται μέθοδοι για εξαρτημένα δείγματα. Στις περισσότερες περιπτώσεις όμως η εξομοίωση δεν είναι εφικτή και για το λόγο αυτό γίνεται διορθωμένη ανάλυση μέσω

1. απλής στρωματοποιημένης ανάλυσης
2. προτυποποίησης
3. ειδικές μεθόδους όπως οι μέθοδοι Mantel-Haenszel
4. με τη χρήση παλινδρομικών μοντέλων

5.2 Ανάλυση σε Μελέτες με Εξομοιωμένα/Εξαρτημένα Δείγματα Δίτιμων Μεταβλητών

5.2.1 Έλεγχος Αναλογιών για Δύο Εξαρτημένα Δείγματα Δίτιμων Κατηγορικών Μεταβλητών

Στην ενότητα 3.6 είδαμε πως ελέγχουμε της υπόθεση της σχέσης μεταξύ δίτιμων μεταβλητών για δύο ανεξάρτητα δείγματα. Εδώ θα δούμε πως ελέγχουμε τη σχέση δυο εξαρτημένων δειγμάτων (ή ζευγαρωτών τιμών). Τέτοια δείγματα εμφανίζονται συχνά στην Ιατρική έρευνα λόγω τη εξομοίωσης (matching) των ασθενών και των μαρτύρων σύμφωνα με κάποιους σημαντικούς συγχυτικούς παράγοντες (συνήθως ηλικία και φύλο).

Παράδειγμα 5.3 Έστω ότι θέλουμε να συγκρίνουμε δύο χημειοθεραπείες που σκοπό έχουν την αντιμετώπιση του καρκίνου του μαστού στις γυναίκες. Οι δύο ομάδες θεραπείας πρέπει να είναι όσο το δυνατόν πιο όμοιες όσον αφορά τους παράγοντες κινδύνου. Για να επιτευχθεί αυτό εξομοιώνουμε (ταιρίαζουμε) τις δύο ομάδες ανάλογα με την ηλικία (σε ομάδες εύρους 5 ετών) και την κλινική κατάσταση. Σε κάθε ασθενή της πρώτης ομάδας αντιστοιχεί ένας ασθενής της δεύτερης ομάδας με την ίδια ηλικιακή ομάδα και κλινική κατάσταση. Οι ασθενείς παρακολουθούνται για 5 έτη και ως μεταβλητή απόκρισης χρησιμοποιείται η επιβίωσή τους στο τέλος των 5 ετών (ΝΑΙ/ΟΧΙ). Αν δε λάβουμε υπόψη την εξομοίωση τότε θα προκύψει ο ακόλουθος πίνακας:

X: Θεραπεία	Επιβίωση στα 5 έτη		Περ.Κατ. X
	1: Ναι	2: Όχι	
1: A	526	95	621
2: B	515	106	621
Περ.Κατ. Y	1041	201	1242

Πίνακας 5.4: Δεδομένα Καρκίνου του Μαστού· Παράδειγμα 5.3 (Rosner, 1994, σελ. 377).

Από τον παραπάνω πίνακα έχουμε μη στατιστικά σημαντική σχέση (σε επίπεδο σημαντικότητας 5%) μεταξύ επιβίωσης και θεραπείας ($\chi^2 = 0.59 < \chi_{1,0.95}^2 = 3.84$, $p\text{-value} = 0.557 > 0.05$). Στην παραπάνω ανάλυση δε λάβαμε όμως υπόψη ότι η δειγματοληψία είναι εξαρτημένη. Δηλαδή υπάρχουν ζεύγη παρατηρήσεων με παρόμοια χαρακτηριστικά. Συνεπώς πρέπει να λάβουμε ως μονάδα μελέτης το εξομοιωμένο ζευγάρι των ασθενών οπότε το μέγεθος δείγματος θα είναι τώρα ίσο με 621. Οι μεταβλητές που θα χρησιμοποιήσουμε στον πίνακα θα είναι οι συνδυασμοί της επιβίωσης για τις δύο θεραπείες. Έτσι τα δυνατά αποτελέσματα θα είναι τέσσερα: επιβίωση των ασθενών και στις δύο θεραπείες, κατάληξη και στις δύο θεραπείες, επιβίωση στη θεραπεία A αλλά όχι στη B και επιβίωση στη θεραπεία B αλλά όχι στη A. Έτσι λοιπόν προκύπτει ο ακόλουθος πίνακας:

Συνεπώς εδώ μας ενδιαφέρει το ποσοστό επιβίωσης των δύο ομάδων/θεραπειών να είναι ίσο. Συνεπώς

$$P(\text{Επιβίωση ασθενών Θεραπείας A}) = P(\text{Επιβίωση ασθενών Θεραπείας B}) \Leftrightarrow \pi_{1.} = \pi_{.1}$$

$$\Leftrightarrow n_{12} = n_{21}.$$

Από τα παραπάνω συμπεραίνουμε ότι, σε ένα 2×2 πίνακα συνάφειας, στην ουσία συγκρίνουμε τις συχνότητες της 2ης διαγωνίου του Πίνακα 5.2.1 που δείχνουν την ασυμφωνία των ζευγαριών ως προς τη μεταβλητή απόκρισης.

X: ΘΕΡΑΠΕΙΑ A Επιβίωση	ΘΕΡΑΠΕΙΑ B Επιβίωση στα 5 έτη		Περ.Κατ.
	1: Ναι	2: Όχι	Περ.Κατ. X
1: Ναι	510	16	526
2: Όχι	5	90	95
Περ.Κατ. Y	515	106	621

Πίνακας 5.5: Ζευγαρωτά Δεδομένα Καρκίνου του Μαστού· Παράδειγμα 5.3 (Rosner, 1994, σελ. 377).

Ορισμός 5.6 Αρμονικό ζευγάρι (*concordant pair*) σε μία εξομοιωμένη μελέτη είναι το ζεύγος των παρατηρήσεων όπου το αποτέλεσμα (δηλαδή η μεταβλητή απόκρισης) είναι το ίδιο και για τα δυο άτομα του ζεύγους.

Ορισμός 5.7 Δυσαρμονικό ζευγάρι (*discordant pair*) σε μία εξομοιωμένη μελέτη είναι το ζεύγος των παρατηρήσεων όπου το αποτέλεσμα (δηλαδή η μεταβλητή απόκρισης) διαφέρει για τα δυο άτομα του ζεύγους.

Εναλλακτικά μπορούμε να ονομάσουμε τα παραπάνω ζευγάρια και «συμφωνούντα» και «διαφωνούντα» ζεύγη αντίστοιχα.

Για να κατασκευάσουμε έναν έλεγχο σε αυτή την περίπτωση, υποθέτουμε ότι αν η H_0 ισχύει τότε

$$n_{21} \sim \text{Binomial}\left(\frac{1}{2}, n_D\right).$$

Συνεπώς $E(n_{21}) = n_D/2$ και $V(n_{21}) = n_D/4$. Υπό την προϋπόθεση ότι $npq \geq 5$ δηλαδή $n_D \geq 20$ μπορούμε να χρησιμοποιήσουμε την κανονική προσέγγιση οπότε

$$\frac{n_{21} - n_D/2}{\sqrt{n_D/4}} \sim N(0, 1).$$

Υψώνοντας στο τετράγωνο προκύπτει ο έλεγχος McNemar που χρησιμοποιεί τη συνάρτηση ελέγχου:

$$\frac{(n_{21} - n_D/2)^2}{n_D/4} \sim \chi_1^2. \quad (5.1)$$

όπου n_D είναι το σύνολο των δυσαρμονικών ζευγαριών δηλαδή $n_D = n_{12} + n_{21}$. Σε μερικές περιπτώσεις για λόγους διόρθωσης χρησιμοποιείται και ο τύπος:

$$\frac{(|n_{21} - n_D/2| - 1/2)^2}{n_D/4} \sim \chi_1^2. \quad (5.2)$$

Η διαδικασία εφαρμογής του ελέγχου McNemar μπορεί να περιγραφεί από τα ακόλουθα βήματα:

1. Φτιάχνουμε τον 2×2 πίνακα συνάφειας για τα ζεύγη των εξομοιωμένων παρατηρήσεων. Στις γραμμές θα είναι το αποτέλεσμα (μεταβλητή απόκρισης) για τον πρώτο ασθενή (που ανήκει στην ομάδα ή θεραπεία A) και στις στήλες θα είναι το αποτέλεσμα (μεταβλητή απόκρισης) για το δεύτερο ασθενή (που ανήκει στην ομάδα ή θεραπεία B).
2. Υπολογίζουμε την παρατηρούμενη συνάρτηση ελέγχου χ_{obs}^2 από τον τύπο 5.1 ή 5.2.
3. Απορρίπτουμε την H_0 : «ποσοστό εμφάνισης νόσου είναι το ίδιο στις δύο ομάδες» αν $\chi_{obs}^2 > \chi_{1,1-a}^2$. Εναλλακτικά μπορούμε να υπολογίσουμε το $p - value = P(X^2 \geq \chi_{obs}^2)$ και απορρίπτουμε την μηδενική υπόθεση αν $p - value < a$.

Σημείωση : ο παραπάνω ασυμπτωτικός έλεγχος ισχύει για $n_D \geq 20$.

Παράδειγμα 5.3 (αποτελέσματα): Εφόσον ο αριθμός των δυσαρμονικών ζευγών $n_D = 21 \geq 20$ μπορούμε να χρησιμοποιήσουμε την παραπάνω διαδικασία οπότε

$$\chi_{obs}^2 = \frac{(|n_{21} - n_D/2| - 1/2)^2}{n_D/4} = \frac{(|5 - 10.5| - 1/2)^2}{10.5/4} = \frac{5^2}{5.25} = 4.76 > \chi_{1,0.95}^2 = 3.84$$

συνεπώς απορρίπτουμε την υπόθεση ότι οι δύο θεραπείες είναι ίδιες (ενώ υποθέτοντας ανεξάρτητα δείγματα είχαμε καταλήξει σε διαφορετικό αποτέλεσμα). Αν θέλουμε να υπολογίσουμε το $p - value$ τότε

$$p - value = P(X^2 \geq \chi_{obs}^2) = P(X^2 \geq 4.76) = 1 - P(X^2 \leq 4.76) = 1 - 0.9708 = 0.029.$$

Για $n_D < 20$ θα πρέπει να χρησιμοποιήσουμε τον ακριβή έλεγχο ο οποίος βασίζεται στη Διωνυμική κατανομή. Έτσι το $p - value$ θα το υπολογίσουμε από τον ακόλουθο τύπο

$$p - value = \begin{cases} 1 & \text{αν } n_{21} = n_D/2 \\ 2 \sum_{k=0}^{n_{21}} \binom{n_D}{k} 2^{-n_D} & \text{αν } n_{21} < n_D/2 \\ 2 \sum_{k=n_{21}}^{n_D} \binom{n_D}{k} 2^{-n_D} & \text{αν } n_{21} > n_D/2 \end{cases} .$$

Παράδειγμα 5.4 Μελέτη σε 20 άτομα με σκοπό τον έλεγχο της καλής λειτουργίας μηχανημάτων μέτρησης της πίεσης έδωσε τα ακόλουθα αποτελέσματα:

Άτομο	C	O	Άτομο	C	O	Άτομο	C	O	Άτομο	C	O
1	-	-	6	+	-	11	+	-	16	+	-
2	-	-	7	-	-	12	+	-	17	+	-
3	+	-	8	+	+	13	-	-	18	-	-
4	+	+	9	+	+	14	+	-	19	-	-
5	-	-	10	-	-	15	-	+	20	-	-

Πίνακας 5.6: Δεδομένα Παραδείγματος 5.4 για τον Έλεγχο της Αξιοπιστίας Μηχανημάτων Μέτρησης Πίεσης (+: Υψηλή Πίεση, -: Χαμηλή Πίεση, C: μέτρηση με μηχανήμα, O μέτρηση από ειδικό Ιατρό).

Εδώ πρώτο βασικό ερώτημα είναι το πραγματικό μέγεθος δείγματος. Είναι 20 ή 40; Φυσικά εδώ είναι 20 απλά έχουμε ζευγάρια τιμών στο ίδιο άτομο (από μηχανήμα και από ένα Ιατρό). Ο πίνακας που προκύπτει είναι ο ακόλουθος:

Πίεση από Μηχάνημα	Μέτρηση Πίεσης από Ιατρό		Περ.Κατ. X
	1: Υψηλή	2: Χαμηλή	
1: Υψηλή	3	7	10
2: Χαμηλή	1	9	10
Περ.Κατ. Y	4	9	20

Πίνακας 5.7: 2×2 Πίνακας των Δεδομένων του Παραδείγματος 5.4.

Υπολογίζουμε το p -value χρησιμοποιώντας τον ακριβή έλεγχο εφόσον $n_D = 1 + 7 = 8 < 20$. Εφόσον $n_{21} < n_D/2 = 8/2 = 4$, τότε το p -value το υπολογίζουμε από τον ακόλουθο τύπο

$$p\text{-value} = 2 \left\{ \binom{8}{0} 2^{-8} + \binom{8}{1} 2^{-8} \right\} = 9 \left(\frac{1}{2} \right)^7 = 0.070 > 0.05$$

άρα δεν απορρίπτουμε την H_0 σε επίπεδο σημαντικότητας 5%.

Άσκηση 5.1 Να γίνει η παραπάνω άσκηση με το χέρι και το SPSS.

SPSS: Ο έλεγχος του McNemar στο SPSS 12.0 μπορεί να γίνει με τα ακόλουθα μενού:

1. Analyse>Descriptives>Crosstabs : Κατασκευή Πίνακα συνάφειας
2. Statistics|McNemar : Επιλογή Εφαρμογής Ελέγχου McNemar στον πίνακα συνάφειας που κατασκευάσαμε
3. Exact|Exact : Υπολογισμός του ακριβούς ελέγχου McNemar.

5.2.2 Ο Δείκτης Συμφωνίας Κάπα (Καπα).

Στην προηγούμενη ενότητα ελέγξαμε τη σχέση μεταξύ δύο δίτιμων μεταβλητών σε εξαρτημένα δείγματα (εξομοιωμένα ζευγάρια). Σε μερικές περιπτώσεις ξέρουμε εκ των προτέρων ότι υπάρχει σχέση μεταξύ των μεταβλητών οπότε στόχος είναι να ποσοτικοποιήσουμε το βαθμό της συσχέτισης. Αυτό κυρίως συμβαίνει στις μελέτες αξιοπιστίας μετρήσεων (reliability studies) όπου σκοπός είναι να μετρήσουμε την αναπαραγωγιμότητα ή επαναληψιμότητα μιας μεταβλητής (reproducibility and reliability). Στις περιπτώσεις αυτές συγκαταλέγουμε λοιπόν την επαναλαμβανόμενη μέτρηση μιας δίτιμης μεταβλητής σε κοντινό χρονικό διάστημα. Άλλη περίπτωση είναι αυτή που θέλουμε να μετρήσουμε ή να συγκρίνουμε την αξιοπιστία μιας μέτρησης που γίνεται με ένα μηχάνημα σε σχέση με μια μέτρηση αναφοράς εμπιστευόμαστε (π.χ. ηλεκτρονικό μηχάνημα μέτρησης πίεσης και μέτρηση πίεσης από Ιατρό) ή τη συμφωνία δύο αξιολογητών, βαθμολογητών ή Ιατρών.

Συνεπώς στις περιπτώσεις αυτές έχουμε πίνακες $I \times I$ (οι δύο μετρήσεις έχουν τον ίδιο αριθμό επιπέδων αφού μετράνε το ίδιο πράγμα) και μας ενδιαφέρει το ποσοστό συμφωνίας $\pi_o = \sum_{i=1}^I \pi_{ii}$ το οποίο εκτιμάται από το παρατηρούμενο ποσοστό συμφωνίας $p_o = \sum_{i=1}^I n_{ii}/n$. Υπό την υπόθεση της ανεξαρτησίας των δύο μετρήσεων (συνεπώς την πλήρη έλλειψη αναπαραγωγιμότητας) αναμένουμε ποσοστό συμφωνίας ίσο με $\pi_e = \sum_{i=1}^I \pi_{i\bullet} \cdot \pi_{\bullet i}$ ο οποίος δίνεται στο δείγμα από την ποσότητα $p_e = \sum_{i=1}^I n_{i\bullet} \cdot n_{\bullet i}/n^2$ δηλαδή το άθροισμα των αναμενόμενων κελιών συμφωνίας. Αυτή η αναλογία συμφωνίας οφείλεται σε τυχαίους παράγοντες και για το λόγο αυτό μπορεί να θεωρηθεί ως σημείο αναφοράς και σύγκρισης. Άρα ένα μέτρο συμφωνίας θα είναι η διαφορά $\pi_o - \pi_e$ όπου το μηδέν δείχνει μηδενική ή τυχαία αναπαραγωγιμότητα. Εμείς όμως θα επιθυμούσαμε να είχαμε ένα μέτρο το οποίο να είναι φραγμένο και η ανώτατη τιμή του να υποδεικνύει πλήρη συμφωνία ή αναπαραγωγιμότητα. Έτσι μπορούμε να διαιρέσουμε με το $\max(\pi_o - \pi_e) = 1 - \pi_e$ εφόσον η καλύτερη δυνατή περίπτωση είναι $\pi_o = 1$ όταν έχουμε πλήρη συμφωνία και προκύπτει ο δείκτης συμφωνίας Κάπα του Cohen (1960):

$$\kappa = \frac{\sum_{i=1}^I \pi_{ii} - \sum_{i=1}^I \pi_{i\bullet} \cdot \pi_{\bullet i}}{1 - \sum_{i=1}^I \pi_{i\bullet} \cdot \pi_{\bullet i}} = \frac{\pi_o - \pi_e}{1 - \pi_e},$$

ο οποίος εκτιμάται από το

$$\hat{\kappa} = \frac{\sum_{i=1}^I n_{ii}/n - \sum_{i=1}^I n_{i\cdot} n_{\cdot i}/n^2}{1 - \sum_{i=1}^I n_{i\cdot} n_{\cdot i}/n^2} = \frac{p_o - p_e}{1 - p_e}.$$

Η μέγιστη τιμή του κ είναι το ένα (πλήρη συμφωνία). Η τιμή μηδέν δίνεται όταν το παρατηρούμενο ποσοστό συμφωνίας είναι ίσο με αυτό που αναμένουμε όταν υπάρχει ανεξαρτησία (άρα οφείλεται σε τυχαιότητα). Αρνητικές τιμές εμφανίζονται όταν η παρατηρούμενη αναλογία συμφωνίας είναι μικρότερη της συμφωνίας που περιμένουμε λόγω της τυχαιότητας. Αρνητικές τιμές σύμφωνα τον Agresti (1990, σελ. 366-367) συμβαίνουν σπάνια στην πράξη.

Η διακύμανση του παραπάνω εκτιμητή δίνεται από τον τύπο

$$\text{Var}(\hat{\kappa}) = \frac{1}{n(1 - p_e)^2} \left(p_e + p_e^2 - \sum_{i=1}^I p_{i\cdot} p_{\cdot i} (p_{i\cdot} + p_{\cdot i}) \right).$$

Ένας ασυμπτωτικός έλεγχος της υπόθεσης $H_0 : \kappa = 0$ έναντι της εναλλακτικής $H_1 : \kappa > 0$ μπορεί να γίνει χρησιμοποιώντας τη συνάρτηση ελέγχου $z = \frac{\hat{\kappa}}{\sqrt{\text{Var}(\hat{\kappa})}}$ η οποία ακολουθεί ασυμπτωτικά τη $N(0, 1)$ κατανομή όταν η H_0 ισχύει. Σημείωση ότι μας ενδιαφέρει ο μονόπλευρος έλεγχος αφού οι αρνητικές τιμές δεν έχουν αξιολογή ερμηνεία και συνήθως λαμβάνονται ως έλλειψη συμφωνίας και επαναληψιμότητας.

Παράδειγμα 5.5 Μελέτες αξιοπιστίας γίνονται κυρίως στις κοινωνικές επιστήμες, στην Ψυχολογία ή στην Ψυχιατρική αθλά και στη διατροφική επιδημιολογία όπου οι μελέτες βασίζονται κυρίως σε μετρήσεις από ερωτηματολόγια. Στις μελέτες καταγραφής διατροφικών συνηθειών δίνεται ένα ερωτηματολόγιο διατροφικών συχνοτήτων (Food Frequency Questionnaire) όπου ο ερωτώμενος καλείται να θυμηθεί τι κατανάλωσε το τελευταίο μήνα ή χρόνο. Είναι προφανές ότι οι μετρήσεις αυτές έχουν μεγάλο σφάλμα μέτρησης το οποίο οφείλεται στην ελλιπή ανάκληση της πληροφορίας. Στο παράδειγμα που ακολουθεί ερωτήθηκαν 537 γυναίκες με ταχυδρομική έρευνα για το αν καταναλώνουν περισσότερες της μιας μερίδας μοσχαρίσιο κρέας ανά εβδομάδα. Οι μετρήσεις έγιναν δύο φορές με απόσταση μερικών μηνών. Τα αποτελέσματα δίνονται στον Πίνακα 5.5 που ακολουθεί.

Από τα παραπάνω έχουμε:

$$\begin{aligned} p_o &= (136 + 240)/537 = 376/537 = 0.70 \\ p_e &= 205 \times 228/537^2 + 332 \times 309/537^2 = 0.382 \times 0.618 + 0.425 \times 0.575 = 0.518 \\ \hat{\kappa} &= \frac{0.70 - 0.518}{1 - 0.518} = 0.378 \\ se(\hat{\kappa}) &= 0.040. \end{aligned}$$

X: Πρώτη Μέτρηση	Δεύτερη Μέτρηση		Περ.Κατ.
	1: ≤ 1 μέριδα/εβδομάδα	2: < 1 μέριδα/εβδομάδα	Περ.Κατ. X
1: ≤ 1 μέριδα/εβδομάδα	136	92	228
2: > 1 μέριδα/εβδομάδα	69	240	309
Περ.Κατ. Y	205	332	537

Πίνακας 5.8: 2×2 Πίνακας των Δεδομένων του Παραδείγματος 5.5.

Συνεπώς $z = 0.378/0.040 = 8.799 > 1.65$ άρα απορρίπτουμε την υπόθεση την πλήρους ασυμφωνίας ή έλλειψης αναπαραγωγισιμότητας ($H_0 : \kappa = 0$).

Στο παραπάνω παράδειγμα είδαμε ότι δεν απορρίψαμε τη μηδενική υπόθεση που σημαίνει ότι υπάρχει κάποια αναπαραγωγισιμότητα ή συμφωνία άλλα αυτό δε σημαίνει ότι είναι υψηλή. Γενικά η μηδενική υπόθεση απορρίπτεται σχετικά για χαμηλές τιμές του $\hat{\kappa}$ ειδικά για μεγάλα δείγματα (ανάλογη είναι η περίπτωση του δείκτη γραμμικής συσχέτισης του Pearson και του αντίστοιχου ελέγχου για $\rho = 0$). Για τον λόγο αυτό οι Landis και Koch (1977) δίνουν έναν εμπειρικό πίνακα για τις τιμές του κ :

- Αν $\hat{\kappa} > 0.75$ τότε έχουμε έξοχη αναπαραγωγισιμότητα και επαναληπτικότητα (ή υψηλή συμφωνία).
- Αν $0.4 \leq \hat{\kappa} \leq 0.75$ τότε έχουμε καλή αναπαραγωγισιμότητα και επαναληπτικότητα (ή ικανοποιητική συμφωνία).
- Αν $0 \leq \hat{\kappa} < 0.4$ τότε έχουμε οριακή αναπαραγωγισιμότητα και επαναληπτικότητα (ή χαμηλή συμφωνία).

Σύμφωνα με τον Rosner (1994, σελ. 426) ο δείκτης κάπα θα πρέπει να προτιμάται για τη μέτρηση της αξιοπιστίας μιας μεταβλητής ή ερωτήματος άρα όταν έχουμε επαναλαμβανόμενες μετρήσεις της ίδιας κατηγορικής μεταβλητής. Στις περιπτώσεις σύγκρισης μιας μέτρησης με μία μέτρηση αναφοράς προτείνει τη χρήση της ειδικότητας και ευαισθησίας.

Εδώ πρέπει να τονίσουμε τις διαφορές μεταξύ της δοκιμασίας McNemar και του δείκτη κάπα. Σκοπός της δοκιμασίας McNemar είναι να ελέγξει αν οι περιθώριες κατανομές είναι ίσες. Έτσι σε 2×2 πίνακες που προκύπτουν από εξομοιωμένα ζεύγη με διαφορετική θεραπεία σε σχέση με την επιβίωση, μας ενδιαφέρει να δούμε αν η πιθανότητα επιβίωσης είναι η ίδια για τις δύο θεραπείες. Από την άλλη μεριά ο δείκτης κάπα συγκρίνει πόσο τα αποτελέσματα των δύο θεραπειών σχετίζονται-συμφωνούν και αναφέρεται κυρίως στην εκ-των-προτέρων γνωστή

εξάρτηση των δειγμάτων λόγω εξομοίωσης. Άρα η πιο συχνή περίπτωση είναι να έχουμε μεγάλο ποσοστό συμφωνίας (λόγω εξομοίωσης) αλλά και στατιστικά σημαντική διαφοροποίηση ως προς την αποτελεσματικότητα της κάθε θεραπείας (σκεφτείτε κατά αναλογία τη σχέση του δείκτη συσχέτισης του Pearson και του t-test ανά ζεύγη· υψηλή συσχέτιση δε σημαίνει και ίσους μέσους). Επίσης μια βασική διαφορά είναι και τα κελιά από τα οποία αντλεί πληροφορία η κάθε διαδικασία. Η δοκιμασία του McNemar βασίζεται στα μη-διαγώνια στοιχεία ενώ ο δείκτης κάπα στα διαγώνια στοιχεία.

5.3 Προτυποποίηση

Η προτυποποίηση είναι μια μέθοδος που χρησιμοποιείται για να κάνει τους δείκτες θνησιμότητας δύο πληθυσμών ή δειγμάτων συγκρίσιμους λαμβάνοντας υπόψη κάποιους συγχυτικούς παράγοντες όπως για παράδειγμα το φύλο και η ηλικία. Στην προτυποποίηση υπολογίζουμε το δείκτη (συνήθως την αναλογία) μίας νόσου υποθέτοντας ότι οι πληθυσμοί ή τα δείγματα που συγκρίνουμε έχουν την ίδια κατανομή ως προς το συγχυτικό παράγοντα. Έτσι αν p_i είναι οι δειγματικές αναλογίες εμφάνισης μιας νόσου για το i επίπεδο ενός συγχυτικού παράγοντα με I επίπεδα τότε η τυποποιημένη αναλογία της νόσου υπολογίζεται από τον τύπο

$$p^{st} = \sum_{i=1}^I \frac{n_i^*}{n^*} p_i = \sum_{i=1}^I w_i^* p_i$$

όπου n_i^* είναι ο συνολικός αριθμός των ατόμων του πρότυπου πληθυσμού που ανήκουν στο i επίπεδο του συγχυτικού παράγοντα, n^* είναι ο συνολικός αριθμός των ατόμων του πρότυπου πληθυσμού και $w_i^* = n_i^*/n^*$ είναι η αναλογία των ατόμων του πρότυπου πληθυσμού που ανήκουν στο i επίπεδο του συγχυτικού παράγοντα.

Η τυποποιημένη αναλογία p^{st} μπορεί να ερμηνευτεί ως η αναλογία ατόμων που θα είχαν τη νόσο αν ο πληθυσμός μας είχε την ίδια κατανομή του συγχυτικού παράγοντα όπως ο πρότυπος πληθυσμός (ή πληθυσμός αναφοράς). Ως πρότυπος πληθυσμός συνήθως λαμβάνεται ο πληθυσμός μίας χώρας στον οποίο πιθανώς ανήκουν οι υπό-σύγκριση πληθυσμοί ή απλά το άθροισμα των πληθυσμών που συγκρίνουμε.

5.4 Συμπερασματολογία για Στρωματοποιημένα Κατηγορικά Δεδομένα

Παράδειγμα 5.6 Μια μελέτη το 1985 εντόπισε 518 περιπτώσεις καρκίνου σε νοσοκομεία των ΗΠΑ και 518 μάρτυρες μέσω ταχυδρομείου (για λεπτομέρειες βλ. Sandler *et. al* , 1985,

Amer.J.Epidem.). Ο κύριος σκοπός της μελέτης ήταν η εξέταση του παθητικού καπνίσματος στην εμφάνιση του καρκίνου του πνεύμονα. Στη μελέτη το παθητικό κάπνισμα ορίστηκε ως «έκθεση στο κάπνισμα του/της συζύγου που κάπνιζε τουλάχιστον 1 τσιγάρο την ημέρα για τους τελευταίους 6 μήνες».

Ένας πιθανός συγχυτικός παράγοντας, είναι το κάπνισμα του ίδιου του ατόμου που εξετάζουμε αφού προφανώς συνδέεται με τη νόσο αλλά και με το κάπνισμα της συζύγου (άρα και το παθητικό κάπνισμα όπως ορίστηκε στη μελέτη αυτή). Τα δεδομένα δίνονται στον Πίνακα 5.6 που ακολουθεί.

X:	Z: Μη Καπνιστές (1)			Z: Καπνιστές (2)		
	Y: Παθ.Καπνιστής	Περ.Κατ.		Y: Παθ.Καπνιστής	Περ.Κατ.	
Καρκίνος	1: Ναι	2: Όχι	X	1: Ναι	2: Όχι	X
1: Ασθενής	120	111	231	161	117	278
2: Μάρτυρας	80	155	235	130	124	254
Περ.Κατ. Y	200	266	466	291	241	532
ΛΣΠ (OR)	2.1			1.3		
95% ΔΕ	1.72-2.47			0.95-1.64		

Πίνακας 5.9: Στρωματοποιημένα ως προς το Κάπνισμα δεδομένα μελέτης της σχέσης Παθητικού Καπνίσματος και Καρκίνου Παραδείγματος 5.6 (Rosner, 1994, σελ. 399–400).

Στο παραπάνω παράδειγμα, από τον περιθώρια πίνακα (αγνοώντας το ατομικό κάπνισμα) βρίσκουμε $\Lambda\Sigma\P = 1.64$ (95% ΔΕ=1.35-1.88). Αντίθετα αν δούμε τα στρωματοποιημένα δεδομένα βλέπουμε ότι στους καπνιστές η σχέση δεν είναι στατιστικά σημαντική ενώ αντίθετα για τους μη καπνιστές ο κίνδυνος είναι σημαντικά αυξημένος (περίπου διπλάσιος για τα άτομα που εκτίθενται σε παθητικό κάπνισμα). Το ερώτημα εδώ είναι πως θα συνδυάσουμε τους παραπάνω στρωματοποιημένους εκτιμητές σε έναν που θα λαμβάνει υπόψη του και το συγχυτικό παράγοντα.

Γενικά η ενότητα αυτή αναφέρεται σε $2 \times 2 \times K$ πίνακες, όπου K είναι τα επίπεδα του συγχυτικού παράγοντα για τον οποίο θέλουμε να διορθώσουμε/ελέγξουμε τα αποτελέσματα (πιο απλά να λάβουμε υπόψη μας). Η υπόθεση που θέλουμε να ελέγξουμε εδώ είναι «ανεξαρτησία των μεταβλητών X και Y για όλα τα επίπεδα του συγχυτικού παράγοντα Z » δηλαδή

$$H_0 : OR_1 = OR_2 = \dots = OR_K = 1 .$$

Αν για κάθε επίπεδο k έχουμε τον ακόλουθο πίνακα

Z=k			
X:	Y		Περ.Κατ. X
	1: Ναι	2: Όχι	
1	n_{11k}	n_{12k}	$n_{1\bullet k}$
2	n_{21k}	n_{22k}	$n_{2\bullet k}$
Περ.Κατ. Y	$n_{\bullet 1k}$	$n_{\bullet 2k}$	$n_{\bullet \bullet k}$

Υποθέτοντας σταθερές περιθώριες κατανομές και ανεξαρτησία συνεπάγεται

$$n_{11k} \sim \text{HyperGeometric}(n_{1\bullet k}, n_{2\bullet k}, n_{\bullet 1k})$$

όπου $\text{HyperGeometric}(m, n, N)$ είναι η υπεργεωμετρική κατανομή με συνάρτηση πιθανότητας

$$f(x) = \frac{\binom{n}{x} \binom{m}{N-x}}{\binom{n+m}{N}},$$

μέση τιμή $E(X) = Np = N \frac{n}{(n+m)}$ και διακύμανση $V(X) = Np(1-p) \frac{n+m-N}{n+m-1} = N \frac{nm}{(n+m)^2} \frac{n+m-N}{n+m-1}$.

Συνεπώς υπό την H_0 θα έχουμε

$$E(n_{11k}) = \frac{n_{1\bullet k} n_{\bullet 1k}}{n_{\bullet \bullet k}}$$

$$V(n_{11k}) = \frac{n_{1\bullet k} n_{2\bullet k} n_{\bullet 1k} n_{\bullet 2k}}{n_{\bullet \bullet k}^2 (n_{\bullet \bullet k} - 1)}.$$

Αν ισχύει η H_0 , ο συνολικός αριθμός (αθροίζοντας ως προς το συγχυτικό παράγοντα) των κελιών 11 θα δίνεται ως

$$E = \sum_{k=1}^K E(n_{11k}) = \sum_{k=1}^K \frac{n_{1\bullet k} n_{\bullet 1k}}{n_{\bullet \bullet k}}$$

και το οποίο θα πρέπει να συγκριθεί με τη αντίστοιχη παρατηρούμενη ποσότητα $O = n_{11\bullet} = \sum_{k=1}^K n_{11k}$. Κάτω από την H_0 , $E(O) = E$ και $V(O) = \sum_{k=1}^K V(n_{11k}) = V$. Έτσι κατασκευάζουμε τον έλεγχο Mantel-Haenszel με τη χρήση της συνάρτησης ελέγχου:

$$\chi_{MH}^2 = \frac{(|O - E| - 1/2)^2}{V} \sim \chi_1^2$$

Παράδειγμα 5.6 (αποτελέσματα).

Στο παραπάνω παράδειγμα έχουμε

$$O = n_{111} + n_{112} = 120 + 161 = 281$$

$$\begin{aligned}
E_1 &= \frac{231 \times 200}{466} = 99.1, E_2 = \frac{278 \times 291}{532} = 152.1 \Leftrightarrow E = 99.1 + 152.1 = 251.2 \\
V_1 &= \frac{231 \times 235 \times 200 \times 266}{466^2 \times 465} = 28.6 \\
V_2 &= \frac{291 \times 241 \times 278 \times 254}{532^2 \times 531} = 32.95 \\
V &= V_1 + V_2 = 28.6 + 32.95 = 61.55
\end{aligned}$$

Συνεπώς $\chi_{MH}^2 = \frac{(|281-251.2|-1/2)^2}{61.55} = 13.94 > 3.84 = \chi_{1,0.95}^2$, p -value = 0.0001887 άρα απορρίπτουμε την μηδενική υπόθεση και συνεπώς υπάρχει σχέση παθητικού καπνίσματος και καρκίνου ακόμα και αν ελέγξουμε/διορθώσουμε για το ατομικό κάπνισμα (ή απλά λάβουμε υπόψη μας το προσωπικό κάπνισμα).

5.4.1 Διαφοροποίηση Επιδράσεων (Effect Modification).

Στον προηγούμενο έλεγχο, εξετάζουμε αν υπάρχει ανεξαρτησία νόσου-παράγοντα κινδύνου για κάθε επίπεδο ενός συγχυτικού παράγοντα. Αν απορριφθεί η υπό-συνθήκη ανεξαρτησία νόσου-παράγοντα κινδύνου μας ενδιαφέρει να δούμε κατά πόσο η σχέση παραμένει ίδια για τα διάφορα επίπεδα του συγχυτικού παράγοντα. Σε αυτή την ενότητα θα δούμε τη δοκιμασία που μπορούμε να χρησιμοποιήσουμε για τον έλεγχο της ομοιογένειας των ΛΣΠ. Πριν προχωρήσουμε στον έλεγχο θα δώσουμε τον ορισμό του «διαφοροποιητή επιδράσεων» (effect modifier).

Ορισμός 5.8 *Αν η μορφή και η ισχύς της σχέσης νόσου-παράγοντα κινδύνου αλληλάζει για διαφορετικά επίπεδα ενός συγχυτικού παράγοντα τότε υπάρχει αλληλεπίδραση παράγοντα κινδύνου και συγχυτικού παράγοντα πάνω στη νόσο ή αλλιώς διαφοροποίηση της επίδρασης του συγχυτικού παράγοντα στη νόσο για διαφορετικά επίπεδα του συγχυτικού παράγοντα (effect modification). Αυτός ο συγχυτικός παράγοντας ονομάζεται διαφοροποιητής επιδράσεων (effect modifier) του παράγοντα κινδύνου στη νόσο.*

Στο παράδειγμα 5.6 φαίνεται ότι το ατομικό κάπνισμα διαφοροποιεί τις επιδράσεις του παθητικού καπνίσματος στον καρκίνο εφόσον οι δύο ΛΣΠ απέχουν αρκετά μεταξύ τους (1.3 για τους καπνιστές και 2.1 για τους μη καπνιστές). Αν κάνουμε τον έλεγχο ομοιογένειας ΛΣΠ που ακολουθεί θα δούμε ότι η διαφορά είναι οριακά στατιστικά σημαντική δηλαδή σημαντική σε επίπεδο σημαντικότητας 10% αλλά όχι σε επίπεδο σημαντικότητας 5%.

Συνεπώς εδώ θα ελέγξουμε την υπόθεση $H_0 : OR_1 = OR_2 = \dots = OR_K$ έναντι της εναλλακτικής $H_1 : OR_i \neq OR_j$ για κάποιο συνδυασμό επιπέδων $i \neq j$ του συγχυτικού παράγοντα.

Ο έλεγχος ομοιογένειας των ΛΣΠ χρησιμοποιεί τον ακόλουθο τύπο

$$\chi_{HOM}^2 = \sum_{k=1}^K w_k \left(\log \widehat{OR}_k - \overline{\log OR} \right)^2 \sim \chi_{K-1}^2$$

με

$$w_k = \frac{1}{\text{Var}(\log \widehat{OR}_k)} = \left(\frac{1}{n_{11k}} + \frac{1}{n_{12k}} + \frac{1}{n_{21k}} + \frac{1}{n_{22k}} \right)^{-1},$$

$$\overline{\log OR} = \frac{\sum_{k=1}^K w_k \log \widehat{OR}_k}{\sum_{k=1}^K w_k} \quad \text{και} \quad \widehat{OR}_k = \frac{n_{11k}n_{22k}}{n_{12k}n_{21k}}.$$

Εναλλακτικός τύπος είναι ο ακόλουθος

$$\chi_{HOM}^2 = \sum_{k=1}^K w_k (\log \widehat{OR}_k)^2 - \frac{(\sum_{k=1}^K w_k \log \widehat{OR}_k)^2}{\sum_{k=1}^K w_k}.$$

Παράδειγμα 5.6 (αποτελέσματα II).

Εφαρμόζοντας τον τελευταίο τελευταίο τύπο έχουμε

$$\log \widehat{OR}_1 = \log \frac{120 \times 155}{80 \times 111} = \log 2.0945 = 0.7394$$

$$\log \widehat{OR}_2 = \log \frac{161 \times 124}{130 \times 117} = \log 1.3126 = 0.2720$$

$$w_1 = \left(\frac{1}{120} + \frac{1}{111} + \frac{1}{80} + \frac{1}{155} \right)^{-1} = (0.0083 + 0.0090 + 0.0125 + 0.0065)^{-1}$$

$$= 1/0.0363 = 27.55$$

$$w_2 = \left(\frac{1}{161} + \frac{1}{117} + \frac{1}{130} + \frac{1}{124} \right)^{-1} = (0.0062 + 0.0085 + 0.0077 + 0.0081)^{-1}$$

$$= 1/0.0305 = 32.79$$

Συνεπώς

$$\chi_{HOM}^2 = 27.55 \times (0.7394)^2 + 32.79 \times (0.272)^2 - \frac{(27.55 \times 0.7394 + 32.79 \times 0.272)^2}{27.55 + 32.79}$$

$$= 15.06 + 2.43 - \frac{857.866}{60.34} = 3.27 \Leftrightarrow p\text{-value} = 0.070$$

Εφόσον $\chi_{HOM}^2 = 3.27 < \chi_{1,0.95}^2 = 3.84$ δεν απορρίπτουμε την H_0 ότι οι ΛΣΠ της σχέσης παθητικού καπνίσματος - καρκίνου είναι ίσοι για τους καπνιστές και μη καπνιστές σε επίπεδο σημαντικότητας 5%.

5.4.2 Εκτίμηση Κοινού ΛΣΠ για Στρωματοποιημένα Δεδομένα.

Όταν ο έλεγχος ομοιογένειας (ή ισότητας) των ΛΣΠ δεν απορριφθεί τότε μας ενδιαφέρει να εκτιμήσουμε τον κοινό ΛΣΠ λαμβάνοντας όμως υπόψη και το συγχυτικό παράγοντα. Αυτό γίνεται με τον εκτιμητή κοινου ΛΣΠ των Mantel-Haenszel και δίνεται από τον τύπο

$$\widehat{OR}_{MH} = \frac{\sum_{k=1}^K n_{11k}n_{22k}/n_{\bullet\bullet k}}{\sum_{k=1}^K n_{12k}n_{21k}/n_{\bullet\bullet k}}.$$

Η διακύμανση του λογαρίθμου του παραπάνω εκτιμητή δίνεται από τον ακόλουθο τύπο

$$\text{Var}(\log \widehat{OR}_{MH}) = \frac{\sum_{k=1}^K P_k R_k}{2 \left(\sum_{k=1}^K R_k \right)^2} + \frac{\sum_{k=1}^K (P_k S_k + Q_k R_k)}{2 \left(\sum_{k=1}^K R_k \right) \left(\sum_{k=1}^K S_k \right)} + \frac{\sum_{k=1}^K Q_k S_k}{2 \left(\sum_{k=1}^K S_k \right)^2}$$

με

$$P_k = n_{11k} + n_{22k}$$

$$Q_k = n_{12k} + n_{21k}$$

$$R_k = n_{11k} n_{22k} / n_{\bullet \bullet k}$$

$$S_k = n_{12k} n_{21k} / n_{\bullet \bullet k} .$$

Και φυσικά υπολογίζουμε ένα $(1 - \alpha)100\%$ διάστημα εμπιστοσύνης για το λογάριθμο του ΛΣΠ από τον τύπο

$$\log \widehat{OR}_{MH} \pm z_{1-\alpha/2} \sqrt{\text{Var}(\log \widehat{OR}_{MH})}$$

ενώ για το ΛΣΠ δίνεται από το διάστημα

$$\left(e^{\log \widehat{OR}_{MH} - z_{1-\alpha/2} \sqrt{\text{Var}(\log \widehat{OR}_{MH})}}, e^{\log \widehat{OR}_{MH} + z_{1-\alpha/2} \sqrt{\text{Var}(\log \widehat{OR}_{MH})}} \right) .$$

Παράδειγμα 5.6 (αποτελέσματα ΙΙΙ).

Στο παράδειγμά μας έχουμε

$$\widehat{OR}_{MH} = \frac{120 \times 155/466 + 161 \times 124/532}{80 \times 111/466 + 130 \times 117/532} = 1.63.$$

Επιπλέον $\text{Var}(\log \widehat{OR}_{MH}) = 0.01646$ συνεπώς ένα 95% ΔΕ δίνεται από τον τύπο

$$\log 1.63 \pm 1.96 \sqrt{0.01646} \Leftrightarrow (0.234, 0.737).$$

Τέλος το 95% ΔΕ είναι ίσο με $(e^{0.234}, e^{0.737}) = (1.26, 2.09)$.

5.4.3 Εφαρμογή Στρωματοποιημένης Ανάλυσης με τη χρήση του SPSS.

Για να εφαρμόσουμε την παραπάνω ανάλυση στο SPSS (version 13) ακολουθούμε τα εξής βήματα

1. Analyze>Descriptives>Crosstabs: Από το αρχικό μενού επιλέγουμε Analyze>Descriptives>Crosstabs.
2. Σε κάθε κουτί της ανάλυσης Crosstabs εισάγουμε τις ακόλουθες μεταβλητές

- (α') ROWS \Rightarrow Cancer: Στο κουτί των γραμμών τοποθετούμε τη μεταβλητή που δηλώνει την παρουσία νόσου (στο παράδειγμα 5.6 είναι ο καρκίνος).
- (β') COLUMNS \Rightarrow Passive Smoker: Στο κουτί των στηλών τοποθετούμε τη μεταβλητή που δηλώνει τον παράγοντα κινδύνου ή την έκθεση στον κίνδυνο (στο παράδειγμα 5.6 είναι το παθητικό κάπνισμα).
- (γ') LAYER \Rightarrow Personal Smoking: Στο κουτί των στηλών τοποθετούμε το συγχυτικό παράγοντα (στο παράδειγμα 5.6 είναι το ατομικό κάπνισμα).

3. STATISTICS > Cochran's and Mantel-Haenszel Statistics : Επιλέγουμε το κουτί επιλογών STATISTICS στο κάτω μέρος του μενού Crosstabs και τσεκάρουμε τις επιλογές για τα στρωματοποιημένα δεδομένα: Cochran's and Mantel-Haenszel Statistics. Επίσης δίνεται και η περίπτωση να ελέγξουμε την υπόθεση $H_0 : OR_{MH} = \delta$ (δηλαδή ότι ο κοινός ΛΣΠ είναι ίσος με δ). Η τιμή που δίνει το πακέτο αν εμείς δεν παρέμβουμε είναι το ένα δηλαδή ανεξαρτησία μεταξύ νόσου-παράγοντα κινδύνου.

Τα αποτελέσματα δίνουν τρεις Πίνακες αποτελεσμάτων ανάλυση των οποίων ακολουθεί:

ΠΙΝΑΚΑΣ 1 Tests of Homogeneity of the Odds Ratio: Ο πίνακας αυτός δίνει δύο ελέγχους (των Breslow-Day και Tarone) για την ομοιογένεια - ισότητα των ΛΣΠ. Στο παράδειγμα 5.6 και οι δύο έλεγχοι δίνουν $\chi_{MH}^2 = 3.27$ και p-value=0.070 όποτε δεν απορρίπτουμε την υπόθεση ισότητας των ΛΣΠ για επίπεδο 5% (τα αποτελέσματα είναι όμοια με αυτά που βρήκαμε και εμείς στην αντίστοιχη ενότητα).

ΠΙΝΑΚΑΣ 2 Tests of Conditional Independence: Ο Πίνακας δίνει τους ελέγχους του Cochran και των Mantel-Haenszel για την υπό-συνθήκη ανεξαρτησία νόσου-παράγοντα κινδύνου. Οι δύο έλεγχοι δίνουν κοντινά αποτελέσματα. Για τον έλεγχο Mantel-Haenszel έχουμε $\chi_{MH}^2 = 13.94$ και p-value=0.000 (όμοια με τα αποτελέσματα της αντίστοιχης ενότητας) συνεπώς υπάρχει σχέση παθητικού καπνίσματος και νόσου ακόμα και αν λάβουμε υπόψη μας το ατομικό κάπνισμα.

ΠΙΝΑΚΑΣ 3 Mantel-Haenszel Common Odds Ratio Estimate: Ο Πίνακας αυτός δίνει τα στοιχεία για τον κοινό εκτιμητή ΛΣΠ. Τον πίνακα αυτό πρέπει να τον συμβουλευόμαστε μόνο όταν δεν απορρίπτουμε την υπόθεση ομοιογένειας των ΛΣΠ στον Πίνακα 1. Τα αποτελέσματα του Πίνακα 3 είναι τα ακόλουθα:

1. **Estimate:** Ο εκτιμητής του κοινού ΛΣΠ (\widehat{OR}_{MH})
2. **ln(Estimate):** Ο εκτιμητής του λογάριθμου του κοινού ΛΣΠ ($\log \widehat{OR}_{MH}$)

3. **Std. Error of ln(Estimate):** Το τυπικό σφάλμα του λογάριθμου του κοινού ΛΣΠ ($\sqrt{\text{Var}(\log \widehat{OR}_{MH})}$)
4. **Asymp. Sig. (2-sided):** Το p-value για τον έλεγχο $H_0 : \widehat{OR}_{MH} = \delta$ (η τιμή δ δίνεται στις επιλογές `Statistics` του μενού `Crosstabs`).
5. **Asymp. 95% Confidence Interval, Common Odds Ratio:** 95% διάστημα εμπιστοσύνης για τον κοινό ΛΣΠ.
6. **Asymp. 95% Confidence Interval, ln(Common Odds Ratio):** 95% διάστημα εμπιστοσύνης για το λογάριθμο του κοινού ΛΣΠ.

Στο παράδειγμα 5.6 παίρνουμε τα ακόλουθα αποτελέσματα

1. **Estimate:** 1.625 .
2. **ln(Estimate):** 0.486 .
3. **Std. Error of ln(Estimate):** 0.128 .
4. **Asymp. Sig. (2-sided):** 0.000 για $\delta = 1$ και 0.903 για $\delta = 1.6$.
5. **Asymp. 95% Confidence Interval, Common Odds Ratio:** 1.264 - 2.090 .
6. **Asymp. 95% Confidence Interval, ln(Common Odds Ratio):** 0.234 - 0.737 .

5.4.4 Εκτίμηση ΛΣΠ σε Μελέτες εξομοιωμένων ζευγών.

Υπάρχει άμεση σχέση της δοκιμασίας McNemar και της προσέγγισης των Mantel-Haenszel. Στις εξομοιωμένες μελέτες έχουμε ζευγάρια ατόμων με ίδια χαρακτηριστικά όσον αφορά τους συγχυτικούς παράγοντες. Άρα φανταστείτε τους 2×2 πίνακες που σχηματίζονται για κάθε εξομοιωμένο ζευγάρι δηλαδή $k = 1, 2, \dots, n$, όπου n είναι το σύνολο των εξομοιωμένων ζευγών. Έτσι θα πρέπει να υπολογίσουμε

$$\widehat{OR}_{MH} = \frac{\sum_{k=1}^n N_{11k}N_{22k}/N_{\bullet\bullet k}}{\sum_{k=1}^n N_{12k}N_{21k}/N_{\bullet\bullet k}}$$

όπου N_{ijk} είναι ο αριθμός των ατόμων σε κάθε ζευγάρι που αντιστοιχεί στο συνδυασμό i, j των μεταβλητών ομάδα/θεραπεία - νόσος για το k εξομοιωμένο ζευγάρι (χρησιμοποιούμε N_{ijk} για τα δεδομένα κάθε ζεύγους και n_{ij} για τα δεδομένα του συνολικού πίνακα των εξομοιωμένων τιμών).

Σε αυτή την περίπτωση έχουμε για τα ζευγάρια των ασθενών ομάδας/θεραπείας A και B τους ακόλουθους συνδυασμούς: έχουν τη νόσο και οι δύο ασθενείς (1,1), έχει τη νόσο μόνο ο A ασθενής (1,2), έχει τη νόσο μόνο ο B ασθενής (2,1) και κανένας από τους δύο δεν έχει τη νόσο (2,2). Έτσι έχουμε τον παρακάτω πίνακα

Ζευγάρι (A = i, B = j)	Ομάδα	ΝΟΣΟΣ		Πλήθος Ζευγών		
		ΝΑΙ (1)	ΟΧΙ (2)	$N_{11k}N_{22k}/N_{..k}$	$N_{11k}N_{22k}/N_{..k}$	n_{ij}
(1,1)	A (1)	1	0	$1 \times 0/2 = 0$	$1 \times 0/2 = 0$	n_{11}
	B (2)	1	0			
(1,2)	A (1)	1	0	$1 \times 1/2 = 1/2$	$0 \times 0/2 = 0$	n_{12}
	B (2)	0	1			
(2,1)	A (1)	0	1	$0 \times 0/2 = 0$	$1 \times 1/2 = 0$	n_{21}
	B (2)	1	0			
(2,2)	A (1)	0	1	$0 \times 1/2 = 0$	$0 \times 1/2 = 0$	n_{22}
	B (2)	0	1			

από τον οποίο προκύπτει ότι

$$\widehat{OR}_M H = \frac{0 \times n_{11} + \frac{1}{2}n_{12} + 0 \times n_{21} + 0 \times n_{22}}{0 \times n_{11} + 0 \times n_{12} + \frac{1}{2}n_{21} + 0 \times n_{22}} = n_{12}/n_{21}$$

όπου n_{ij} είναι το πλήθος των ζευγών των ατόμων των ομάδων A και B με αποτέλεσμα (νόσο) i και j αντίστοιχα. Συνεπώς

$$\widehat{OR} = \frac{n_{12}}{n_{21}}$$

ενώ η διακύμανση του λογαρίθμου του παραπάνω εκτιμητή δίνεται από τον τύπο

$$Var(\log \widehat{OR}) = \frac{1}{n_{12}} + \frac{1}{n_{21}}.$$

Χρησιμοποιώντας τα παραπάνω μπορούμε να υπολογίσουμε $100(1 - \alpha)\%$ διαστήματα εμπιστοσύνης για το λογάριθμο του ΛΣΠ από τον τύπο

$$\log \frac{n_{12}}{n_{21}} \pm z_{1-\alpha/2} \sqrt{\frac{1}{n_{12}} + \frac{1}{n_{21}}}$$

και για το ΛΣΠ από τον τύπο

$$\left(e^{\log \frac{n_{12}}{n_{21}} - z_{1-\alpha/2} \sqrt{\frac{1}{n_{12}} + \frac{1}{n_{21}}}}, e^{\log \frac{n_{12}}{n_{21}} + z_{1-\alpha/2} \sqrt{\frac{1}{n_{12}} + \frac{1}{n_{21}}}} \right).$$

Παράδειγμα 5.3 (Αποτελέσματα).

Γυρίζοντας στο παράδειγμα 5.3, Πίνακα 5.2.1 βρίσκουμε ότι ο ΛΣΠ είναι ίσος με $\widehat{OR} = 16/5 = 3.2$ συνεπώς «η σχετική πιθανότητα επιβίωσης για την ομάδα/θεραπεία A είναι (περίπου) τριπλάσια της αντίστοιχης σχετικής πιθανότητας για την ομάδα B».

Επιπλέον $Var(\log \widehat{OR}) = 1/5 + 1/16 = 0.2625$ ενώ το 95% ΔΕ για το $Var(\log OR)$ δίνεται από τον τύπο

$$\log 3.2 \pm 1.96\sqrt{0.2625} \Leftrightarrow (0.159, 2.170)$$

και το 95% ΔΕ για το OR δίνεται από τον τύπο

$$(e^{0.159}, e^{2.170}) \Leftrightarrow (1.17, 8.73) .$$

Παρόμοια είναι η διαδικασία αν πάρουμε τον αντίστροφο ΛΣΠ. Συνεπώς $\widehat{OR}^{-1} = 5/16 = 0.3125$ το οποίο σημαίνει ότι «η σχετική πιθανότητα επιβίωσης για την ομάδα/θεραπεία Β είναι περίπου 70% μικρότερη της αντίστοιχης σχετικής πιθανότητας για την ομάδα Α» ή εναλλακτικά «η σχετική πιθανότητα θανάτου για την ομάδα/θεραπεία Α είναι περίπου 70% μικρότερη της αντίστοιχης σχετικής πιθανότητας για την ομάδα Β». Το διαστήμα εμπιστοσύνης για την ποσότητα αυτή είναι ίσο με

$$(0.3125 \times e^{-1.96\sqrt{0.2625}}, 0.3125 \times e^{1.96\sqrt{0.2625}}) \Leftrightarrow (0.11, 0.85) .$$

Κεφάλαιο 6

Υπολογισμός Μεγέθους Δείγματος

6.1 Υπολογισμός Μεγέθους Δείγματος για 2 Διωνυμικά Ποσοστά

Για να υπολογίσουμε το απαιτούμενο μέγεθος δείγματος πρέπει να προκαθορίσουμε το επίπεδο σημαντικότητας α (συνήθως 0.05) και την ισχύ $(1 - \beta)$ του ελέγχου με την οποία εντοπίζει σωστά μια απόκλιση από την H_0 ίση με δ .

6.1.1 Ανεξάρτητα Δείγματα

Στην περίπτωση αυτή θέλουμε να ελέγξουμε την υπόθεση $H_0 : \pi_1 = \pi_2$ έναντι της εναλλακτικής $H_1 : \pi_1 \neq \pi_2$. Αν η H_0 είναι αληθής τότε δεν την απορρίπτουμε με πιθανότητα $1 - \alpha$ (από τον τρόπο κατασκευής του ελέγχου). Επιπλέον αν στην πραγματικότητα ισχύει ότι $H_1 : |\pi_1 - \pi_2| = \delta > 0$ τότε θα πρέπει να απορρίπτουμε την H_0 με πιθανότητα $1 - \beta$. Έτσι υπολογίζουμε το μέγεθος δείγματος το οποίο ικανοποιεί τη συνθήκη

$$P(\text{Απόρριψη } H_0 \mid |\pi_1 - \pi_2| = \delta) = 1 - \beta$$

και καταλήγουμε στον τύπο

$$n_1 = \frac{1}{\delta^2} \left(z_{1-\beta} \sqrt{\pi_1(1 - \pi_1) + \pi_2(1 - \pi_2)/k} + z_{1-\alpha/2} \sqrt{\pi(1 - \pi)(1 + 1/k)} \right)^2 \quad (6.1)$$

όπου $n_2 = kn_1$, π_1 και π_2 είναι τα ποσοστά επιτυχία σε κάθε δείγμα. Οι ποσότητες αυτές συνήθως βρίσκονται από τη βιβλιογραφία ή από μικρές πιλοτικές έρευνες. Επιπλέον το κοινό ποσοστό π δίνεται από τον τύπο $\pi = (\pi_1 + k\pi_2)/(1 + k)$.

Άσκηση 6.1 Να αποδειχθεί ο παραπάνω τύπος.

Άσκηση 6.2 Να βρεθεί το μέγεθος δείγματος για τον αντίστοιχο μονόπλευρο έλεγχο.

Παράδειγμα 6.1 Έστω ότι το ποσοστό εμφάνισης μίας νόσου είναι 150 ανά 100000 άτομα και θέλουμε να εντοπίσουμε μια μείωση του κινδύνου κατά $\delta = 20$ ποσοστιαίες μονάδες. Να βρεθεί το μέγεθος δείγματος αν υποθέσουμε ίσο αριθμό ατόμων στις 2 ομάδες.

Άσκηση 6.3 Στο παραπάνω παράδειγμα να βρεθεί το μέγεθος δείγματος αν η πρώτη ομάδα έχει το μισό αριθμό ατόμων από ότι η δεύτερη.

6.1.2 Εξαρτημένα Δείγματα

Στην περίπτωση αυτή θέλουμε να ελέγξουμε την υπόθεση $H_0 : \pi_D^A = 1/2$ έναντι της εναλλακτικής $H_1 : \pi_D^A \neq 1/2$, όπου $\pi_D^A = n_{21}/n_D$. Αν η H_0 είναι αληθής τότε δεν την απορρίπτουμε με πιθανότητα $1 - \alpha$ (από τον τρόπο κατασκευής του ελέγχου). Επιπλέον αν στην πραγματικότητα ισχύει ότι $H_1 : \pi_D^A = \pi_{1D}^A$ τότε θα πρέπει να απορρίπτουμε την H_0 με πιθανότητα $1 - \beta$. Έτσι υπολογίζουμε το μέγεθος δείγματος το οποίο ικανοποιεί τη συνθήκη

$$P(\text{Απόρριψη } H_0 | \pi_D^A = \pi_{1D}^A) = 1 - \beta$$

και καταλήγουμε στον τύπο

$$n_D = \frac{(2z_{1-\beta}\sqrt{\pi_{1D}^A(1 - \pi_{1D}^A)} + z_{1-\alpha/2})^2}{4(1/2 - \pi_{1D}^A)^2}.$$

Όμως από το παραπάνω μπορούμε να υπολογίσουμε το n_D και όχι το συνολικό μέγεθος του δείγματος. Έτσι μπορούμε να χρησιμοποιήσουμε μια εκτίμηση του ποσοστού των δυσαρμονικών ζευγών δηλαδή να θέσουμε $\pi_D = n_D/n$ συνεπώς

$$n = n_D/\pi_D = \frac{(2z_{1-\beta}\sqrt{\pi_{1D}^A(1 - \pi_{1D}^A)} + z_{1-\alpha/2})^2}{4(1/2 - \pi_{1D}^A)^2\pi_D}.$$

Προσοχή το παραπάνω νούμερο αναφέρεται σε αριθμό ζευγαριών άρα το πραγματικό μέγεθος δείγματος που θα πάρουμε στην πράξη είναι ίσο με $2n$.

Άσκηση 6.4 Να αποδειχθεί ο παραπάνω τύπος.

Άσκηση 6.5 Να βρεθεί ο αντίστοιχος τύπος για τους αντίστοιχους μονόπλευρους ελέγχους.

Παράδειγμα 6.2 Έστω ότι έχουμε 2 θεραπείες A και B σε μια εξομοιωμένη μελέτη. Από προηγούμενες μελέτες γνωρίζουμε ότι το ποσοστό συμφωνίας των δύο θεραπειών είναι ίσο με 85%. Επιπλέον θέλουμε να απορρίπτεται η H_0 με ισχύ 90% όταν η αναλογία μεταξύ των δύο τύπων ασυμφωνίας είναι 2 προς 1. Να βρεθεί ο αριθμός των ζευγών και το συνολικό μέγεθος του δείγματος.

6.1.3 Μέγεθος δείγματος σε Κλινικές Δοκιμές: Ανεξάρτητα Δείγματα, Δίτιμες Μεταβλητές με Αλλαγές στην Θεραπεία των Ασθενών

Οι παραπάνω υπολογισμοί υποθέτουν το μέγεθος δείγματος χωρίς να λαμβάνουν υπόψη τους ότι στην πράξη πολλοί ασθενείς δεν ακολουθούν ακριβώς τις οδηγίες του πρωτοκόλλου. Έτσι πολλοί ασθενείς έχουν ανατεθεί στην ομάδα θεραπείας τελικά μπορεί στην πραγματικότητα να μην λάβουν θεραπεία διότι πολύ απλά δεν ακολούθησαν τις απαιτούμενες οδηγίες (ξεχνούσαν δόσεις, δεν έπαιρναν τα φάρμακα στα σωστά χρονικά σημεία ή στις απαιτούμενες δόσεις). Αντίστοιχα υπάρχουν και οι περιπτώσεις (οι οποίες είναι λιγότερες) όπου κάποιος από την ομάδα ελέγχου λαμβάνει λανθασμένα να θεραπεία.

Ορισμός 6.1 Ρυθμός Εγκατάλειψης (*drop-out rate*) *θεραπείας είναι η αναλογία των ασθενών της ομάδας ενεργής θεραπείας που δε συμμορφώνονται με τους κανόνες του πρωτοκόλλου, δηλαδή αποτυγχάνουν να λάβουν σωστή και ολοκληρωμένη θεραπεία.*

Ορισμός 6.2 Ρυθμός Εισαγωγής (*drop-in rate*) *στην ομάδα θεραπείας είναι η αναλογία των ασθενών της ομάδας ελέγχου (μη ενεργής θεραπείας ή ψευδοφαρμάκου) που τελικά έλαβαν ενεργή θεραπεία.*

Χρησιμοποιούμε τους ακόλουθους συμβολισμούς

- β_1 : Ρυθμός Εγκατάλειψης Θεραπείας (*drop-out rate*).
- β_2 : Ρυθμός εισαγωγής στην ομάδα θεραπείας (*drop-in rate*).
- π_1 : Πιθανότητα ενός δυσμενούς γεγονότος (π.χ. θανάτου ή κάποιων συμπτωμάτων) για τα άτομα που όντως έλαβαν θεραπεία.
- π_2 : Πιθανότητα ενός δυσμενούς γεγονότος (π.χ. θανάτου ή κάποιων συμπτωμάτων) για τα άτομα που όντως δεν έλαβαν θεραπεία.
- π_1^* : Παρατηρούμενος κίνδυνος στην ομάδα θεραπείας.
- π_2^* : Παρατηρούμενος κίνδυνος στην ομάδα ελέγχου.

Εμείς στην πράξη συγκρίνουμε π_1^* και π_2^* διότι δεν μπορούμε να εντοπίσουμε τους ασθενείς που δε συμμορφώθηκαν με το πρωτόκολλο και άλλαξαν θεραπεία τα οποία δίνονται από τους τύπους:

$$\pi_1^* = P(\text{γεγονός}|\text{θεραπεία})$$

$$\begin{aligned}
&= P(\text{γεγονός, συμμόρφωση ασθενή|θεραπεία}) \\
&\quad + P(\text{γεγονός, μη συμμόρφωση ασθενή|θεραπεία}) \\
&= P(\text{γεγονός|συμμόρφωση ασθενή, θεραπεία})P(\text{συμμόρφωση ασθενή|θεραπεία}) \\
&\quad + P(\text{γεγονός|μη συμμόρφωση ασθενή, θεραπεία})P(\text{μη συμμόρφωση ασθενή|θεραπεία}) \\
&= \pi_1(1 - \hat{\eta}_1) + \pi_2\hat{\eta}_1.
\end{aligned}$$

Όμοια βρίσκουμε

$$\begin{aligned}
\pi_2^* &= P(\text{γεγονός|ομ.ελέγχου}) \\
&= P(\text{γεγονός, συμμόρφωση ασθενή|ομ.ελέγχου}) \\
&\quad + P(\text{γεγονός, μη συμμόρφωση ασθενή|ομ.ελέγχου}) \\
&= P(\text{γεγονός|συμμόρφωση ασθενή, ομ.ελέγχου})P(\text{συμμόρφωση ασθενή|ομ.ελέγχου}) \\
&\quad + P(\text{γεγονός|μη συμμόρφωση ασθενή, ομ.ελέγχου})P(\text{μη συμμόρφωση ασθενή|ομ.ελέγχου}) \\
&= \pi_2(1 - \hat{\eta}_2) + \pi_1\hat{\eta}_2.
\end{aligned}$$

Στα πλαίσια μιας κλινικής δοκιμής εμείς παρατηρούμε και εκτιμούμε τις ποσότητες π_1^* και π_2^* . Αν τις αφαιρέσουμε έχουμε

$$\pi_1^* - \pi_2^* = \pi_1(1 - \hat{\eta}_1) + \pi_2\hat{\eta}_1 - \pi_2(1 - \hat{\eta}_2) - \pi_1\hat{\eta}_2 = (\pi_1 - \pi_2)(1 - \hat{\eta}_1 - \hat{\eta}_2).$$

Η παραπάνω ποσότητα ονομάζεται και «αποδιδόμενος κίνδυνος υπό το καθεστώς μη συμμόρφωσης πρωτοκόλλου» (compliance-adjusted risk difference, βλ. Rosner, 1994, σελ. 390). Προφανώς όταν δεν υπάρχει πρόβλημα μη συμμόρφωσης των ασθενών στους κανόνες του πρωτοκόλλου τότε $\hat{\eta}_1 = \hat{\eta}_2 = 0$ οπότε και $\pi_1^* = \pi_1$, $\pi_2^* = \pi_2$.

Έστω ότι θέλουμε να εντοπίσουμε μια διαφορά $\delta = |\pi_1 - \pi_2|$. Εμείς βασιζόμαστε στον τύπο για δύο ανεξάρτητα δείγματα με $n_1 = n_2 = n$ (τύπος 6.1) οπότε

$$n = \frac{1}{\delta^{*2}} \left(z_{1-\beta} \sqrt{\pi_1^*(1 - \pi_1^*) + \pi_2^*(1 - \pi_2^*)} + z_{1-\alpha/2} \sqrt{2\pi^*(1 - \pi^*)} \right)^2.$$

Στον παραπάνω τύπο αντικαθιστούμε τις παρακάτω ισότητες

$$\begin{aligned}
\pi_1^* &= \pi_1(1 - \hat{\eta}_1) + \pi_2\hat{\eta}_1 \\
\pi_2^* &= \pi_2(1 - \hat{\eta}_2) + \pi_1\hat{\eta}_2 \\
\pi^* &= (\pi_1^* + \pi_2^*)/2 = (\pi_1 + \pi_2)/2 - (\hat{\eta}_1 - \hat{\eta}_2)(\pi_1 - \pi_2)/2 = \pi - (\hat{\eta}_1 - \hat{\eta}_2)(\pi_1 - \pi_2)/2 \\
\delta^* &= |\pi_1^* - \pi_2^*| = |\pi_1 - \pi_2|(1 - \hat{\eta}_1 - \hat{\eta}_2) = \delta(1 - \hat{\eta}_1 - \hat{\eta}_2)
\end{aligned}$$

και βρίσκουμε το μέγεθος δείγματος.

Για μικρούς ρυθμούς μη συμμόρφωσης ($\beta_1, \beta_2 \leq 0.10$) μπορούμε να θέσουμε $\pi_1^* \approx \pi_1$, $\pi_2^* \approx \pi_2$ και $\pi^* \approx \pi$ οπότε

$$n \approx \frac{\left(z_{1-\beta} \sqrt{\pi_1(1-\pi_1) + \pi_2(1-\pi_2)} + z_{1-\alpha/2} \sqrt{2\pi(1-\pi)}\right)^2}{\delta^2} \times \frac{1}{(1-\beta_1-\beta_2)^2}$$

δηλαδή

$$n_1^{(\text{μη συμ.})} = n_2^{(\text{μη συμ.})} \approx n^{(\text{απόλ. συμ.})} \times \frac{1}{(1-\beta_1-\beta_2)^2}.$$

όπου $n_1^{(\text{μη συμ.})}$, $n_2^{(\text{μη συμ.})}$ είναι τα μεγέθη κάθε ομάδας υπό το καθεστώς μη συμμόρφωσης στους κανόνες του πρωτοκόλλου και $n^{(\text{απόλ. συμ.})}$ είναι το μέγεθος κάθε ομάδας υπό το καθεστώς α-πόλυτης (ή πλήρους) συμμόρφωσης στους κανόνες του πρωτοκόλλου (για λεπτομέρειες βλέπε Rosner, 1994, σελ. 389-392). Η ποσότητα $\frac{1}{(1-\beta_1-\beta_2)^2}$ μπορεί να θεωρηθεί ως διόρθωση λόγω της μη συμμόρφωσης στους κανόνες του πρωτοκόλλου.

Παράδειγμα 6.3 Ένα ιατρικό κέντρο θέλει να διενεργήσει μία πενταετούς διάρκειας κλινική δοκιμή για να μελετήσει την πιθανή προστατευτική δράση της ασπιρίνης στην μυοκαρδιακή ανεπάρκεια (myocardial infraction). Η πιθανότητα εμφάνισης της νόσου σε κανονικό πληθυσμό είναι ίση με 0.5% ανά έτος. Επιπλέον, από τη βιβλιογραφία είναι γνωστό ότι η ασπιρίνη εμποδίζει την εμφάνιση του 20% των επεισοδίων. Υποθέτοντας ρυθμούς εγκατάλειψης θεραπείας β_1 και εισαγωγής στην ομάδα θεραπείας β_2 ίσους με 10% και 5% αντίστοιχα να βρεθεί το μέγεθος δείγματος που εντοπίζει την παραπάνω διαφορά με ισχύ 80% για έναν δίπλευρο έλεγχο με επίπεδο σημαντικότητας 5%.

Κεφάλαιο 7

Λογιστική Παλινδρόμηση (Logistic Regression)

7.1 Εισαγωγή

Στο κεφάλαιο 3 είδαμε πως μπορεί να αναλυθεί μια κατηγορική μεταβλητή, για παράδειγμα η παρουσία ή όχι μιας ασθένειας, σε σχέση με την έκθεση ή μη σε ένα παράγοντα κινδύνου. Είδαμε επίσης μέτρα κινδύνου, όπως ο αποδιδόμενος και σχετικός κίνδυνος και τον λόγο συμπληρωματικών πιθανοτήτων. Η μορφή αυτή στατιστικής ανάλυσης κυριαρχεί σε πολλές επιδημιολογικές μελέτες. Στην πραγματικότητα όμως, οι περισσότερες ασθένειες είναι σύνθετες και πολυπαραγοντικές. Αυτό σημαίνει πως πολύ σπάνια η παρουσία ή εξέλιξη μιας ασθένειας εξαρτάται μόνο από ένα παράγοντα κινδύνου.

Σε σύγχρονες επιδημιολογικές μελέτες καθώς και σε κλινικές δοκιμές, οι ερευνητές μελετούν και καταγράφουν μια σειρά από παραμέτρους που μπορεί να επηρεάζουν την εξέλιξη μιας νόσου, όπως δημογραφικά στοιχεία, κλινικά και γενετικά χαρακτηριστικά και άλλα. Σε μια μελέτη με σκοπό να αξιολογηθεί ο κίνδυνος επανεμφάνισης εμφράγματος, σε ασθενείς που έχουν ήδη υποστεί ένα επεισόδιο εμφράγματος, η διακοπή ή όχι του καπνίσματος αποτελεί ένα σημαντικό παράγοντα κινδύνου, παράλληλα όμως με την ηλικία του ασθενούς, τις διατροφικές συνήθειες, σωματική άσκηση καθώς και την πιθανή φαρμακευτική αγωγή που ακολουθεί ο ασθενής. Ως εκ τούτου, θα θέλαμε να υπολογίσουμε πως συνεπιδρούν, αλληλεπιδρούν και σχετίζονται αυτά τα χαρακτηριστικά, τόσο μεμονωμένα όσο και από κοινού, με την πιθανότητα εμφάνισης ενός δεύτερου εμφράγματος.

Για το σκοπό αυτό είναι απαραίτητη η χρήση πιο εξειδικευμένων εργαλείων στατιστικής ανάλυσης. Στην κλασική στατιστική, η χρήση μοντέλων γραμμικής παλινδρόμησης είναι η πιο

συνηθισμένη μέθοδος πολυμεταβλητής ανάλυσης. Σε βιοστατιστικές μελέτες, όταν η μεταβλητή απόκρισης είναι συνήθως κατηγορική, παρουσία ασθένειας ή όχι, τότε είναι απαραίτητη η γενίκευση του γραμμικού μοντέλου σε ένα μοντέλο λογιστικής παλινδρόμησης. Στο κεφάλαιο αυτό θα παρουσιάσουμε την μεθοδολογία εφαρμογής μοντέλων λογιστικής παλινδρόμησης, την ερμηνεία των παραμέτρων, όπως και την στρατηγική δημιουργίας του καταλληλότερου μοντέλου για τα δεδομένα. Θα μιλήσουμε επίσης για ελέγχους υποθέσεων και θα παρουσιάσουμε παραδείγματα ανάλυσης δεδομένων με τη χρήση στατιστικών πακέτων.

7.2 Η λογιστική παλινδρόμηση

7.2.1 Εισαγωγή

Θα ξεκινήσουμε με ένα κλασσικό παράδειγμα επιδημιολογικής μελέτης. Οι Dixon and Massey (1983) παρουσίασαν μια μελέτη για την εμφάνιση στεφανιαίας νόσου σε 200 άνδρες σε ένα νοσοκομείο του Λος Άντζελες. Για τα άτομα αυτά κατέγραψαν αν είχαν υποστεί κάποιο καρδιακό επεισόδιο κατά τη διάρκεια των τελευταίων 10 ετών (CNT=1) ή όχι (CNT=0) Επίσης ήταν γνωστά μια σειρά από άλλα χαρακτηριστικά όπως η ηλικία των ατόμων, μετρήσεις της συστολικής πίεσης και διαστολικής πίεσης του αίματος, τα επίπεδα χοληστερόλης, το εύρος και το βάρος των ατόμων. Τα δεδομένα είναι διαθέσιμα στο διαδίκτυο στη σελίδα [statlib](#). Μια κλασσική επιδημιολογική προσέγγιση θα προσπαθούσε να αναλύσει την πιθανότητα εμφάνισης στεφανιαίας νόσου σε σχέση με έναν παράγοντα κινδύνου, όπως π.χ. η ηλικία. Στη μελέτη αυτή η ηλικία μετριέται σε συνεχή κλίμακα αλλά ως υποθέσουμε πως μπορούμε να χωρίσουμε τα άτομα σε άνδρες ηλικίας κάτω των 40 ετών και άνω των 40. Τότε μπορούμε να φτιάξουμε τον παρακάτω 2×2 πίνακα και να υπολογίσουμε μέτρα κινδύνου.

Πίνακας 7.1: Συχνότητες εμφάνισης στεφανιαίας νόσου ανά ηλικιακή ομάδα

Ηλικία	Νόσος		
	0: Υγιής	1: Ασθενής	Σύνολο
≤ 40	86	6	92
> 40	88	20	108
Σύνολο	174	26	200

Ο σχετικός κίνδυνος εδώ είναι ίσος με $RR = 0.185/0.065 = 2.846$. Αυτό σημαίνει ότι ο κίνδυνος εμφάνισης στεφανιαίας νόσου σε άνδρες άνω των σαράντα ετών είναι ίσος με 2.84

φορές (ή 184% μεγαλύτερος) από τον αντίστοιχο κίνδυνο που έχουν οι άνδρες νεώτεροι των 40 ετών. Ο ΛΣΠ υπολογίζεται ως $OR = 86 \times 20/88 \times 6 = 3.25$. Δηλαδή, η σχετική πιθανότητα εμφάνισης της νόσου σε άνδρες άνω των σαράντα είναι 3.25 φορές τον ίδιο κίνδυνο που έχουν οι νεώτεροι.

Αν και αυτή η μορφή ανάλυσης παρέχει κάποιες πληροφορίες για τον κίνδυνο εμφάνισης της νόσου ανά ηλικία, ωστόσο είναι ελλιπής αφού δε λαμβάνονται υπόψιν άλλοι σημαντικοί παράγοντες που μπορεί να σχετίζονται με τη στεφανιαία νόσο, όπως μετρήσεις για τη συστολική και διαστολική πίεση των ανθρώπων αυτών, τα επίπεδα χοληστερόλης, τις διατροφικές συνήθειες και άλλες μεταβλητές που μπορεί να επηρεάζουν την υγεία των ατόμων που συμμετέχουν στη μελέτη. Επιπλέον, ο τρόπος που δημιουργήθηκε ο πίνακας αυτός βασίζεται στην αυθαίρετη διχοτόμηση μιας συνεχούς μεταβλητής. Υπάρχουν αρκετές μελέτες στη ιατρική βιβλιογραφία (Royston, Altman, Sauerbrei) που τονίζουν τα μειονεκτήματα και προβλήματα αυτής της στρατηγικής. Πρώτον, δεν υπάρχει καμία επιστημονική επιχειρηματολογία που να υπαγορεύει ότι ο διαχωρισμός των ατόμων πρέπει να γίνει στα 40 έτη. Αν αντί αυτής της ηλικίας επιλεγεί η ηλικία των πενήντα ετών τότε οι εκτιμήσεις αλλάζουν. Επιπλέον, η απώλεια πληροφορίας από την κατηγοριοποίηση αυτή είναι πολύ σημαντική. Ενώ υπάρχουν στοιχεία για την ηλικία των 200 ατόμων, η επιλογή να χωριστούν αυτά τα άτομα σε 2 μεγάλες ομάδες οδηγεί σε αφελή συμπεράσματα. Έτσι, ένας άνδρας 41 ετών θεωρείται ότι έχει 225% μεγαλύτερο κίνδυνο να πάθει στεφανιαία νόσο σε σχέση με έναν άνδρα που είναι μόλις ένα χρόνο νεώτερος του, ενώ έχει τον ίδιο ακριβώς κίνδυνο με ένα άνδρα που μπορεί να είναι 30 χρόνια μεγαλύτερος. Μια εναλλακτική προσέγγιση θα ήταν να χωρισθεί η συνεχής μεταβλητή της ηλικίας σε περισσότερες από δύο ομάδες, για παράδειγμα σε ομάδες εύρους 10 ετών. Με τον τρόπο αυτό η απώλεια πληροφορίας είναι μικρότερη, αλλά και πάλι υπάρχει.

Είναι φανερό λοιπόν, για να ξεπεραστούν αυτά τα προβλήματα θα ήταν χρήσιμη μια μορφή ανάλυσης που να είναι στην ίδια λογική με την ανάλυση παλινδρόμησης. Σε ένα γραμμικό μοντέλο η ανεξάρτητη μεταβλητή, στο παράδειγμα μας η ηλικία, μπορεί να χρησιμοποιηθεί ως συνεχής, ενώ ταυτόχρονα μπορούν να συμπεριληφθούν στο στατιστικό μοντέλο και άλλες μεταβλητές που πιθανών να επηρεάζουν την εξαρτημένη μεταβλητή. Σε ένα απλό μοντέλο γραμμικής παλινδρόμησης ο μέσος μιας συνεχούς απόκρισης y μπορεί να περιγραφεί ως μια σχέση με μια ανεξάρτητη μεταβλητή X (συνεχής ή κατηγορική) με την ακόλουθη σχέση:

$$E(y|X) = \beta_0 + \beta_1 X \quad (7.1)$$

Μια τέτοια μορφή αναπαράστασης θα ήταν πολύ χρήσιμη για να περιγραφεί η σχέση της ηλικίας με τον κίνδυνο εμφάνισης στεφανιαίας νόσου. Ωστόσο, η βασική υπόθεση που πρέπει να ισχύει για να έχει νόημα το γραμμικό μοντέλο (7.1) είναι ότι η εξαρτημένη μεταβλητή y ακολουθεί

κανονική κατανομή. Προφανώς, στο παράδειγμα μας, εξαρτημένη μεταβλητή είναι η εμφάνιση ή όχι στεφανιαίας νόσου και φυσικά μια δίτιμη μεταβλητή δεν μπορεί να ακολουθεί κανονική κατανομή. Έπειτα, αυτό που ενδιαφέρει τον ερευνητή εδώ είναι ο κίνδυνος εμφάνισης της νόσου, με άλλα λόγια η δεσμευμένη πιθανότητα $P(Y = 1|X)$ να εμφανιστεί η νόσος με βάση μια οποιαδήποτε έκθεση Q . Συνεπώς, θα χρειαζόταν ένα μοντέλο της μορφής:

$$P(Y = 1|X) = \beta_0 + \beta_1 X$$

Στη περίπτωση αυτή δύο είναι τα προβλήματα που πρέπει να αντιμετωπισθούν. Πρώτον, το στατιστικό πρόβλημα ότι η εξαρτημένη μεταβλητή δεν ακολουθεί κανονική κατανομή, και δεύτερον, το αριθμητικό πρόβλημα ότι το δεξί μέρος της εξίσωσης θα πρέπει να περιοριστεί να δίνει τιμές στο διάστημα $(0,1)$. Και τα δύο αυτά θέματα αντιμετωπίζονται με την χρήση της λογιστικής παλινδρόμησης, μιας μεθοδολογίας που ανήκει στα Γενικευμένα Γραμμικά Μοντέλα.

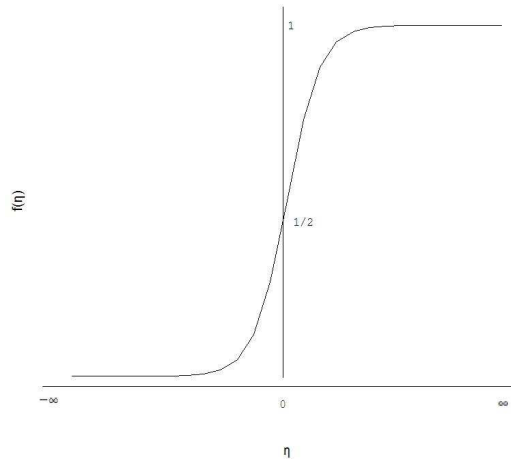
7.2.2 Η Λογιστική Συνάρτηση

Όπως είδαμε, το μοντέλο παλινδρόμησης που περιγράφει τον κίνδυνο εμφάνισης μιας ασθένειας θα πρέπει να δίνει τιμές μέσα στο διάστημα $(0,1)$. Θα πρέπει λοιπόν να χρησιμοποιηθεί ο κατάλληλος μαθηματικός μετασχηματισμός της δεξιάς πλευράς της εξίσωσης (7.1) έτσι ώστε οποιαδήποτε εκτίμηση και αν προκύπτει για κάποιον κίνδυνο αυτό ποτέ να μη βρίσκεται κάτω από το μηδέν ή πάνω από το ένα. Έστω $\eta = \beta_0 + \beta_1 X$ το συστηματικό μέρος της εξίσωσης. Η λογιστική συνάρτηση (logistic function) του η ορίζεται ως:

$$f(\eta) = \frac{1}{1 + \exp(-\eta)} \quad (7.2)$$

Η λογιστική συνάρτηση έχει μερικές πολύ χρήσιμες ιδιότητες για το λόγο αυτό κυριαρχεί στην ανάλυση κατηγορικών δεδομένων.

Η συνάρτηση έχει σιγμο-ειδή μορφή όπως φαίνεται στο διάγραμμα (7.2.2), όταν το $\eta = -\infty$ τότε $f(\eta) = 0$ ενώ όταν $\eta = \infty$ τότε $f(\eta) = 1$. Καθώς προχωράμε από το $\eta = -\infty$ η $f(\eta)$ βρίσκεται κοντά στο 0 έως ότου από κάποιο σημείο και έπειτα αυξάνει με γρήγορο ρυθμό. Προς το δεξί μέρος του διαγράμματος η συνάρτηση πλησιάζει προς τη μονάδα και εκεί ο ρυθμός αύξησης μειώνεται σημαντικά. Αυτή η χαρακτηριστική μορφή της λογιστικής παλινδρόμησης περιγράφει με άριστο τρόπο πολλά επιδημιολογικά φαινόμενα. Για παράδειγμα, ο κίνδυνος εμφάνισης καρδιακής νόσου είναι χαμηλός για τις μικρές ηλικίες και αυξάνει ανεπαίσθητα ως ότου ένας άνδρας ξεπεράσει ένα κατώφλι, την ηλικία που θεωρείται κρίσιμη για έμφραγμα. Από εκεί και έπειτα ο κίνδυνος αυξάνει δραματικά μέχρι το επόμενο κατώφλι και παραμένει σταθερά υψηλός στα τελευταία ηλικιακά χρόνια.



Σχήμα 7.1: Μορφή λογιστικής συνάρτησης

7.2.3 Το λογιστικό μοντέλο

Στην προηγούμενη ενότητα είδαμε τη λογιστική συνάρτηση $f(\eta) = 1/(1 + \exp(-\eta))$. Η τελευταία σχέση είναι ισοδύναμη με

$$P(Y = 1|X) = \frac{\exp(\eta)}{1 + \exp(\eta)} \quad (7.3)$$

Η σχέση (7.3) ορίζει το μοντέλο λογιστικής παλινδρόμησης. Το μοντέλο μπορεί να γραφεί ως λόγος σχετικών πιθανοτήτων στη μορφή:

$$\frac{P(Y = 1|X)}{1 - P(Y = 1|X)} = \exp(\beta_0 + \beta_1 X)$$

Παίρνοντας το λογάριθμο της αριστερής πλευράς της εξίσωσης καταλήγουμε στη σχέση

$$\log\left[\frac{P(Y = 1|X)}{1 - P(Y = 1|X)}\right] = \text{logit}(P(Y = 1|X)) = \beta_0 + \beta_1 X = \eta \quad (7.4)$$

Το μοντέλο αυτό συνδέει το λογάριθμο του λόγου σχετικών πιθανοτήτων εμφάνισης της ασθένειας γραμμικά με την ανεξάρτητη μεταβλητή Q . Για να δημιουργηθεί το μοντέλο αυτό λογιστικής παλινδρόμησης το συστηματικό μέρος της εξίσωσης η συνδέεται με τη χρήση της λογιστικής συνάρτησης για τον μετασχηματισμό της εξαρτημένης μεταβλητής. Για το λόγο αυτό η λογιστική συνάρτηση ονομάζεται εδώ *συνάρτηση δεσμού* (link function) και η $f(\cdot) = \exp(\cdot)/(1 + \exp(\cdot))$ είναι η αντίστροφη της συνάρτησης δεσμού.

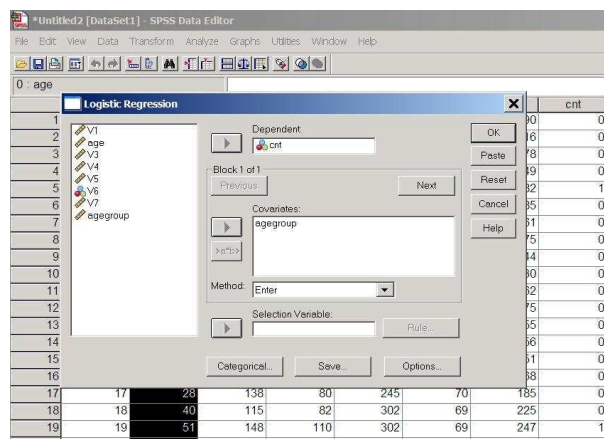
Το αριστερό μέρος του μοντέλου (7.3) ορίζει το στοχαστικό κομμάτι του υποδείγματος, την κατανομή της εξαρτημένης μεταβλητής y δεδομένης μιας παρατηρηθείσας τιμής για την ανεξάρτητη μεταβλητή x . Εφόσον η y εδώ είναι δίτιμη τότε ισχύουν οι ακόλουθες υποθέσεις για το μοντέλο λογιστικής παλινδρόμησης:

- Το y_i ακολουθεί διωνυμική κατανομή (όπου ο δείκτης i υποδηλώνει την i -οστή παρατήρηση)
- Ο μέσος $E(Y|X) = P(Y = 1|X)$ δίνεται από τη λογιστική συνάρτηση
- Οι τιμές της εξαρτημένης μεταβλητής είναι στατιστικά ανεξάρτητες

7.3 Ερμηνεία Παραμέτρων

7.3.1 Παλινδρόμηση με μια ανεξάρτητη κατηγορική μεταβλητή

Στον πίνακα (7.2.1) παρουσιάστηκαν οι συχνότητες εμφάνισης στεφανιαίας νόσου ανά ηλικιακή ομάδα και υπολογίστηκε ο λόγος σχετικών πιθανοτήτων. Η ίδια ανάλυση μπορεί να γίνει με την εφαρμογή ενός μοντέλου λογιστικής παλινδρόμησης και τη χρήση του πακέτου SPSS. Από το μενού **Analyze>Regression> Binary logistic** προκύπτει το παρακάτω παράθυρο :



Σχήμα 7.2: Παράθυρο διαλόγου λογιστικής παλινδρόμησης

Η μεταβλητή cnt που δηλώνει την παρουσία (1=ναι) ή όχι (0=όχι) στεφανιαίας νόσου επιλέγεται στο δεξί μέρος του παραθύρου ως εξαρτημένη Dependent ενώ η μεταβλητή agegroup εισέρχεται ως συμμεταβλητή (covariate) στο μοντέλο. Οι εκτιμήσεις των παραμέτρων υπολογίζονται από το πρόγραμμα και παρουσιάζονται στον παρακάτω πίνακα. Ο συντελεστής B στον

Variables in the Equation							
Step		B	S. E.	Wald	df	Sig.	Exp(B)
1 ^a	agegroup	1.181	.490	5.820	1	.016	3.258
	Constant	-2.663	.422	39.762	1	.000	.070

a. Variable(s) entered on step 1: agegroup.

Σχήμα 7.3: Εκτίμηση παραμέτρων λογιστικής παλινδρόμησης

πίνακα με το όνομα constant δίνει την εκτίμηση της σταθεράς β_0 του μοντέλου (ή τομής της ευθείας λογιστικής παλινδρόμησης), ενώ ο συντελεστής που αντιστοιχεί στην μεταβλητή agegroup είναι η εκτίμηση για την παράμετρο β_1 , την κλίση της ευθείας λογιστικής παλινδρόμησης. Η μεταβλητή agegroup, έχει δύο κατηγορίες, τα άτομα ηλικίας ως 40 ετών ($Q = 0$) και τα άτομα ηλικίας άνω των 40 ετών ($Q = 1$). Με τον τρόπο που έγινε η κωδικοποίηση προκύπτουν από το μοντέλο (7.4) δύο συναρτήσεις. Αν ένα άτομο ανήκει στην πρώτη κατηγορία (κάτω των 40 ετών) τότε ο λογαριθμικός λόγος σχετικών πιθανοτήτων να αναπτύξει στεφανιαία νόσο υπολογίζεται ως:

$$\log \frac{P(Y = 1|X = 0)}{P(Y = 1|X = 0)} = \hat{\beta}_0 + \hat{\beta}_1 X = \hat{\beta}_0 = -2.663$$

ενώ ο αντίστοιχος λόγος για άτομα ηλικίας άνω των 40 ετών είναι

$$\log \frac{P(Y = 1|X = 1)}{P(Y = 1|X = 0)} = \hat{\beta}_0 + \hat{\beta}_1 X = \hat{\beta}_0 + \hat{\beta}_1 = -2.663 + 1.181 = -1.482$$

Η διαφορά μεταξύ αυτών των δύο λόγων

$$\text{logit}\{P(Y = 1|X = 1)\} - \text{logit}\{P(Y = 1|X = 0)\} = 1.181$$

είναι φυσικά ο συντελεστής β_1 . Ο συντελεστής β_1 όπως υπολογίζεται από το μοντέλο, δίνει μια εκτίμηση του λογαριθμικού λόγου σχετικών πιθανοτήτων να νοσήσει ένα άτομο ηλικίας άνω των 40 ετών, σε σχέση με ένα άτομο νεότερης ηλικίας. Παρατηρήστε ότι η σχέση $\exp(1.181) = 3.258$ είναι ο λόγος σχετικός πιθανοτήτων, όπως ακριβώς είχε υπολογιστεί στην προηγούμενη ενότητα από τον 2×2 πίνακα συνάφειας.

Με βάση τις εκτιμήσεις που παρουσιάζονται στον πίνακα μπορούν να υπολογιστούν οι πιθανότητες εμφάνισης στεφανιαίας νόσου σε κάποια άτομα, ανάλογα με την ηλικιακή ομάδα στην οποία ανήκουν. Έτσι, για ένα άτομο ηλικίας 35 ετών η πιθανότητα εμφάνισης στεφανιαίας νόσου εκτιμάται ως:

$$P(Y = 1|X = 0) = \frac{\exp(-2.663)}{1 + \exp(-2.663)} = 0.065$$

ενώ για ένα άτομο ηλικίας 55 ετών η αντίστοιχη πιθανότητα γίνεται:

$$P(Y = 1|X = 1) = \frac{\exp(-2.663 + 1.181)}{1 + \exp(-2.663 + 1.181)} = 0.185$$

7.3.2 Παλινδρόμηση με μια ανεξάρτητη συνεχή μεταβλητή

Στην προηγούμενη ενότητα η συνεχής μεταβλητή ηλικία διχοτομήθηκε έτσι ώστε να δημιουργηθούν δύο διαφορετικές ηλικιακές ομάδες. Ωστόσο, θα ήταν ορθότερη τακτική να συμπεριληφθεί η ηλικία στο μοντέλο λογιστικής παλινδρόμησης ως συνεχής μεταβλητή. Στην περίπτωση αυτή οι εκτιμήσεις που προκύπτουν παρουσιάζονται στον πίνακα που ακολουθεί:

Variables in the Equation							
	B	S.E.	Wald	df	Sig.	Exp(B)	
Step 1 ^a	age	.063	.019	10.828	1	.001	1.065
	Constant	-4.793	.960	24.949	1	.000	.008

a. Variable(s) entered on step 1: age.

Σχήμα 7.4: Εκτίμηση παραμέτρων λογιστικής παλινδρόμησης

Με βάση το νέο μοντέλο που δημιουργήθηκε η εκτίμηση για την τομή της ευθείας (β_0) είναι -4.793 , ενώ για την κλίση β_1 είναι 0.063 . Ο συντελεστής β_1 εκφράζει στην περίπτωση αυτή τον κίνδυνο που έχει ένα άτομο κάποιας ηλικίας να αναπτύξει στεφανιαία νόσο, σε σχέση με ένα άτομο ηλικίας μικρότερης κατά ένα χρόνο. Ο αντίστοιχος λόγος σχετικών πιθανοτήτων για την αύξηση ηλικίας κατά ένα χρόνο υπολογίζεται αν πάρουμε το εκθετικό του β_1 , δηλαδή $\exp(0.063) = 1.065$. Αυτός ο λόγος δείχνει μια αύξηση του κινδύνου κατά 6.5% για εμφάνιση στεφανιαίας νόσου για κάθε χρόνο ηλικίας που περνά. Για παράδειγμα, ένα άτομο ηλικίας 45 ετών έχει κίνδυνο να αναπτύξει την ασθένεια:

$$P(Y = 1|X = 45) = \frac{\exp(-4.793 + 0.063 \times 45)}{1 + \exp(-4.793 + 0.063 \times 45)} = 0.123$$

Όπως είδαμε προηγουμένως, για κάθε ένα χρόνο ηλικίας που μεγαλώνει ένας ασθενής ο κίνδυνος αυξάνεται κατά περίπου 6.5% . Συνεπώς, ένα άτομο 46 ετών έχει (κατά προσέγγιση) κίνδυνο $0.123 + (0.123 \times 0.065) = 0.131$. Εναλλακτικά, ο κίνδυνος για ένα άτομο 46 ετών υπολογίζεται ως:

$$P(Y = 1|X = 46) = \frac{\exp(-4.793 + 0.063 \times 46)}{1 + \exp(-4.793 + 0.063 \times 46)} = 0.131$$

Μπορούμε επίσης να υπολογίσουμε τον κίνδυνο που αφορά μια διαφορά ηλικίας 10 ετών μεταξύ δύο ασθενών ως $\exp(0.063 \times 10) = 1.877$. Γενικότερα, για να υπολογισθεί ο λόγος σχετικών πιθανοτήτων για μια αύξηση κατά Δ μονάδες μιας συνεχούς μεταβλητής x χρησιμοποιείται ο τύπος:

$$\frac{P(Y = 1|X + \Delta)}{1 - P(Y = 1|X + \Delta)} / \frac{P(Y = 1|X)}{1 - P(Y = 1|X)} = \exp(\beta_1 \Delta)$$

7.3.3 Παλινδρόμηση με περισσότερες ανεξάρτητες μεταβλητές

Μέχρι τώρα είδαμε πως επηρεάζεται ο κίνδυνος εμφάνισης στεφανιαίας νόσου σε σχέση με την ηλικία των ανθρώπων που συμμετείχαν στη μελέτη. Ωστόσο, είναι γνωστό στη ιατρική βιβλιογραφία πως υπάρχουν μια σειρά από άλλους παράγοντες που μπορεί να αυξάνουν ή να μειώνουν τον κίνδυνο εμφάνισης της νόσου. Για παράδειγμα το κάπνισμα μπορεί να είναι ένας επιβαρυντικός παράγοντας, το βάρος ενός ανθρώπου, η συστολική ή διαστολική πίεση, η άσκηση, η χοληστερόλη και άλλα. Με τη χρήση ενός μοντέλου λογιστικής παλινδρόμησης

μπορεί να αναλυθεί πως μια σειρά από μετρήσεις μπορεί να συνεπιδρούν για την εμφάνιση μιας ασθένειας.

Για τα άτομα του παραδείγματος που αναλύσαμε μέχρι τώρα, είναι επίσης γνωστές μετρήσεις που έγιναν για το βάρος και τα επίπεδα χοληστερόλης. Ένας ερευνητής θα μπορούσε να αναλύσει τον κίνδυνο εμφάνισης στεφανιαίας νόσου σε άτομα με αυξημένα επίπεδα χοληστερόλης έναντι ατόμων που βρίσκονται μέσα στο φυσιολογικά όρια. Επίσης, να κάνει μια δεύτερη ανάλυση που να αφορά τη σχέση του βάρους με την στεφανιαία νόσο και μια τρίτη που να μελετά την ηλικία. Αντί για τρεις διαφορετικές αναλύσεις, σε κάθε μια από τις οποίες αδιαφορούμε για τις πληροφορίες που υπάρχουν πάνω σε κάθε μεταβλητή, είναι σκόπιμο να δημιουργηθεί ένα συνολικό μοντέλο το οποίο θα περιλαμβάνει όλες τις μεταβλητές που ενδιαφέρουν τον ερευνητή.

Ένα πολυμεταβλητό μοντέλο λογιστικής παλινδρόμησης με m ανεξάρτητες μεταβλητές έχει τη γενική μορφή:

$$\log\left[\frac{P(Y = 1|X_1, X_2, \dots, X_m)}{1 - P(Y = 1|X_1, X_2, \dots, X_m)}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m \quad (7.5)$$

Για λόγους μαθηματικής ευκολίας συνηθίζεται το μοντέλο να αναπαριστάται με τη μορφή πινάκων:

$$\log\left[\frac{P(Y = 1|\mathbf{X})}{1 - P(Y = 1|\mathbf{X})}\right] = \text{logit}[P(Y = 1|\mathbf{X})] = \beta\mathbf{X} \quad (7.6)$$

όπου X εδώ είναι ένας πίνακας διαστάσεων $n \times (p + 1)$, με n το πλήθος των παρατηρήσεων και p το πλήθος των ανεξάρτητων μεταβλητών που εισάγονται στο μοντέλο. Το διάνυσμα β είναι διαστάσεων $1 \times (p + 1)$ που περιλαμβάνει τις παραμέτρους που πρέπει να εκτιμηθούν από το μοντέλο.

Στο παράδειγμα με τις τρεις ανεξάρτητες μεταβλητές, αν συμβολίσουμε με X_1 τη μεταβλητή που περιέχει την ηλικία των ασθενών, X_2 την μεταβλητή βάρος και X_3 τη μεταβλητή που περιέχει τα επίπεδα χοληστερόλης, ο πίνακας \mathbf{X} γράφεται:

$$\mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & X_{31} \\ 1 & X_{12} & X_{22} & X_{32} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} & X_{3n} \end{bmatrix}$$

όπου το στοιχείο X_{21} δηλώνει την πρώτη παρατήρηση από τη δεύτερη μεταβλητή. Ο πίνακας \mathbf{X} έχει στην πρώτη στήλη μονάδες έτσι ώστε να συμπεριλαμβάνεται στο μοντέλο και ο σταθερός όρος, η τομή δηλαδή της εξίσωσης λογιστικής παλινδρόμησης. Το διάνυσμα β εδώ είναι $\beta = [\beta_0, \beta_1, \beta_2, \beta_3]$.

Για να εφαρμοστεί το πολυμεταβλητό μοντέλο λογιστικής παλινδρόμησης με τη χρήση του SPSS χρησιμοποιούνται οι ίδιες εντολές όπως πριν:

Analyze>Regression> Binary logistic και οι τρεις ανεξάρτητες μεταβλητές εισέρχονται στο δεξί μέρος του παραθύρου **Dependent** . Τα αποτελέσματα που προκύπτουν συνοψίζονται στον πίνακα που ακολουθεί:

Variables in the Equation							
	B	S.E.	Wald	df	Sig.	Exp(B)	
Step							
1 ^a	age	.053	.021	6.476	1	.011	1.054
	cholesterol	.007	.004	3.299	1	.069	1.007
	weight	.018	.008	4.496	1	.034	1.018
	Constant	-9.256	2.072	19.960	1	.000	.000

a. Variable(s) entered on step 1: age, cholesterol, weight.

Σχήμα 7.5: Εκτίμηση παραμέτρων πολυμεταβλητής λογιστικής παλινδρόμησης

Οι εκτιμήσεις των παραμέτρων που παρουσιάζονται στον πίνακα οδηγούν στα ακόλουθα συμπεράσματα:

- Ο λογαριθμικός λόγος σχετικών πιθανοτήτων που αφορά την ηλικία εκτιμάται πως είναι 0.053. Αυτό σημαίνει πως για κάθε ένα έτος που αυξάνει η ηλικία και αν οι άλλες μεταβλητές βάρος και χοληστερόλη παραμείνουν σταθερά, ο κίνδυνος εμφάνισης στεφανιαίας νόσου αυξάνει κατά 5.4% (διότι $\exp(0.053) = 1.054$).
- Αν το βάρος ενός ατόμου αυξηθεί κατά ένα κιλό, ενώ η ηλικία και τα επίπεδα χοληστερόλης παραμείνουν σταθερά, ο λόγος σχετικών πιθανοτήτων για την εμφάνιση της νόσου εκτιμάται ως $\exp(0.018) = 1.018$. Αυτό σημαίνει πως για κάθε πρόσθετο κιλό βάρους που λαμβάνει ένα άτομο και αν τα επίπεδα χοληστερόλης και η ηλικία παραμείνουν σταθερά ο κίνδυνος να νοσήσει από στεφανιαία νόσο αυξάνεται περίπου 2%
- Ο λόγος σχετικών πιθανοτήτων που σχετίζεται με την αύξηση των επιπέδων χοληστερόλης εκτιμάται $\exp(0.007) = 1.007$. Παρόλο που η αύξηση στον κίνδυνο δε φαίνεται να είναι σημαντική στην ουσία έχει πολύ μεγάλη σημασία. Η αύξηση κατά 0.7% στον κίνδυνο εμφάνισης της νόσου σχετίζεται με την αύξηση των επιπέδων χοληστερόλης κατά μια μονάδα. Ωστόσο, αν ένα άτομο βρίσκεται μέσα στα φυσιολογικά όρια των επιπέδων χοληστερόλης (λιγότερο από 200mg/dl) έστω 180mg/dl και στην επόμενη μέτρηση τα επίπεδα έχουν μετρηθεί πως είναι 240mg/dl (διαφορά δηλαδή 60mg/dl τότε ο ΛΣΠ είναι $\exp(0.007 * 60) = 1.522$ ήτοι αύξηση του κινδύνου 52.2%.

Με βάση τον πίνακα (7.5.1) οι πιθανότητες εμφάνισης στεφανιαίας νόσου μπορούν να υπολογιστούν με καλύτερη ακρίβεια (κάτω από την υπόθεση βέβαια ότι το μοντέλο είναι σωστά προσδιορισμένο, όπως θα δούμε στις επόμενες ενότητες). Με την ένταξη περισσότερων μεταβλητών σε ένα μοντέλο μπορούν να απαντηθούν πιο συγκεκριμένα ερωτήματα που ενδεχομένως να

αφορούν τους ερευνητές. Εδώ μπορούμε να υπολογίσουμε την πιθανότητα να ασθενήσει ένα άτομο ηλικίας 53 ετών, που έχει χοληστερόλη 173 mg/dl και ζυγίζει 85 κιλά:

$$P(Y = 1|\mathbf{X}) = \frac{\exp(-9.256 + 0.053 * 53 + 0.007 * 173 + 0.018 * 85)}{1 + \exp(-9.256 + 0.053 * 53 + 0.007 * 173 + 0.018 * 85)} = 0.024$$

ενώ η αντίστοιχη πιθανότητα για ένα άτομο 53 ετών που ζυγίζει 95 κιλά και έχει επίπεδα χοληστερόλης 230 mg/dl υπολογίζεται με αντίστοιχο τρόπο πως είναι περίπου 4.2%.

7.4 Μέθοδοι εκτίμησης

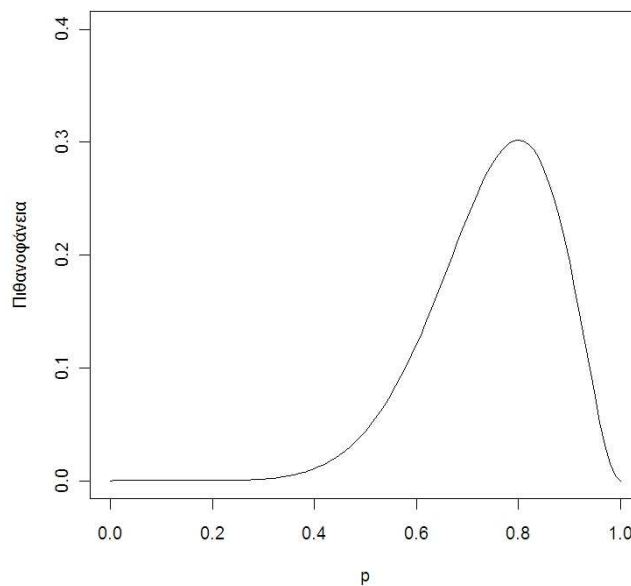
Το μοντέλο λογιστικής παλινδρόμησης ανήκει στη γενικότερη ομάδα μοντέλων που ονομάζεται Γενικευμένα Γραμμικά Μοντέλα (Generalized Linear Models, GLMs). Σε αντίθεση με την απλή γραμμική παλινδρόμηση, οι κανονικές εξισώσεις που προκύπτουν στα γενικευμένα γραμμικά μοντέλα δεν είναι απλές γραμμικές συναρτήσεις που έχουν απλή λύση. Συνεπώς, η μέθοδος των ελαχίστων τετραγώνων δε μπορεί να χρησιμοποιηθεί με τον ίδιο τρόπο όπως στα γραμμικά μοντέλα. Αντίθετα, η εκτίμηση των παραμέτρων απαιτεί μια επαναληπτική διαδικασία, έναν αλγόριθμο που μετά από μια σειρά βημάτων θα συγκλίνει σε μια συγκεκριμένη τιμή, την εκτίμηση των παραμέτρων του μοντέλου. Στα γενικευμένα γραμμικά μοντέλα δύο είναι οι βασικοί τρόποι εκτίμησης των παραμέτρων και βασίζονται στην συνάρτηση πιθανοφάνειας. Οι μέθοδοι αυτοί είναι η μέθοδος Newton-Raphson και η μέθοδος επαναλαμβανόμενων σταθμισμένων ελαχίστων τετραγώνων (Iterative Weighted Least Squares). Πριν όμως παρουσιαστούν οι αλγόριθμοι αυτοί ας ορίσουμε τη συνάρτηση μέγιστης πιθανοφάνειας.

7.4.1 Συνάρτηση πιθανοφάνειας

Η συνάρτηση πιθανοφάνειας L είναι μια συνάρτηση των αγνώστων παραμέτρων β που περιέχονται στο μοντέλο λογιστικής παλινδρόμησης και γράφεται $L(\beta; Y)$. Το β είναι (στη γενική του μορφή) ένα διάνυσμα γραμμή που περιλαμβάνει τις p άγνωστες παραμέτρους. Η συνάρτηση πιθανοφάνειας εκφράζει την από κοινού πιθανότητα (αλλιώς την πιθανοφάνεια) να παρατηρηθούν τα δεδομένα που έχουν προκύψει από μια συγκεκριμένη μελέτη. Για παράδειγμα, έστω μια μελέτη στην οποία επιτυχία θεωρείται η εμφάνιση Κορώνα (K) μετά τη ρίψη ενός νομίσματος. Έστω ότι μετά από 10 ρίψεις τα αποτελέσματα είναι 8 φορές κορώνα και 2 φορές γράμματα. Η πιθανοφάνεια εμφάνισης οκτώ επιτυχιών ορίζεται από τη διωνυμική κατανομή ως:

$$L(p; x = 8, n = 10) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

Η τιμή της πιθανοφάνειας του προβλήματος εξαρτάται από την τιμή της παραμέτρου p . Αν θεωρήσουμε λοιπόν ότι το νόμισμα είναι αμερόληπτο, τότε $p = 0.5$ και η τιμή της πιθανοφάνειας είναι 0.044. Με άλλα λόγια, η από κοινού πιθανότητα εμφάνισης οκτώ επιτυχιών σε 10 ρίψεις ενός αμερόληπτου νομίσματος είναι πολύ μικρή. Θα ήταν συνεπώς λογικό να υποθέσουμε ότι αφού η πιθανότητα των δεδομένων μας είναι μικρή, υπάρχουν ενδείξεις ότι το νόμισμα είναι μεροληπτικό. Συνεπώς, ίσως μια μεγαλύτερη τιμή για την άγνωστη παράμετρο p θα οδηγούσε σε αυτά τα αποτελέσματα. Η τιμή της πιθανοφάνειας αν $p = 0.7$ είναι 0.233 ενώ για $p = 0.8$ η πιθανοφάνεια γίνεται 0.302. Συνεπώς, τα δεδομένα του πειράματος θα ήταν πιο πιθανά για τιμή της παραμέτρου κοντά στο 0.8. Στην πράξη, δοκιμάζονται μια σειρά από ενδεχόμενες τιμές της παραμέτρου p έτσι ώστε να βρεθεί η τιμή που μεγιστοποιεί την συνάρτηση πιθανοφάνειας. Η τιμή αυτή ονομάζεται εκτιμητής μέγιστης πιθανοφάνειας της παραμέτρου p και συμβολίζεται με \hat{p} . Το διάγραμμα 7.4.1 παρουσιάζει τις τιμές που παίρνει η συνάρτηση πιθανοφάνειας για διαφορετικές τιμές της παραμέτρου p . Η συνάρτηση μεγιστοποιείται όταν $p = 0.8$. Στο



Σχήμα 7.6: Τιμές πιθανοφάνειας για διαφορετικές τιμές της παραμέτρου p .

σημείο αυτό πρέπει να τονιστεί η διαφορά μεταξύ της πιθανότητας που προκύπτει από μια διωνυμική συνάρτηση και της συνάρτησης πιθανοφάνειας. Αν γνωρίζουμε πως το νόμισμα είναι αμερόληπτο, αν δηλαδή είναι γνωστό πως $p = 0.5$ τότε με τον τύπο της διωνυμικής κατανομής μπορούμε να υπολογίσουμε την πιθανότητα να παρουσιαστούν οκτώ επιτυχίες σε δέκα ρίψεις. Η συνάρτηση πιθανοφάνειας ορίζει το αντίστροφο, δηλαδή, με δεδομένο πως σε δέκα ρίψεις

ενός νομίσματος παρουσιάστηκαν οκτώ επιτυχίες, ποια είναι η πιθανότητα το νόμισμα να είναι αμερόληπτο.

7.4.2 Μέθοδος μέγιστης πιθανοφάνειας

Σε ένα πρόβλημα λογιστικής παλινδρόμησης αυτό που ενδιαφέρει συνήθως είναι η πιθανότητα εμφάνισης μιας ασθένειας, η πιθανότητα θανάτου, η πιθανότητα υποτροπής, η πιθανότητα ίασης ενός ασθενούς κ.ο.κ. Κάθε μια από τις παρατηρήσεις (άτομα, ασθενής) μπορεί να "επιτύχει" ή "αποτύχει", συνεπώς για κάθε άτομο ξεχωριστά η πιθανότητα επιτυχίας ακολουθεί Bernoulli κατανομή. Κατά συνέπεια, ο αριθμός των επιτυχιών σε μια μελέτη ακολουθεί διωνυμική κατανομή με πιθανότητα p . Κάτω από το μοντέλο λογιστικής παλινδρόμησης, η πιθανότητα να εμφανιστεί το γεγονός (επιτυχία) ορίζεται, όπως είδαμε, ως $\pi(X) = P(Y = 1|X) = \frac{\exp(\beta X)}{1 + \exp(\beta X)}$, όπου X είναι ο πίνακας που περιλαμβάνει σε στήλες τις μεταβλητές και β το διάνυσμα των αγνώστων παραμέτρων. Αυτό που καθορίζει συνεπώς την πιθανότητα επιτυχίας είναι η ποσότητα β η οποία και θα πρέπει να εκτιμηθεί από τη μέθοδο μέγιστης πιθανοφάνειας.

Ας θεωρήσουμε πως η απόκριση y κωδικοποιείται ως 0 (αποτυχία) ή 1 (επιτυχία) και $y_i = 1$ όταν το i άτομο παρουσιάσει το γεγονός. Έστω επίσης X_i η i -οστή γραμμή του πίνακα X που περιγράφει τις τιμές των συμμεταβλητών για το i άτομο, και $\pi(X_i) = \frac{\exp(\beta X_i)}{1 + \exp(\beta X_i)}$ η πιθανότητα το i άτομο να "επιτύχει". Τότε η συνάρτηση πιθανοφάνειας L ορίζεται ως:

$$L(\beta X) = \prod_{i=1}^n \pi(X_i)^{y_i} (1 - \pi(X_i))^{1-y_i} \quad (7.7)$$

Με άλλα λόγια, ένα άτομο συνεισφέρει στην πιθανοφάνεια την ποσότητα $\pi(X_i)$ όταν εμφανίζει το γεγονός (όταν δηλαδή το $y = 1$) και $1 - \pi(X_i)$ όταν το $y = 0$. Η πιθανοφάνεια προκύπτει ως το γινόμενο των πιθανοτήτων αυτών πάνω σε όλα τα άτομα. Εφόσον η πιθανότητα επιτυχίας για κάθε άτομο ξεχωριστά είναι ανεξάρτητη από την πιθανότητα επιτυχίας των άλλων ατόμων, η από κοινού πιθανότητα ορίζεται ως το γινόμενο όλων αυτών των πιθανοτήτων.

Στην πράξη είναι προτιμότερο να χρησιμοποιείται η λογαριθμική πιθανοφάνεια κυρίως λόγω μαθηματικής ευκολίας. Η συνάρτηση λογαριθμικής πιθανοφάνειας συμβολίζεται με $\log L(\beta)$ και ορίζεται ως ο νεπέριος λογάριθμος της L .

$$\log L(\beta; X) = \ln[L(\beta; X)] = \sum_{i=1}^n y_i \ln(\pi(X_i)) + (1 - y_i) \ln(1 - \pi(X_i)) \quad (7.8)$$

Όπως είδαμε προηγουμένως $\pi = \exp(\beta X)/(1 + \exp(\beta X))$ ενώ προκύπτει επίσης ότι $(1 - \pi) = 1/(1 + \exp(\beta X))$. Από τη σχέση (7.8) υπολογίζονται τα scores της λογαριθμικής συνάρτησης πιθανοφάνειας ως:

$$U(\beta_b) = \frac{\partial \log L(\beta; X_i)}{\partial \beta_b}$$

με $b = 1, 2, \dots, p$ Παρουσιάζουμε τα ενδιάμεσα βήματα που χρειάζονται για τον καθορισμό των scores :

$$\frac{\partial \log(\pi(X_i))}{\partial \beta_b} = \frac{\partial \pi(X_i)}{\partial \beta_b} / \pi(X_i) \quad (7.9)$$

Αλλά

$$\begin{aligned} \frac{\partial \pi}{\partial \beta_b} &= \frac{\partial}{\partial \beta_b} \frac{\exp(\beta X_i)}{1 + \exp(\beta X_i)} \\ &= \frac{\frac{\partial \exp(\beta X_i)}{\partial \beta_b} (1 + \exp(\beta X_i)) - \exp(\beta X_i) \frac{\partial 1 + \exp(\beta X_i)}{\partial \beta_b}}{[1 + \exp(\beta X_i)]^2} \\ &= \frac{X_{ib} \exp(\beta X_i) (1 + \exp(\beta X_i)) - \exp(\beta X_i) X_{ib} \exp(\beta X_i)}{[1 + \exp(\beta X_i)]^2} \\ &= \frac{X_{ib} \exp(\beta X_i)}{[1 + \exp(\beta X_i)]^2} \end{aligned} \quad (7.10)$$

Συνεπώς από τις σχέσεις (7.9) και (7.10)

$$\frac{\partial \log(\pi)}{\partial \beta_b} = \frac{X_{ib} \exp(\beta X_i)}{[1 + \exp(\beta X_i)]^2} / \frac{\exp(\beta X_i)}{1 + \exp(\beta X_i)} = X_{ib} \frac{1}{1 + \exp(\beta X_i)} = X_{ib} (1 - \pi(X_i)) \quad (7.11)$$

Επίσης

$$\begin{aligned} \frac{\partial \log(1/1 + \exp(\beta X_i))}{\partial \beta_b} &= \frac{-X_{ib} \exp(\beta X_i)}{[1 + \exp(\beta X_i)]^2} / \frac{1}{1 + \exp(\beta X_i)} \\ &= -\frac{X_{ib} \exp(\beta X_i)}{1 + \exp(\beta X_i)} = -X_{ib} \pi(X_i) \end{aligned} \quad (7.12)$$

Τέλος, από τις (7.11 και 7.12) προκύπτει ότι τα scores είναι :

$$U(\beta_b) = \sum_{i=1}^n X_{ib} [y_i - p_i(X_i)] \quad (7.13)$$

Με τον ίδιο τρόπο ορίζεται ο πίνακας πληροφορίας I ως

$$I(\beta_{bq}) = \frac{\partial U(\beta_b)}{\partial \beta_q} = \sum_{i=1}^n -X_{iq} \frac{\exp(\beta X_i)}{1 + \exp(\beta X_i)} X_{iq} \quad (7.14)$$

Οι εκτιμητές μέγιστης πιθανοφάνειας για το διάνυσμα β προκύπτουν αν λυθεί επαναληπτικά η παρακάτω εξίσωση :

$$\beta^m = \beta^{m-1} + I(\beta^{m-1})^{-1} U(\beta^{m-1}) \quad (7.15)$$

7.4.3 Μέθοδος Επαναλαμβανόμενων Σταθμισμένων Ελαχίστων Τετραγώνων

Η μέθοδος μεγίστης πιθανοφάνειας συγκλίνει στην τελική εκτιμώμενη τιμή των παραμέτρων μετά έναν αριθμό βημάτων. Οι McCullag και Nelder (ΠΑΡΑΠΟΜΠΗ, σελίδα 40) πρότειναν τη χρήση ενός άλλου αλγορίθμου για να πάρουν τις ίδιες εκτιμήσεις που θα προέκυπταν με τη μεγιστοποίηση της πιθανοφάνειας. Η μέθοδος αυτή ονομάζεται *Μέθοδος επαναλαμβανόμενων σταθμισμένων ελαχίστων τετραγώνων* (Iterative Weighted Least Squares, ITWLS). Η μέθοδος αυτή είναι πιο απλή στην εφαρμογή της και στον προγραμματισμό, καθώς επίσης συγκλίνει στις εκτιμήσεις σε λιγότερο αριθμό βημάτων από αυτόν που απαιτείται από τη μέθοδο μεγίστης πιθανοφάνειας.

Έστω, για ένα μοντέλο λογιστικής παλινδρόμησης, ότι ορίζονται κάποιες αυθαίρετες αρχικές τιμές για τις παραμέτρους, που συμβολίζονται με $\tilde{\beta}$. Με τις τιμές αυτές υπολογίζεται η τιμή $\tilde{\eta}_i = X_i \tilde{\beta}$ και στη συνέχεια οι τιμές του $\tilde{\mu}_i = g^{-1}(\tilde{\eta}_i)$. Με βάση αυτές τις ποσότητες ορίζεται μια ενδιάμεση μεταβλητή z η οποία υπολογίζεται από τον τύπο :

$$z_i = \tilde{\eta}_i + (y_i - \tilde{\mu}_i) \frac{\partial \eta_i}{\partial \mu_i} \quad (7.16)$$

όπου $\frac{\partial \eta_i}{\partial \mu_i}$ είναι η μερική παράγωγος της συνάρτησης δεσμού ως προς το μ . Στη περίπτωση της λογιστικής παλινδρόμησης, η συνάρτηση δεσμού είναι

$$\eta_i = \log\left(\frac{\pi_i}{1 - \pi_i}\right) = \log\left(\frac{\mu_i}{n_i - \mu_i}\right)$$

αφού $\mu_i = n_i \pi_i$. Συνεπώς, η ενδιάμεση μεταβλητή γίνεται :

$$z_i = \tilde{\eta}_i + \frac{y_i - \tilde{\mu}_i}{\tilde{\mu}_i(n_i - \tilde{\mu}_i)} n_i$$

Στη συνέχεια, για να εκτιμηθεί το διάνυσμα των αγνώστων παραμέτρων, επαναλαμβάνεται η διαδικασία επίλυσης της εξίσωσης

$$\hat{\beta} = (X'WX)^{-1}X'Wz \quad (7.17)$$

όπου W είναι ένας διαγώνιος πίνακας που περιέχει τα σταθμά $w_i = \tilde{\mu}_i(n_i - \tilde{\mu}_i)/n_i$ (για τον τρόπο με τον οποίο προκύπτουν τα w χρησιμοποιείται θεωρία από την εκθετική οικογένεια κατανομών, δείτε McCullag και Nelder (ΠΑΡΑΠΟΜΠΗ, σελίδα 40)).

7.5 Έλεγχοι υποθέσεων

Το επόμενο βήμα μετά την εφαρμογή ενός μοντέλου στα δεδομένα και την εκτίμηση των σχετικών παραμέτρων, είναι η εφαρμογή στατιστικών ελέγχων και η εκτίμηση διαστημάτων εμπιστοσύ-

νης. Η διαδικασία αυτή έχει ως σκοπό να ελεγχθεί η καταλληλότητα του μοντέλου, δηλαδή η ικανότητα του να εφαρμόζει επαρκώς στα δεδομένα, έτσι ώστε να περιγράφεται με τον καλύτερο δυνατό τρόπο το πρόβλημα που μελετάται. Πολλές φορές το ζητούμενο είναι η καλή πρόβλεψη, με άλλα λόγια η ικανότητα του μοντέλου να εκτιμά την πραγματική πιθανότητα που έχει ένα άτομο να ασθενήσει εφόσον έχει συγκεκριμένα χαρακτηριστικά υγείας. Άλλες φορές, το στατιστικό μοντέλο είναι αυτό που μπορεί να υποδείξει τον βιολογικό μηχανισμό που κρύβεται πίσω από μια ασθένεια. Για παράδειγμα, ένα μοντέλο λογιστικής παλινδρόμησης μπορεί να αποκαλύψει αν ένα συγκεκριμένο γονίδιο σχετίζεται με την εμφάνιση μιας ασθένειας. Η λογική είναι απλή: αν δημιουργηθεί ένα μοντέλο το οποίο δεν περιλαμβάνει την επίδραση του γονιδίου και ένα δεύτερο που περιλαμβάνει την επίδραση του γονιδίου, μπορεί το μοντέλο που περιλαμβάνει την παραπάνω πληροφορία να μας πει περισσότερα για την εξαρτημένη μεταβλητή από το μοντέλο που δεν περιλαμβάνει το γονίδιο.

Εφόσον έχει υπολογισθεί η πιθανοφάνεια και οι τιμές των παραμέτρων ενός μοντέλου υπάρχουν τρεις βασικοί έλεγχοι που μπορούν να εφαρμοστούν, ο έλεγχος λόγου πιθανοφανειών, ο έλεγχος κατά Wald και ο έλεγχος των scores .

7.5.1 Έλεγχος λόγου πιθανοφανειών

Ο έλεγχος του λόγου των πιθανοφανειών βασίζεται στη σύγκριση των τιμών που παίρνει η συνάρτηση πιθανοφάνειας από δύο διαφορετικά μοντέλα. Απαραίτητη προϋπόθεση για να εφαρμοστεί ο έλεγχος είναι τα μοντέλα να εμπεριέχουν το ένα το άλλο (nested). Δηλαδή, έστω ένα μοντέλο m_1 το οποίο έχει στο συστηματικό μέρος μια ανεξάρτητη μεταβλητή, τη X_1 και ένα δεύτερο μοντέλο m_2 με περισσότερες παραμέτρους. Για να γίνει ο έλεγχος, το μεγαλύτερο μοντέλο (m_2) θα πρέπει να περιλαμβάνει στο συστηματικό του μέρος την μεταβλητή X_1 . Έστω λοιπόν τα μοντέλα m_1 και m_2 με $m_1 \subset m_2$ και $L(\hat{\beta}_{m_1}, y)$ η πιθανοφάνεια του πρώτου μοντέλου, όπου $\hat{\beta}_{m_1}$ είναι η εκτίμηση μέγιστης πιθανοφάνειας του διανύσματος των παραμέτρων κάτω από το μοντέλο m_1 , και $L(\hat{\beta}_{m_2}, y)$ η αντίστοιχη πιθανοφάνεια που υπολογίζεται για το μοντέλο m_2 . Η αναλογία αυτών των δύο ποσοτήτων

$$\hat{\lambda} = \frac{L(\hat{\beta}_{m_1}, y)}{L(\hat{\beta}_{m_2}, y)} \quad (7.18)$$

παίρνει τιμές ανάμεσα στο 0 και το 1. Τιμές κοντά στο μηδέν δείχνουν ότι το μικρότερο μοντέλο δεν είναι αποδεκτό σε σχέση με το μεγαλύτερο μοντέλο, γιατί τότε οι παρατηρούμενες τιμές θα ήταν λιγότερο πιθανές. Από την άλλη, τιμές κοντά στη μονάδα δείχνουν ότι το μικρό μοντέλο είναι (σχεδόν) το ίδιο καλό με το μεγαλύτερο.

Μπορεί να αποδειχθεί ότι η ποσότητα μείον δύο φορές του λόγου πιθανοφανειών ακολουθεί (σε μεγάλα δείγματα) X^2 κατανομή, με βαθμούς ελευθερίας την διαφορά ανάμεσα στον αριθμό

των παραμέτρων των δύο μοντέλων υπό σύγκριση :

$$-2 \log \hat{\rho} = 2 \log L(\hat{\beta}_{m_2}, y) - 2 \log L(\hat{\beta}_{m_1}, y) \sim X^2_\nu \quad (7.19)$$

όπου ν είναι οι διαστάσεις του μεγαλύτερου μοντέλου μείον τις διαστάσεις του μικρότερου.

Το κριτήριο του λόγου πιθανοφανειών μπορεί να εφαρμοστεί σε οποιοδήποτε μοντέλο εκτιμάται με τη μέθοδο μέγιστης πιθανοφάνειας. Ωστόσο, στα γενικευμένα γραμμικά μοντέλα (όπως η λογιστική παλινδρόμηση), ο έλεγχος μπορεί να κατασκευασθεί αν ορισθεί μια ιδιαίτερα σημαντική στατιστική συνάρτηση, η συνάρτηση της *αποκλίουσας* (deviance). Έστω ένα θεωρητικό μοντέλο m_s που περιλαμβάνει μια παράμετρο για κάθε παρατήρηση. Το μοντέλο αυτό ονομάζεται *κορεσμένο* (saturated). Έστω επίσης ένα μικρότερο μοντέλο m το οποίο θα θέλαμε να συγκρίνουμε με το κορεσμένο. Το κριτήριο του λόγου πιθανοφανειών γίνεται :

$$\hat{\rho} = \frac{L(\hat{\beta}_m, y)}{L(\hat{\beta}_{m_s}, y)}$$

και η ποσότητα

$$D = 2 \log L(\hat{\beta}_{m_s}, y) - 2 \log L(\hat{\beta}_m, y) \quad (7.20)$$

ονομάζεται αποκλίουσα. Στην περίπτωση της λογιστικής παλινδρόμησης, η συνάρτηση πιθανοφάνειας όπως αυτή ορίστηκε στην εξίσωση (7.8) για το κορεσμένο μοντέλο γίνεται :

$$\log L(\beta; X) = \sum_{i=1}^n y_i \ln(y_i) + (1 - y_i) \ln(1 - y_i)$$

διότι ένας εκτιμητής μέγιστης πιθανοφάνειας για το μοντέλο με αριθμό παραμέτρων όσος και ο αριθμός των παρατηρήσεων θα ήταν $\hat{\mu}_i = y_i$. Άρα, η συνάρτηση της αποκλίουσας για το διωνυμικό μοντέλο προκύπτει από τις εξισώσεις (7.8) και (7.20):

$$D = 2 \sum_{i=1}^n \left[y_i \ln\left(\frac{y_i}{\pi(\hat{X}_i)}\right) + (1 - y_i) \ln\left(\frac{1 - y_i}{1 - \pi(\hat{X}_i)}\right) \right] \quad (7.21)$$

Είναι σύνηθες να υπολογίζονται οι αποκλίουσες διαφορετικών μοντέλων και στη συνέχεια να γίνεται η σύγκριση χρησιμοποιώντας αυτές τις τιμές. Η διαφορά ανάμεσα στις αποκλίουσες nested μοντέλων ακολουθεί X^2 κατανομή με ν βαθμούς ελευθερίας, όπως ακριβώς ορίστηκαν πριν στον έλεγχο του λόγου πιθανοφανειών.

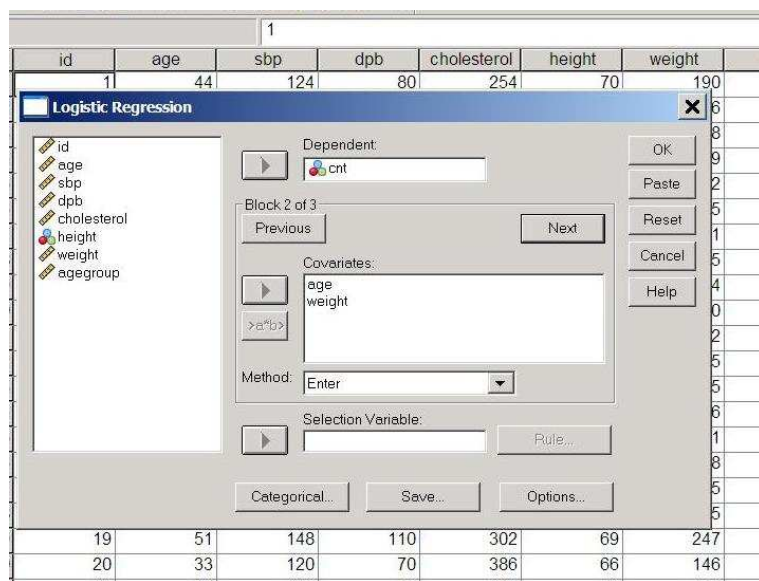
Η συνάρτηση της αποκλίουσας έχει ιδιαίτερη σημασία. Με απλά λόγια, η αποκλίουσα είναι ένας δείκτης της απόστασης των εκτιμηθέντων τιμών (αυτών που προβλέπει το μοντέλο) από τις παρατηρούμενες, όπως ακριβώς είναι το τετραγωνικό άθροισμα των σφαλμάτων στην απλή γραμμική παλινδρόμηση ($SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$). Επίσης, η αποκλίουσα χρησιμοποιείται για

ελέγχους καλής προσαρμογής του μοντέλου, καθώς επίσης για να ελεγχθεί η ύπαρξη υπερδιασποράς (overdispersion).

ΠΑΡΑΔΕΙΓΜΑ Είδαμε προηγουμένως πως η ηλικία αυξάνει τον κίνδυνο εμφάνισης στεφανιαίας νόσου, χρησιμοποιώντας ένα μοντέλο που περιλαμβάνει μια ανεξάρτητη μεταβλητή, το οποίο εδώ ονομάζουμε m_1 . Έπειτα, εφαρμόστηκε ένα μοντέλο με τρεις ανεξάρτητες μεταβλητές, την ηλικία, το βάρος και τα επίπεδα χοληστερόλης (έστω m_3 . Έστω και ένα ενδιάμεσο μοντέλο το οποίο περιλαμβάνει δύο ανεξάρτητες μεταβλητές, την ηλικία και το βάρος (m_2). Το μεγαλύτερο μοντέλο m_3 εμπεριέχει τα m_2 και m_1 , ενώ το m_2 εμπεριέχει το m_1 ($m_1 \subset m_2 \subset m_3$). Για να αποφασίσουμε ποιο μοντέλο από τα τρία είναι προτιμότερο για τα δεδομένα μπορούμε να χρησιμοποιηθεί ο έλεγχος πιθανοφανειών, ή ισοδύναμα να συγκριθούν οι αποκλίνοσες από τα τρία μοντέλα.

Το SPSS εφαρμόζει αυτόματα τον έλεγχο των πιθανοφανειών, μεταξύ διαφορετικών μοντέλων, αν αυτά οριστούν σαν διαφορετικά blocks. Τα βήματα είναι τα ακόλουθα:

Analyze > Regression > Binary logistic



Σχήμα 7.7:

στο παράθυρο διαλόγου ορίζεται η εξαρτημένη μεταβλητή CNT και η ηλικία ως ανεξάρτητη. Έπειτα, πατώντας την επιλογή NEXT ορίζεται το δεύτερο μοντέλο, με την ίδια ανεξάρτητη μεταβλητή και ανεξάρτητες τη φορά αυτή την ηλικία και το βάρος. Τέλος, ορίζεται ένα τρίτο block με τις τρεις ανεξάρτητες μεταβλητές. Η διαδικασία αυτή είναι σα να εφαρμόζονται τρία διαφορετικά μοντέλα γραμμικής παλινδρόμησης, ωστόσο αν οριστούν ως block τότε εμφανίζονται και οι

έλεγχοι μεταξύ των διαφορετικών μοντέλων. Το παράθυρο του *outruttwora* περιλαμβάνει τους πίνακες (7.3.2 και 7.5.1) για τα μοντέλα m_1 και m_3 , αντίστοιχα, καθώς και έναν πίνακα που παραθέτει τις εκτιμήσεις για τις παραμέτρους του ενδιάμεσου μοντέλου.

Το ενδιαφέρον εδώ στρέφεται στον πίνακα που περιλαμβάνει την τιμή της αποκλίνουσας (για κάθε μοντέλο όπως αυτά ορίζονται από τα διαφορετικά block) καθώς και στον πίνακα που παρουσιάζει την διαφορά της αποκλίνουσας του μοντέλου από ένα μηδενικό μοντέλο (το μοντέλο που δεν περιλαμβάνει καμία ανεξάρτητη μεταβλητή). Στο πρώτο βήμα, παρουσιάζεται η αποκλίνουσα του μοντέλου με την ηλικία ως ανεξάρτητη μεταβλητή στον παρακάτω πίνακα :

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	142.742 ^a	.057	.107

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Omnibus Tests of Model Coefficients

Step 1		Chi-square	df	Sig.
	Step	11.813	1	.001
	Block	11.813	1	.001
	Model	11.813	1	.001

Σχήμα 7.8: Πίνακας αποκλίνουσας και διαφοράς αποκλινουσών

Η τιμή της αποκλίνουσας είναι 142.742 (-2Loglikelihood) που σημαίνει ότι η πιθανοφάνεια του μοντέλου θα είναι -71.371. Ο δεύτερος πίνακας παρουσιάζει τη διαφορά της αποκλίνουσας από το μηδενικό μοντέλο :

$$\hat{\lambda} = 2 \log L(\hat{\beta}_1; y) - 2 \log L(\hat{\beta}_0; y) = 11.813$$

Αν αυτή η διαφορά συγκριθεί με μια X^2 κατανομή με ένα βαθμό ελευθερίας τότε προκύπτει η κρίσιμη τιμή του ελέγχου $\text{Sig}=0.001$. Συνεπώς, η διαφορά των μοντέλων είναι στατιστικά σημαντική, το μοντέλο με την ηλικία ως ανεξάρτητη μεταβλητή εφαρμόζει καλύτερα στα δεδομένα από ένα μοντέλο που δεν εμπεριέχει καμία πληροφορία. Από τον πίνακα φαίνεται επίσης πως η λογαριθμική πιθανοφάνεια του μηδενικού μοντέλου είναι -77.227.

Στον επόμενο πίνακα παρουσιάζονται οι αντίστοιχες ποσότητες που υπολογίστηκαν για το μοντέλο με την ηλικία και το βάρος ως ανεξάρτητες μεταβλητές. Η αποκλίνουσα του μοντέλου m_2 είναι τώρα μικρότερη (138,768)εφόσον το μοντέλο περιλαμβάνει περισσότερες παραμέτρους. Ο δεύτερος πίνακας παρουσιάζει την διαφορά της αποκλίνουσας από το προηγούμενο μοντέλο m_1

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	138.768 ^a	.076	.141

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	3.974	1	.046
	Block	3.974	1	.046
	Model	15.787	2	.000

Σχήμα 7.9: Πίνακας αποκλίνουσας και διαφοράς αποκλινουσών

(Block Chi square =3.974) καθώς και την διαφορά της αποκλίνουσας του m_2 από την αντίστοιχη του μηδενικού (Model Chi square =15.786). Η τελευταία διαφορά των 15.786 μονάδων είναι στατιστικά σημαντική αν συγκριθεί με την τιμή που δίνει η χ -τετράγωνο κατανομή με δύο βαθμούς ελευθερίας. Συνεπώς το μοντέλο με δύο ανεξάρτητες μεταβλητές εφαρμόζει καλύτερα στις παρατηρήσεις από ότι το μηδενικό. Επιπλέον, η διαφορά 3.974 μονάδων μεταξύ των δύο ενδιάμεσων μοντέλων είναι επίσης στατιστικά σημαντική αν συγκριθεί με την τιμή της χ -τετράγωνο κατανομής με ένα βαθμό ελευθερίας. Άρα το μοντέλο που περιλαμβάνει και το βάρος ως ανεξάρτητη μεταβλητή είναι καλύτερο από το μοντέλο που περιέχει μόνο την ηλικία ως προγνωστικό παράγοντα.

Τέλος, παρουσιάζεται ο πίνακας και για το τρίτο βήμα, που αφορά τη σύγκριση του m_3 με το m_2 και το μηδενικό μοντέλο. Σύμφωνα με τις εκτιμήσεις το μοντέλο που περιλαμβάνει τρεις ανεξάρτητες μεταβλητές είναι καλύτερο από το απλό μηδενικό μοντέλο (Model Chi square = 19.031). Ωστόσο, η διαφορά της αποκλίνουσας του m_3 με το m_2 δεν είναι αρκετά μεγάλη για να προτιμηθεί το μοντέλο με τις τρεις μεταβλητές έναντι του πιο απλού μοντέλου με τις δύο. Αν και υπάρχει βελτίωση στην τιμή της αποκλίνουσας η διαφορά αυτή δεν είναι στατιστικά σημαντική.

Η σύγκριση του μοντέλου με την ηλικία μόνο ως ανεξάρτητη μεταβλητή με το μοντέλο με τρεις μεταβλητές είναι εύκολο να γίνει με το χέρι. Γνωρίζουμε ότι η αποκλίνουσα του m_1 είναι 142.742 ενώ η αποκλίνουσα του m_3 είναι 135.523. Συνεπώς η διαφορά τους 7.219 πρέπει να συγκριθεί με την τιμή της χ -τετράγωνο κατανομής με δύο βαθμούς ελευθερίας $\chi^2 = 5.99$ (σε επίπεδο στατιστικής σημαντικότητας $\alpha=5\%$). Το συμπέρασμα είναι πως το μοντέλο με τρεις μεταβλητές εφαρμόζει καλύτερα στα δεδομένα από το μοντέλο που περιλαμβάνει μόνο την ηλικία.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	135.523 ^a	.091	.169

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	3.245	1	.072
	Block	3.245	1	.072
	Model	19.031	3	.000

Σχήμα 7.10: Πίνακας αποκλίνουσας και διαφοράς αποκλινουσών

7.5.2 Έλεγχοι Wald

Ο έλεγχος κατά τον Wald χρησιμοποιείται όταν θέλουμε να ελέγξουμε την υπόθεση ότι το διάνυσμα των παραμέτρων β είναι ίσο με το μηδέν. Δηλαδή, η αρχική υπόθεση είναι

$$H_0 : \beta = 0$$

Αποδεικνύεται, για τη λογιστική παλινδρόμηση, ότι το διάνυσμα που περιέχει τους εκτιμητές μεγίστης πιθανοφάνειας $\hat{\beta}$ των πραγματικών παραμέτρων β ακολουθεί πολυδιάστατη κανονική κατανομή (με διαστάσεις όσες και οι άγνωστες παράμετροι που εκτιμούνται) με μέση τιμή ίση με την πραγματική τιμή των παραμέτρων και πίνακα διακύμανσης-συνδιακύμανσης που προκύπτει από τον αντίστροφο του πίνακα πληροφορίας, δηλαδή:

$$\hat{\beta} \sim N_p(\beta, I^{-1}(\beta)) \quad (7.22)$$

Με αφετηρία το συμπέρασμα αυτό προκύπτει ότι η τετραγωνική μορφή

$$W = (\hat{\beta} - \beta_0)' I^{-1}(\hat{\beta}) (\hat{\beta} - \beta_0) \quad (7.23)$$

ακολουθεί, κάτω από την μηδενική υπόθεση, X^2 κατανομή με p βαθμούς ελευθερίας, όπου p είναι το μήκος του διαστήματος β .

Ο έλεγχος του Wald χρησιμοποιείται συνήθως για να ελέγξουμε αν ένα μέρος των διανύσματος των παραμέτρων είναι ίσο με το μηδέν. Για παράδειγμα, για το μοντέλο m_1 ο πίνακας 7.3.2 περιλαμβάνει την τιμή του Wald στατιστικού στην τρίτη στήλη, μαζί με τους βαθμούς ελευθερίας που αντιστοιχούν στον έλεγχο και την κρίσιμη τιμή (p-value). Για τον έλεγχο λοιπόν

$H_0 : \beta_{age} = 0$ η τιμή της στατιστικής W είναι 10.828, μεγαλύτερη από την τιμή της X^2 κατανομής με ένα βαθμό ελευθερίας ($X^2_{1,\alpha=5\%} = 3.84$). Συνεπώς, υπάρχουν ισχυρές ενδείξεις υπέρ της απόρριψης της αρχικής υπόθεσης. Με άλλα λόγια η επίδραση της μεταβλητής ηλικία, έχει στατιστικά σημαντική σχέση με τον κίνδυνο εμφάνισης στεφανιαίας νόσου. Το ίδιο προκύπτει και από την κρίσιμη τιμή του ελέγχου που υπολογίζεται από το SPSS Sig.=0.001.

Στο μοντέλο με μία ανεξάρτητη μεταβλητή ο έλεγχος του Wald είναι ισοδύναμος με τον έλεγχο του λόγου πιθανοφάνειων. Σε γενικές γραμμές, αν το μέγεθος του δείγματος είναι αρκετό, τότε οι δύο έλεγχοι θα πρέπει να δίνουν το ίδιο αποτέλεσμα. Διαφορετικά, ο έλεγχος λόγου πιθανοφάνειων είναι πιο ισχυρός.

Για το μοντέλο m_3 οι εκτιμήσεις των παραμέτρων παρουσιάζονται στον πίνακα 7.5.1. Για κάθε μια από τις τρεις μεταβλητές που έχουν συμπεριληφθεί στο μοντέλο, έχουν υπολογισθεί οι τιμές του ελέγχου και οι αντίστοιχες κρίσιμες τιμές. Σύμφωνα λοιπόν με τις εκτιμήσεις οδηγούμαστε στο συμπέρασμα πως δεν υπάρχουν αρκετές ενδείξεις ότι η μεταβλητή που μετρά τα επίπεδα χοληστερόλης επηρεάζει τον κίνδυνο εμφάνισης στεφανιαίας νόσου. Εφόσον το p-value είναι 0.069 τότε δεν μπορούμε με ασφάλεια να απορρίψουμε την μηδενική υπόθεση $H_0 : \beta_{chol} = 0$.

7.5.3 Διαστήματα εμπιστοσύνης

Οι τιμές των παραμέτρων $\hat{\beta}$ που υπολογίζονται αποτελούν μια απλή σημειακή εκτίμηση για την πραγματική τιμή. Είναι σκόπιμο πολλές φορές να παρουσιάζονται διαστήματα εμπιστοσύνης γύρω από αυτή την εκτίμηση. Για μια συγκεκριμένη παράμετρο, έστω $\hat{\beta}_j$ ένα $100(1 - \alpha)\%$ διάστημα εμπιστοσύνης υπολογίζεται από τη σχέση:

$$\hat{\beta}_j \pm z_{1-\alpha/2} \sqrt{\text{var}(\hat{\beta}_j)} \quad (7.24)$$

όπου η διακύμανση του $\hat{\beta}_j$ προκύπτει από τον αντίστροφο πίνακα πληροφορίας $I^{-1} = (X'WX)^{-1}$ ως το jj στοιχείο. Στο μοντέλο m_3 η εκτίμηση του λογαριθμικού λόγου σχετικών πιθανοτήτων ήταν 0.007. Ένα 95% διάστημα εμπιστοσύνης για τον λόγο σχετικών πιθανοτήτων προκύπτει αν πάρουμε το εκθετικό της σχέσης 7.24, (0.999-1.014). Η εμφάνιση των διαστημάτων εμπιστοσύνης γύρω από τις τιμές του λόγου σχετικών πιθανοτήτων προσφέρεται ως επιλογή στο SPSS. Στο παράθυρο διαλόγου της λογιστικής παλινδρόμησης υπάρχει η επιλογή **Options>CI for Exp(B)**.

7.6 Έλεγχοι καλής προσαρμογής

Hoσμερ ανδ Λεμεσηωω

7.7 Αλληλεπιδράσεις και συγχυτικοί παράγοντες

7.8 Στρατηγικές μοντελοποίησης

7.9 Πολύτομη λογιστική παλινδρόμηση

7.10 Ordinal

Κεφάλαιο 8

Ανάλυση Επιβίωσης (Survival Analysis)

8.1 Εισαγωγή

Η *ανάλυση επιβίωσης* (Survival Analysis) είναι ο κλάδος της στατιστικής επιστήμης που ασχολείται μέχρι την εμφάνιση ενός γεγονότος, όπως για παράδειγμα την υποτροπή ή τον θάνατο ενός ασθενούς, τη βλάβη ενός μηχανήματος, τη χρεωκοπία μιας επιχείρησης ή το χρόνο ως τη νίκη σε μια παρτίδα σκάκι.

Στις περισσότερες ιατρικές έρευνες, η ανάλυση επιβίωσης μελετά τον χρόνο ως την υποτροπή ή τον θάνατο ενός ασθενούς. Είναι προφανές πως σε αυτές τις περιπτώσεις, είναι αδύνατο για τον ερευνητή να περιμένει όλο το χρονικό διάστημα που απαιτείται ώστε να υποτροπιάζουν όλοι οι ασθενείς. Σε πολλές περιπτώσεις ο χρόνος αυτός μπορεί να ξεπερνά αρκετά χρόνια, ενώ ενδιάμεσα μπορεί να μεσολαθήσει ένα ανταγωνιστικό γεγονός (competing risk) που θα αποκλείσει την περίπτωση υποτροπής από την ασθένεια. Για παράδειγμα, αν μια ασθενής πεθάνει σε αυτοκινητιστικό δυστύχημα μετά από 4 χρόνια παρακολούθησης, αυτό δεν αποκλείει τη πιθανότητα να είχε υποτροπιάσει στον πέμπτο χρόνο παρακολούθησης.

Το χαρακτηριστικό αυτό συνιστά και τη βασικότερη διαφορά της ανάλυσης επιβίωσης από τις υπόλοιπες στατιστικές μεθόδους, το γεγονός δηλαδή πως το αποτέλεσμα της έρευνας είναι μόνο κατά ένα μέρος διαθέσιμο. Οι παρατηρήσεις για τις οποίες δεν είναι γνωστό το τελικό αποτέλεσμα ονομάζονται λογοκριμένες παρατηρήσεις (censored cases). Οι αιτίες που τελικά δεν παρατηρείται το τελικό γεγονός μπορεί να είναι διαφορετικές. Η βασικότερη είναι αυτή που αναφέρθηκε, μια μελέτη που ξεκινά για την αποτίμηση ενός συγκεκριμένου φαρμάκου έχει ένα σαφώς καθορισμένο χρονικό ορίζοντα, έστω πέντε χρόνια. Με το πέρας της μελέτης, οι ασθενείς που δεν έχουν υποτροπιάσει θεωρούνται λογοκριμένοι, εφόσον μπορεί να υποτροπιάσουν στο μέλλον αλλά δεν γνωρίζουμε πότε. Ένα άλλο παράδειγμα είναι περιπτώσεις ασθενών που για

κάποιο τυχαίο λόγο χάθηκαν από την παρακολούθηση, κάτι που είναι σύνηθες στις ιατρικές έρευνες. Παρόλο που λογοκριμένες παρατηρήσεις είναι ελλιπείς μπορούν να συνεισφέρουν πληροφορία για την έρευνα. Αν ένας ασθενής καταφέρει να επιβιώσει για μερικά χρόνια χωρίς να εμφανίσει την ασθένεια, τότε ο χρόνος αυτός της επιβίωσης θα πρέπει να συμπεριληφθεί στη μελέτη αντί να αγνοηθεί. Στην ουσία, ο πραγματικός χρόνος υποτροπής δεν έχει παρατηρηθεί πλήρως, αφού το μόνο που γνωρίζουμε είναι ότι ξεπερνά ένα συγκεκριμένο κατώφλι, για παράδειγμα τα πέντε χρόνια.

Ο στατιστικός ερευνητής θα ήθελε να εκτιμήσει πόσο χρόνο μπορεί να επιβιώσει ένας ασθενής, δεδομένων κάποιων συγκεκριμένων χαρακτηριστικών που μπορεί να αφορούν τον χρόνο που έχει ήδη επιβιώσει, όσο και κλινικά χαρακτηριστικά της ασθένειας, της θεραπείας ή ακόμα και γονιδιακά χαρακτηριστικά. Βασικά ερωτήματα που μπορούν να απαντηθούν με την ανάλυση επιβίωσης είναι: α) ποιο είναι το ποσοστό των ασθενών που μπορεί να επιβιώσουν πέρα από ένα συγκεκριμένο χρονικό σημείο, πχ ποια μπορεί να είναι η 5ετής επιβίωση γυναικών με καρκίνο των ωοθηκών β) ποια είναι η μέση επιβίωση μιας ομάδας ασθενών με καρδιαγγειακά προβλήματα. Πιο ουσιαστικά όμως, με την ανάλυση επιβίωσης μπορεί να μελετηθεί ο τρόπος με τον οποίο ένα συγκεκριμένο χαρακτηριστικό επηρεάζει τον χρόνο θανάτου και την πιθανότητα να συμβεί σε κάθε χρονική στιγμή. Για παράδειγμα, σε μια μελέτη ασθενών με καρκίνο του μαστού, το μέγεθος του όγκου ή ο αριθμός των διηθημένων λεμφαδένων ή η φαρμακευτική αγωγή και παρέμβαση μπορεί να επηρεάσει το χρόνο επιβίωσης. Κάθε πληροφορία που είναι διαθέσιμη για τους ασθενείς, μπορεί και θα πρέπει να συμπεριληφθεί σε ένα στατιστικό μοντέλο, κατάλληλα φτιαγμένο έτσι ώστε να μπορεί να περιγράψει το μηχανισμό της ασθένειας, τον τρόπο με τον οποίο οι παράγοντες αλληλεπιδρούν και να είναι σε θέση να προσφέρει ασφαλείς εκτιμήσεις των κινδύνων και του χρόνου επιβίωσης. Στο κεφάλαιο αυτό θα παρουσιαστεί μια σύντομη και πρακτική περιγραφή των βασικών εννοιών της ανάλυσης επιβίωσης, καθώς και των βασικών μεθόδων που παρουσιάζονται στην ιατρική βιβλιογραφία.

8.2 Βασικές συναρτήσεις

Έστω T ο χρόνος μέχρι την εμφάνιση ενός προκαθορισμένου γεγονότος. Το γεγονός αυτό μπορεί να είναι ο θάνατος ενός ασθενούς, η υποτροπή από κάποια ασθένεια, ή κάτι θετικό όπως ο χρόνος ως την ίαση, ο χρόνος ως την οριστική διακοπή του καπνίσματος κ.ο.κ. Ο χρόνος ορίζεται ως μια μη αρνητική τυχαία μεταβλητή για την οποία δεν γνωρίζουμε (προς το παρόν) αν προέρχεται από κάποια συγκεκριμένη κατανομή πιθανότητας. Οι βασικές έννοιες που αφορούν την μεταβλητή του χρόνου είναι η συνάρτηση επιβίωσης, η συνάρτηση κινδύνου, η πυκνότητα πιθανότητας και ο μέσος υπολειπόμενος χρόνος. Στις ενότητες που ακολουθούν θα ορίσουμε

τις παραπάνω έννοιες και θα δείξουμε πως συνδέονται μεταξύ τους.

8.2.1 Η συνάρτηση επιβίωσης

Η συνάρτηση επιβίωσης (Survival Function) είναι από τις πλέον βασικές έννοιες στην ανάλυση επιβίωσης. Εδώ θα θεωρήσουμε ως γεγονός το θάνατο. Έστω λοιπόν, T ο χρόνος ζωής ενός ασθενούς, τότε η συνάρτηση επιβίωσης, που συμβολίζεται με S (από το Survival) δηλώνει την πιθανότητα του ασθενούς να επιβιώσει πέρα από ένα συγκεκριμένο χρονικό σημείο t . Η συνάρτηση επιβίωσης ορίζεται ως

$$S(t) = P(T > t) \quad (8.1)$$

Αν ο χρόνος είναι μια συνεχής τυχαία μεταβλητή τότε η συνάρτηση επιβίωσης είναι συμπληρωματική της αθροιστικής συνάρτησης κατανομής, δηλαδή $S(t) = 1 - F(t)$ όπου $F(t) = P(T \leq t)$ δείχνει την πιθανότητα να πεθάνει ένας ασθενής πριν τον χρόνο t . Επίσης, η συνάρτηση επιβίωσης ορίζεται ως το ολοκλήρωμα της συνάρτησης πυκνότητας πιθανότητας, $f(t)$:

$$S(t) = P(T > t) = \int_t^{\infty} f(s) ds \quad (8.2)$$

συνεπώς

$$f(t) = -\frac{dS(t)}{dt}$$

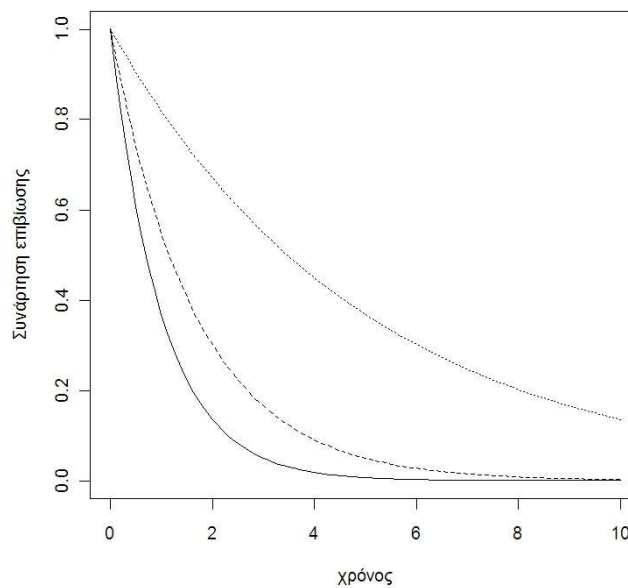
Η $f(t)$ δίνει την συνάρτηση πυκνότητας πιθανότητας για να συμβεί το προκαθορισμένο γεγονός ακριβώς την χρονική στιγμή t , ενώ το ολοκλήρωμα της $f(t)$ από 0 ως το άπειρο δίνει όλη την επιφάνεια που ορίζει η συνάρτηση και ισούτε με ένα.

Η συνάρτηση επιβίωσης εκτιμάται με παραμετρικούς ή παραμετρικούς τρόπους (όπως θα παρουσιαστεί παρακάτω) για ένα χρονικό διάστημα. Η εκτίμηση αυτή χρησιμοποιείται για την κατασκευή ενός διαγράμματος επιβίωσης. Στον οριζόντιο άξονα τοποθετείται ο χρόνος ενώ στον κάθετο οι τιμές της συνάρτησης επιβίωσης. Είναι προφανές ότι στον χρόνο 0 η συνάρτηση επιβίωσης παίρνει την τιμή 1, ενώ με το πέρασμα του χρόνου η συνάρτηση φθίνει μονότονα, ως που να πάρει την τιμή 0 σε κάποιο σημείο στο χρόνο. Το διάγραμμα της καμπύλης επιβίωσης είναι το βασικότερο διάγραμμα που χρησιμοποιείται στην ανάλυση επιβίωσης για την περιγραφή των δεδομένων.

ΠΑΡΑΔΕΙΓΜΑ Δίνεται ότι ο χρόνος ως τον θάνατο μιας ομάδας ασθενών ακολουθεί την εκθετική κατανομή. Η συνάρτηση πυκνότητας πιθανότητας της εκθετικής κατανομής είναι $f(x) = \lambda \exp(-\lambda x)$, $\lambda > 0$, $x > 0$. Για να υπολογιστεί η πιθανότητα ένας από τους ασθενείς να επιβιώσει πέραν του χρόνου t , δεδομένου πως είναι γνωστή η παράμετρος λ χρησιμοποιούμε τη σχέση (8.2).

$$\begin{aligned}
 S(t) &= P(T > t) = \int_t^{\infty} f(s) ds \\
 &= \int_t^{\infty} \lambda \exp(-\lambda s) ds \\
 &= \lambda \left[\frac{\exp(-\lambda s)}{-\lambda} \right]_t^{\infty} \\
 &= \exp(-\lambda \infty) + \exp(-\lambda t) = \exp(-\lambda t)
 \end{aligned}$$

Η παράμετρος λ δείχνει το ρυθμό με τον οποίο η ασθένεια "γεννάει" τα γεγονότα. Για διαφορετικές τιμές της παραμέτρου υπολογίζονται διαφορετικές καμπύλες επιβίωσης. Στο διάγραμμα 8.2.1 παρουσιάζονται οι καμπύλες επιβίωσης για χρόνο από 0.1 ως 10 έτη και διαφορετικές τιμές του λ , 0.2, 0.6, και 1. Παρατηρούμε πως όσο πιο μεγάλη είναι η τιμή της παραμέτρου, τόσο πιο γρήγορα συμβαίνουν τα γεγονότα, με αποτέλεσμα η καμπύλη επιβίωσης που αντιστοιχεί στην τιμή $\lambda = 1$ να βρίσκεται πολύ πιο χαμηλά από τις υπόλοιπες.



Σχήμα 8.1: Καμπύλες επιβίωσης από εκθετική κατανομή. — $\lambda=1$, ... $\lambda=0.6$, - - - $\lambda=0.2$

8.2.2 Η συνάρτηση κινδύνου

Μια άλλη βασική ποσότητα στην ανάλυση επιβίωσης είναι η συνάρτηση κινδύνου (hazard function). Η συνάρτηση κινδύνου αποτελεί ένα βασικό εργαλείο για να υπολογιστούν άλλες ποσότητες, όπως η συνάρτηση επιβίωσης ή η αθροιστική συνάρτηση κινδύνου (cumulative hazard).

αν και στην πράξη λίγες φορές παρουσιάζονται συναρτήσεις κινδύνου αυτούσιες. Η συνάρτηση αυτή είναι επίσης γνωστή και ως ένταση θνησιμότητας (force of mortality) ή ρυθμός αποτυχίας ανά ηλικία (age-specific failure rate) στην δημογραφία. Ο ρυθμός κινδύνου δίνεται από

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (8.3)$$

Αν το T είναι μια συνεχής τυχαία μεταβλητή τότε ισχύει:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d \ln(S(t))}{dt} \quad (8.4)$$

Ως αθροιστική συνάρτηση κινδύνου ορίζεται η ποσότητα:

$$H(t) = \int_0^t h(s) ds = -\ln(S(t)) \quad (8.5)$$

Η σχέση που συνδέει αυτές τις ποσότητες είναι:

$$S(t) = \exp[-H(t)] = \exp\left(-\int_0^t h(s) ds\right) \quad (8.6)$$

Από την 8.3 προκύπτει πως η συνάρτηση κινδύνου μπορεί να νοηθεί ως η πιθανότητα ένα άτομο ηλικίας t να βιώσει το προκαθορισμένο γεγονός στην αμέσως επόμενη χρονική στιγμή. Η συνάρτηση κινδύνου είναι ιδιαίτερα χρήσιμη για την κατανόηση του μηχανισμού που μπορεί να οδηγεί στο προκαθορισμένο γεγονός, κάτω από μια ασθένεια, αφού περιγράφει τον τρόπο ή τον ρυθμό με τον οποίο τα άτομα βιώνουν το γεγονός καθώς περνάει ο χρόνος. Ως συνάρτηση το $h(t)$ μπορεί να πάρει οποιαδήποτε μορφή, με τον μόνο περιορισμό ότι οι τιμές είναι μη αρνητικές για όλες τις τιμές του t .

Σε βιολογικά φαινόμενα, όταν για παράδειγμα μελετάται ο χρόνος ως τον θάνατο, η συνάρτηση κινδύνου είναι συνήθως αύξουσα, καθώς η ηλικία αυξάνει την πιθανότητα θανάτου. Από την άλλη, αν το γεγονός που ενδιαφέρει τον ερευνητή είναι η υποτροπή από κάποια μολυσματική ασθένεια, είναι πιθανό καθώς περνά ο χρόνος ο οργανισμός του ατόμου να αναπτύσσει άμυνες και συνεπώς η συνάρτηση κινδύνου να δίνει μικρότερες τιμές προς το τέλος του χρόνου παρακολούθησης. Ενδιάμεσες πιο πολύπλοκες μορφές για την συνάρτηση κινδύνου είναι αποδεκτές, αλλά αυτές σπάνια συναντώνται στην πράξη σε ιατρικές εφαρμογές.

ΠΑΡΑΔΕΙΓΜΑ Για την ομάδα ασθενών του παραδείγματος 2.1.1 γνωρίζουμε ότι ο χρόνος ακολουθεί εκθετική κατανομή με κάποια παράμετρο λ . Επίσης είδαμε ότι η συνάρτηση επιβίωσης είναι $S(t) = \exp(-\lambda t)$. Από τη σχέση (4) προκύπτει ότι η συνάρτηση κινδύνου είναι η

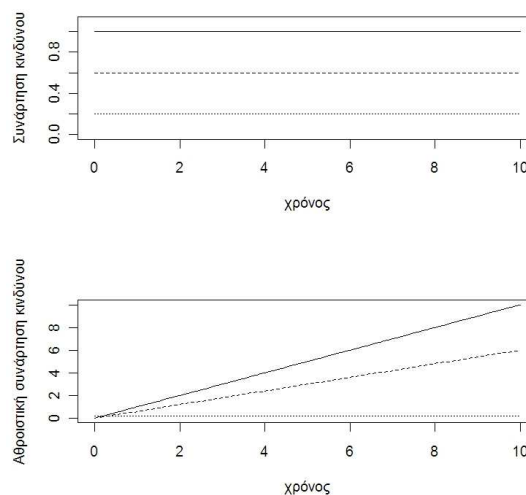
$$h(t) = \frac{f(t)}{S(t)} = \frac{\lambda \exp(-\lambda t)}{\exp(-\lambda t)} = \lambda$$

Συνεπώς, όταν ο χρόνος ακολουθεί εκθετική κατανομή η συνάρτηση κινδύνου ή η ένταση θνησιμότητας είναι σταθερή και δίνεται από την παράμετρο λ , ανεξάρτητα από τον χρόνο που έχει ζήσει ένας ασθενής. Αυτό κάνει την επιλογή της εκθετικής κατανομής ως συνάρτηση χρόνου επιβίωσης πολύ χρήσιμη από μαθηματικής πλευράς, αλλά ταυτόχρονα αποτελεί μια απλοϊκή προσέγγιση. Είναι λίγες οι περιπτώσεις που θα είχε νόημα η υπόθεση ότι ο χρόνος δεν επηρεάζει την ένταση της θνησιμότητας.

Η αθροιστική συνάρτηση επιβίωσης προκύπτει εύκολα από τη σχέση

$$H(t) = \int_0^t h(s) ds = \lambda t$$

Στο Διάγραμμα (8.2.2) παρουσιάζονται η συνάρτηση κινδύνου και η αθροιστική συνάρτηση κινδύνου. Η συνάρτηση κινδύνου δίνεται ως μια ευθεία γραμμή παράλληλη με τον άξονα του χρόνου και παραμένει σταθερή, ανάλογα με τις τιμές που παίρνει η παράμετρος λ .



Σχήμα 8.2: Συνάρτηση κινδύνου και αθροιστική συνάρτηση κινδύνου όταν ο χρόνος επιβίωσης ακολουθεί εκθετική κατανομή για διαφορετικές τιμές της παραμέτρου λ . — $\lambda=1$, ... $\lambda=0.6$, - - - $\lambda=0.2$

8.2.3 Μέσος υπολειπόμενος χρόνος

Ο μέσος υπολειπόμενος χρόνος ζωής μετράει τον χρόνο ζωής που αναμένεται, κατά μέσο όρο, να έχει ένα άτομο ηλικίας t . Ορίζεται ως

$$mlr(t) = E(T - t | T > t)$$

Αποδεικνύεται ότι προκύπτει από την επιφάνεια που ορίζεται από την καμπύλη επιβίωσης από το σημείο t και έπειτα διαιρεμένη με $S(t)$. Ταυτόχρονα, ο μέσος χρόνος ζωής ενός ατόμου μ ορίζεται ως η συνολική επιφάνεια κάτω από την καμπύλη επιβίωσης. Δηλαδή, ο υπολειπόμενος χρόνος ζωής είναι ο αναμενόμενος χρόνος ζωής δεδομένου ότι ο ασθενής έχει ζήσει μέχρι τη χρονική στιγμή t . Ισχύει ότι $\mu = mrl(0)$.

Συνεπώς, για μια συνεχή τυχαία μεταβλητή ισχύει:

$$\begin{aligned} mrl(t) &= \frac{\int_t^\infty (s-t)f(s)ds}{S(t)} = \\ &= \frac{\int_t^\infty sf(s)ds - t \int_t^\infty f(s)ds}{S(t)} = \frac{\int_t^\infty sf(s)ds - tS(t)}{S(t)} = \\ &= \frac{\int_t^\infty S(s)ds}{S(t)} \end{aligned}$$

αφού από τη σχέση (2) ισχύει $\int_t^\infty f(s)ds = S(t)$. Επίσης,

$$\begin{aligned} \int_t^\infty sf(s)ds &= \int_t^\infty s[-S'(s)]ds \\ &= [-sS(s)]_t^\infty - \int_t^\infty s'(-S(s))ds \\ &= -\lim_{s \rightarrow \infty} sS(s) + tS(t) + \int_t^\infty S(s)ds = tS(t) + \int_t^\infty S(s)ds \end{aligned}$$

αφού με λίγη άλγεβρα μπορεί να φανεί ότι

$$-\lim_{s \rightarrow \infty} sS(s) = 0$$

συνεπώς, ισχύει

$$mrl(t) = \frac{\int_t^\infty (s-t)f(s)ds}{S(t)} = \frac{\int_t^\infty S(s)ds}{S(t)} \quad (8.7)$$

Με παρόμοιο τρόπο προκύπτει:

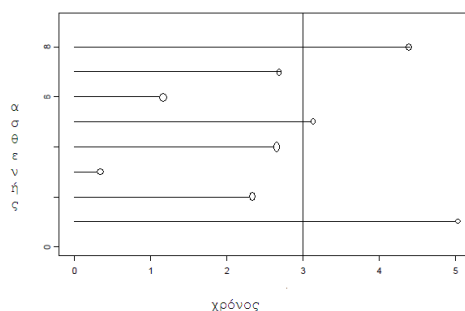
$$\mu = E(T) = \int_0^\infty sf(s)ds = \int_0^\infty S(s)ds \quad (8.8)$$

8.3 Λογοκρισία

Λόγω των ιδιοτήτων που παρουσιάζουν οι ιατρικές μελέτες, τα δεδομένα που συγκεντρώνονται μπορεί να είναι λογοκριμένα (censored). Με τον όρο λογοκρισία εννοούμε τις παρατηρήσεις εκείνες για τις οποίες το τελικό αποτέλεσμα, το προκαθορισμένο γεγονός, δεν έχει παρατηρηθεί κατά τη διάρκεια της μελέτης. Ας υποθέσουμε για άλλη μια φορά ότι το γεγονός που μας ενδιαφέρει είναι ο θάνατος από μια ασθένεια. Ο συχνότερος λόγος για τον οποίο υπάρχουν

λογοκριμένα δεδομένα είναι όταν σε μια ιατρική μελέτη ο χρόνος παρακολούθησης ξεκινά από $t = 0$ και σταματά σε ένα προκαθορισμένο σημείο, έπειτα από κάποιο χρόνο. Κατά τη διάρκεια της μελέτης κάποιοι από τους ασθενείς θα αποβιώσουν ενώ κάποιοι άλλοι όχι. Αν δεν παρατηρηθεί το γεγονός μέσα στο προκαθορισμένο χρονικό διάστημα της μελέτης αυτό δε σημαίνει ότι το άτομο δεν πρόκειται να πεθάνει μέσα στο επόμενο χρονικό διάστημα. Η μορφή αυτή λογοκρισίας ονομάζεται *δεξιά λογοκρισία* και αναπαριστάται γραφικά στο διάγραμμα (8.3).

ΠΑΡΑΔΕΙΓΜΑ Μια μελέτη που αφορά τον χρόνο ζωής ατόμων με καρκίνο του πνεύμονα είχε διάρκεια τρία χρόνια. Στο διάγραμμα 8.3 παρουσιάζεται ο χρόνος ζωής για οκτώ άτομα. Κατά την διάρκεια των τριών χρόνων πέντε από τα οκτώ άτομα πέθαναν από καρκίνο του πνεύμονα. Μετά το πέρας της μελέτης τρία άτομα ακόμα πέθαναν από καρκίνο του πνεύμονα (άτομα 1, 5, 8). Για τους ασθενείς 1, 5, και 8 ο θάνατος θα επέλθει μετά το τέλος της μελέτης. Συνεπώς, το γεγονός δεν είναι παρατηρήσιμο, αλλά ξέρουμε ότι ο χρόνος επιβίωσης ήταν μεγαλύτερος από 3 έτη.



Σχήμα 8.3: Χρόνοι ζωής για οκτώ ασθενείς με καρκίνο του πνεύμονα. Η κάθετη γραμμή δείχνει το σημείο που σταμάτησε η μελέτη.

Στην πράξη, όταν ξεκινάει μια μελέτη δε στρατολογούνται όλα τα άτομα στον χρόνο μηδέν, αλλά εισέρχονται στη μελέτη σε μια διάρκεια χρόνου. Σε αυτές τις περιπτώσεις ο χρόνος δεν ταυτίζεται με τον ημερολογιακό χρόνο, αλλά με την διάρκεια παρακολούθησης. Ένα παράδειγμα είναι μια μελέτη που διοργανώνεται σε ένα νοσοκομείο που αφορά τον χρόνο που απαιτείται ως την εξαγωγή ενός ασθενή από την στιγμή που θα αρχίσει η νοσηλεία. Η μελέτη μπορεί να ξεκινήσει μια συγκεκριμένη χρονική στιγμή, αλλά τα δεδομένα που συγκεντρώνονται έχουν την ακόλουθη μορφή:

Το Διάγραμμα (8.3) παρουσιάζει τη μορφή των δεδομένων όπως δίνονται στον πίνακα. Ε-

Πίνακας 8.1: Χρόνος εισαγωγής και εξαγωγής τεσσάρων υποθετικών ασθενών σε ένα νοσοκομείο.

ασθενής	εισαγωγή	εξαγωγή	χρόνος
1	15-4-2008	26-4-2008	11
2	20-4-2008	24-4-2008	4
3	25-4-2008	02-5-2008	7
4	26-4-2008	30-4-2008	4

φόσον η μελέτη σταμάτησε μετά από το πέρας 12 ημερών, το τρίτο άτομο στη μελέτη, παρά το γεγονός ότι παρέμεινε στο νοσοκομείο για επτά ημέρες, παρουσιάζει λογοκρισία. Όταν τα δεδομένα είναι αυτής της μορφής, τότε ο ημερολογιακός χρόνος αντικαθιστάτε από τον χρόνο παρακολούθησης. Μια γραφική απεικόνιση των δεδομένων αυτών παρουσιάζεται στο Διάγραμμα 8.3. Σε κάθε ένα από τους ασθενείς το ενδιαφέρον μετατοπίζεται από τον ημερολογιακό χρόνο, στον χρόνο που παρέμειναν στο νοσοκομείο (σε ημέρες). Όλες οι παρατηρήσεις ξεκινούν από το μηδέν και για αυτούς που βίωσαν το προκαθορισμένο γεγονός (στη συγκεκριμένη περίπτωση την έξοδο από το νοσοκομείο) μετράμε τον χρόνο T_i , $i = 1, 2, 3$ ως που να συμβεί. Αντίστοιχα, για το άτομο που παρέμεινε στο νοσοκομείο στο τέλος της μελέτης μετράμε τον χρόνο C_i , το χρόνο δηλαδή ως τη λογοκρισία.

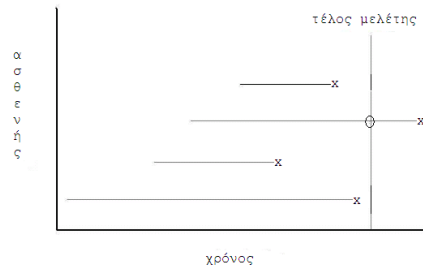
Σε κάθε μελέτη ανάλυσης επιβίωσης με δεξιά λογοκρισία, για κάθε άτομο που βρίσκεται υπό μελέτη θεωρείται ότι υπάρχει ένας συγκεκριμένος χρόνος ζωής T . Αν αυτός ο χρόνος δεν παρατηρηθεί επειδή η μελέτη τελείωσε νωρίτερα, τότε το άτομο είναι λογοκριμένο στον χρόνο που συμβολίζεται ως C . Τα δεδομένα από ιατρικές έρευνες παρουσιάζονται ανά ζεύγη (T_i, d_i) όπου i είναι ο δείκτης για το κάθε άτομο ($i = 1, 2, \dots, n$, το πλήθος των ατόμων υπό παρακολούθηση) και D μια δείκτρια συνάρτηση που παίρνει την τιμή 1 αν το υπό μελέτη γεγονός έχει συμβεί ή 0 διαφορετικά, δηλαδή:

$$d_i = \begin{cases} 1 & \text{αν } T_i < C_i \\ 0 & \text{αν } T_i \geq C_i \end{cases}$$

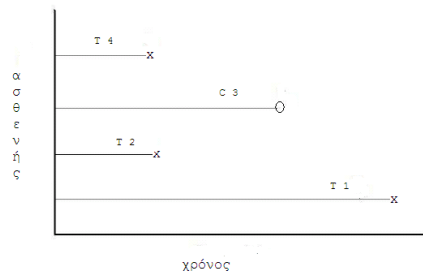
Υπάρχουν κι άλλοι λόγοι για τους οποίους μπορεί να προκύψουν λογοκριμένες παρατηρήσεις. Ένας από αυτούς είναι το γεγονός ότι κάποιοι από τους ασθενείς χάνονται από τη μελέτη. Αν υπάρχει κάποιος συγκεκριμένος μηχανισμός σύμφωνα με τον οποίο οι ασθενείς μπορεί να φεύγουν από μια μελέτη (για παράδειγμα, νιώθουν υγιείς, συνεπώς σταματάνε την φαρμακευτική αγωγή) τότε τα αποτελέσματα μιας ανάλυσης θα είναι μεροληπτικά. Για να είναι

έγκυρη μια ανάλυση επιβίωσης θα πρέπει η λογοκρισία να μην πληροφοριακή, δηλαδή ο τρόπος που προκύπτουν οι λογοκριμένες παρατηρήσεις να μην μη συνδέεται με την εμφάνιση του γεγονότος.

Άλλες μορφές λογοκρισίας περιλαμβάνουν την λογοκρισία από *ανταγωνιστικούς κινδύνους*, (*competing risks*), *αριστερή* λογοκρισία και λογοκρισία *διαστήματος* (*interval censoring*). Για μια αναλυτική περιγραφή δείτε Andersen et al. (σελ. 128-165).



Σχήμα 8.4: Χρόνοι ζωής για τέσσερις ασθενείς που εισάγονται σε ένα νοσοκομείο. Η κάθετη γραμμή δείχνει το σημείο που σταμάτησε η μελέτη.



Σχήμα 8.5: Χρόνοι ζωής για τέσσερις ασθενείς που εισάγονται σε ένα νοσοκομείο.

8.4 Μη παραμετρική εκτίμηση συναρτήσεων

8.4.1 Εκτιμητής για τη συνάρτηση επιβίωσης

Ο εκτιμητής κατά τους Kaplan και Meier χρησιμοποιείται για εκτιμηθεί η συνάρτηση επιβίωσης σε μια ομάδα παρατηρήσεων. Έστω n παρατηρήσεις για τις οποίες είναι γνωστά ο χρόνος T και μια δίτιμη μεταβλητή D που παίρνει τις τιμές $d_i = 1$ αν το άτομο i παρουσιάσει το προκαθορισμένο γεγονός και $d_i = 0$ αν το άτομο i είναι λογοκριμένο. Θεωρούμε πάντα πως από τα άτομα που συμμετέχουν σε μια μελέτη, κάποια από αυτά θα παρουσιάσουν το γεγονός που ενδιαφέρει για τη συγκεκριμένη μελέτη και κάποια άλλα όχι. Μια βασική υπόθεση για να είναι ο εκτιμητής αμερόληπτος είναι πως η κατανομή της μεροληψίας που δίνεται από τη μεταβλητή D είναι μη πληροφοριακή. Αυτό σημαίνει πως τα άτομα που λογοκρίνονται σε μια μελέτη παρουσιάζονται τυχαία και όχι με συστηματικό τρόπο.

Για να εκτιμηθεί η συνάρτηση επιβίωσης, τα δεδομένα διατάσσονται κατά αύξουσα σειρά του χρόνου, $t_{(1)}, t_{(2)}, \dots, t_{(D)}$ (προς το παρόν θεωρείται ότι κάθε χρόνος που έχει καταγραφεί είναι μοναδικός και δεν υπάρχουν ταυτίσεις μεταξύ δύο παρατηρήσεων). Πρέπει να σημειωθεί ότι ο χρόνος εκτείνεται ως τη στιγμή που συμβαίνει το τελευταίο γεγονός, δηλαδή ως το σημείο $t_{(D)}$. Αν από εκεί και πέρα υπάρχουν άτομα τα οποία λογοκρίνονται σε ύστερο χρόνο, αυτά δε συνεισφέρουν πληροφορία για την εκτίμηση της συνάρτησης επιβίωσης. Σε κάθε έναν από τους χρόνους αυτούς $t_i, i = 1, 2, \dots, D$ αντιστοιχεί μια τιμή $d_i, i = 1, 2, \dots, D$ που δηλώνει αν η παρατήρηση είναι λογοκριμένη ή όχι. Έστω N_i το σύνολο των παρατηρήσεων που βρίσκονται σε κίνδυνο στον χρόνο t_i . Ο εκτιμητής Kaplan-Meier χωρίζει το συνολικό χρόνο σε υποδιαστήματα στα οποία συμβαίνει ένα γεγονός. Στον χρόνο t η συνάρτηση επιβίωσης υπολογίζεται ως:

$$\hat{S}(t) = \begin{cases} 1, & \text{αν } t < t_1 \\ \prod_{t_i \leq t} \left[\frac{N_i - d_i}{N_i} \right] = \prod_{t_i \leq t} \left[1 - \frac{d_i}{N_i} \right], & \text{αν } t_1 \leq t \end{cases} \quad (8.9)$$

όπου t_1 είναι ο μικρότερος χρόνος αποτυχίας που παρατηρείται στα δεδομένα.

Ο εκτιμητής ονομάζεται και product limit estimator γιατί προκύπτει ως το γινόμενο πιθανοτήτων επιβίωσης. Πρακτικά αυτό σημαίνει πως για να επιβιώσει ένα άτομο 4 μέρες μετά από μια επέμβαση θα πρέπει πρώτα να έχει επιβιώσει τουλάχιστον 3. Συνεπώς, η πιθανότητα επιβίωσης $S(4)$ προκύπτει ως το γινόμενο των πιθανοτήτων να επιβιώσει το άτομο την πρώτη, δεύτερη και τρίτη μέρα. Πρόκειται δηλαδή για μια δεσμευμένη πιθανότητα. Στην αρχή της μελέτης, όλα τα Y άτομα είναι σε κίνδυνο και κανένας από αυτούς δεν έχει υποστεί το προκαθορισμένο γεγονός. Συνεπώς, ο εκτιμητής είναι 1, ως την στιγμή που θα συμβεί το πρώτο γεγονός, γιατί ακριβώς πριν τη στιγμή που θα συμβεί το πρώτο γεγονός, στο χρόνο δηλαδή $t_{(1)}$ έχουμε $1 - \frac{0}{N_1} = 1$, όπου N_1 είναι ο αριθμός των ατόμων σε κίνδυνο ακριβώς πριν συμβεί το πρώτο γεγονός. Γενικότερα,

$1 - \frac{d_i}{N_i}$ είναι η πιθανότητα επιβίωσης στο διάστημα $(t_{(i-1)}, t_{(i)})$.

Οι τιμές της συνάρτησης παραμένουν σταθερές μεταξύ δύο διαδοχικών χρόνων και παρουσιάζουν άλματα ακριβώς τη στιγμή που θα συμβεί το επόμενο γεγονός. Αν δηλαδή σε μια μελέτη ο πρώτος θάνατος συμβαίνει 4 ημέρες μετά από μια επέμβαση η πιθανότητα επιβίωσης στο διάστημα 0 ως 4 μέρες μετά την επέμβαση παραμένει σταθερή στο 1. Από εκεί και έπειτα η συνάρτηση επιβίωσης επανυπολογίζεται όποτε παρατηρείται ένας θάνατος. Σε κάθε χρονική στιγμή που εκτιμάται η πιθανότητα επιβίωσης, η τιμή του εκτιμητή εξαρτάται όχι μόνο από τον αριθμό των γεγονότων που συμβαίνουν αλλά και από τον αριθμό των ατόμων που έχουν λογοκριθεί και συνεπώς δεν βρίσκονται πλέον σε κίνδυνο.

ΠΑΡΑΔΕΙΓΜΑ Έστω 10 ασθενείς με καρκίνο του πνεύμονα για τους οποίους είναι γνωστός ο χρόνος ως την υποτροπή από την ασθένεια. Οι δέκα διαφορετικοί χρόνοι (σε μήνες) για τους ασθενείς αυτούς είναι 6, 8, 12+, 14, 16, 16, 16+, 19, 21+, 24. Οι παρατηρήσεις με το σύμβολο (+) αντιπροσωπεύουν λογοκριμένους ασθενείς. Για να υπολογιστεί η καμπύλη επιβίωσης δημιουργείται ο παρακάτω πίνακας:

Πίνακας 8.2: Χρόνοι ζωής 10 υποθετικών ασθενών με καρκίνο του πνεύμονα

Χρόνος (t)	# ασθενών (N)	# γεγονότων (d)	d/N	S(t)
0	10	0	0	1
6	10	1	1/10	0.900
8	9	1	1/9	0.800
12	8	0	0/8	0.800
14	7	1	1/7	0.686
16	6	2	2/4	0.457
19	3	1	1/3	0.305
21	2	0	0/2	0.305
24	1	1	1/1	0.000

Στον χρόνο $t = 14$ επτά άτομα βρίσκονται σε κίνδυνο και ένα παθαίνει υποτροπή. Συνεπώς, η πιθανότητα επιβίωσης είναι

$$S(14) = \prod_{t \leq 12} [1 - \frac{d_i}{N_i}] = (1 - 0/10)(1 - 1/10)(1 - 1/9)(1 - 0/8)(1 - 1/7) = 0.686.$$

Μαζί με την εκτίμηση της συνάρτησης επιβίωσης $\hat{S}(t)$ θα πρέπει να υπολογιστεί και η διακύμανση της εκτίμησης, δηλαδή απαιτείται μια σχέση που να εκτιμά τη διακύμανση του εκτιμητή

επιβίωσης, $\hat{V}(\hat{S}(t))$. Για το σκοπό αυτό χρησιμοποιείται η μέθοδος *δέλτα*, μια διαδικασία που εφαρμόζεται συχνά όταν ο εκτιμητής της διακύμανσης δεν αφορά απλά ένα άθροισμα παρατηρήσεων και όπως είδαμε προηγουμένως, ο εκτιμητής Kaplan-Meier στο χρόνο t προκύπτει ως ένα γινόμενο αναλογιών.

Η μέθοδος δέλτα μπορεί να χρησιμοποιηθεί σε οποιαδήποτε συνάρτηση η οποία προσεγγίζεται από ένα ανάπτυγμα Taylor. Έστω λοιπόν μια συνάρτηση $f(X)$ πάνω σε μια τυχαία μεταβλητή Q . Η μέθοδος χρησιμοποιεί τους δύο πρώτους όρους ενός αναπτύγματος Taylor, γύρω από το μέσο, για να προσεγγίσει την τιμή της συνάρτησης

$$f(X) = f(\mu) + (X - \mu)f'(\mu) \quad (8.10)$$

όπου

$$f'(\mu) = \left. \frac{df(X)}{dX} \right|_{X=\mu}$$

είναι η παράγωγος της συνάρτησης ως προς X , στο σημείο όπου αυτό παίρνει τη μέση τιμή μ . Από αυτή τη σχέση προκύπτει ότι η διακύμανση της συνάρτησης είναι κατά προσέγγιση:

$$V(f(X)) = V(X - \mu)[f'(\mu)]^2 = \sigma^2 f'(\mu)^2 \quad (8.11)$$

όπου σ^2 είναι η διακύμανση του Q . Αν αντικατασταθούν τα σ^2 και μ με τους αντίστοιχους εκτιμητές $\hat{\sigma}^2$ και $\hat{\mu}$ τότε ο εκτιμητής της διακύμανσης προκύπτει από τη σχέση

$$\hat{V}(f(X)) = \hat{\sigma}^2 f'(\hat{\mu})^2 \quad (8.12)$$

Για να βρούμε τον εκτιμητή της διακύμανσης του Kaplan-Meier εκτιμητή χρησιμοποιούμε την εκθετική συνάρτηση $f(X) = \exp(X)$. Έχουμε δηλαδή ότι $\hat{S}(t) = \exp[\ln(\hat{S}(t))]$. Τότε από το ανάπτυγμα Taylor προκύπτει

$$\exp(X) = \exp(\mu) + (X - \mu) \exp(\mu)$$

και η κατά προσέγγιση διακύμανση είναι

$$\hat{V}(\exp(X)) \cong \sigma \exp(\mu)^2 \quad (8.13)$$

Αλλά το $\exp(\mu)^2$ μπορεί να αντικατασταθεί με το $\exp(\hat{\mu})^2 = \exp(\ln \hat{S}(t))^2 = \hat{S}(t)^2$ και το σ με $\hat{\sigma}$. Μένει λοιπόν να δείξουμε τι είναι το $\hat{\sigma}$.

Εφόσον $\hat{S}(t) = \prod_{t_i \leq t} \left[\frac{N_i - d_i}{N_i} \right]$ ο λογάριθμος είναι

$$\ln \hat{S}(t) = \sum_{t_i \leq t} \ln \frac{N_i - d_i}{N_i} = \sum_{t_i \leq t} \ln(\hat{p}_i)$$

όπου $\hat{p}_i = (N_i - d_i)/N_i$. Βασική υπόθεση σε αυτό το σημείο είναι ότι οι παρατηρήσεις για την επιβίωση των N_i ατόμων σε κίνδυνο είναι ανεξάρτητες Bernoulli παρατηρήσεις με σταθερή πιθανότητα p_i . Κάτω από αυτή την υπόθεση ο εκτιμητής της πιθανότητας p_i είναι ο \hat{p}_i με εκτιμητή της διακύμανσης του το $\hat{p}_i(1 - \hat{p}_i)/N_i$. Έτσι, από την σχέση (8.12) προκύπτει

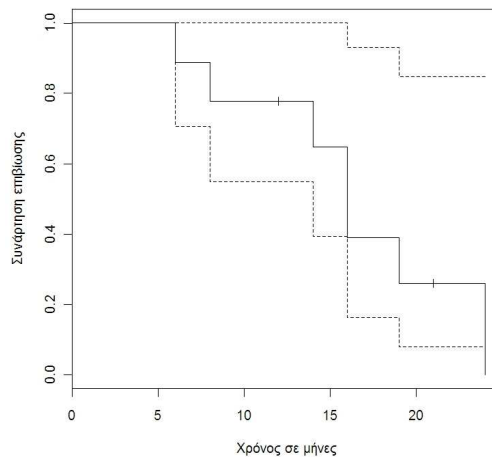
$$\hat{V}(\ln(\hat{p}_i)) = \frac{1}{\hat{p}_i^2} \frac{\hat{p}_i(1 - \hat{p}_i)}{N_i} \cong \frac{d_i}{N_i(N_i - d_i)}$$

Η τελευταία σχέση μπορεί να αντικατασταθεί στην φόρμουλα (8.13) και συνεπώς προκύπτει

$$\hat{V}[\hat{S}(t)] = \hat{S}(t)^2 \sum_{t_i \leq t} \frac{d_i}{N_i(N_i - d_i)} \quad (8.14)$$

Ο τύπος αυτός είναι γνωστός και ως φόρμουλα του Greenwood (1926) .

Στο παράδειγμα παραπάνω $\hat{V}(\hat{S}(14)) = 0.686^2 * [1/10*9 + 1/9*8 + 1/7*6] = 0.02296$. Από τα δεδομένα του πίνακα κατασκευάζεται η καμπύλη επιβίωσης στο διάγραμμα 8.4.1 με 95% διάστημα εμπιστοσύνης.



Σχήμα 8.6: Καμπύλη επιβίωσης για την υποτροπή από καρκίνο του πνεύμονα σε 10 ασθενείς. Οι διακεκομμένες γραμμές δίνουν ένα 95% διάστημα εμπιστοσύνης γύρω από την εκτίμηση.

8.4.2 Ανάλυση δεδομένων με τη χρήση του SPSS

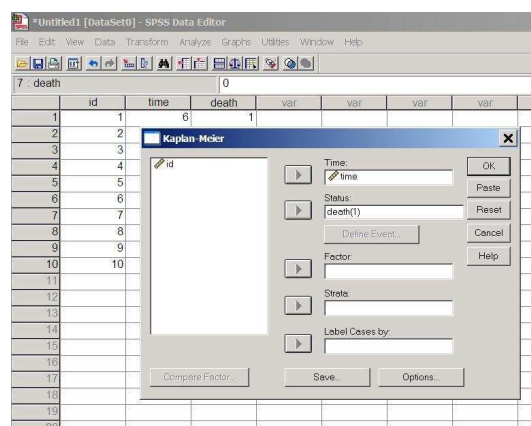
Το στατιστικό πακέτο SPSS περιλαμβάνει μια σειρά από εντολές για την ανάλυση δεδομένων επιβίωσης. Τα δεδομένα του παραδείγματος πρέπει να περαστούν σε δύο στήλες, μία που θα περιλαμβάνει το χρόνο και μια που θα περιλαμβάνει τη δείκτρια συνάρτηση. Κάθε ένα από τα άτομα που περιλαμβάνονται στη μελέτη θα πρέπει να έχει μια γραμμή. Στο διάγραμμα

(2.3.2) παρουσιάζεται η μορφή που θα έχει το φύλλο εργασίας του SPSS για τα δεδομένα αυτά. Σημειώνεται πως υπάρχουν τρεις ξεχωριστές γραμμές για τον χρόνο 16, εφόσον στον χρόνο αυτό υπήρχαν δύο άτομα που παρουσίασαν το γεγονός και ένα που λογοκρίθηκε.

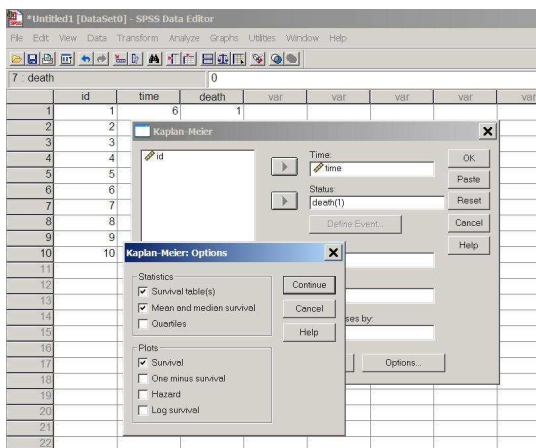
	id	time	death	var	var	var
1	1	6	1			
2	2	8	1			
3	3	12	0			
4	4	14	1			
5	5	16	1			
6	6	16	1			
7	7	16	0			
8	8	19	1			
9	9	21	0			
10	10	24	1			
11						
12						
13						

Σχήμα 8.7: Παράδειγμα φύλλου δεδομένων SPSS

Από το μενού *Analyze > Survival > Kaplan-Meier* προκύπτει το παράθυρο με τις επιλογές για την εκτίμηση της καμπύλης επιβίωσης. Στο αριστερό μέρος βρίσκεται το παράθυρο που περιλαμβάνει τις μεταβλητές που έχουν δημιουργηθεί στο φύλλο δεδομένων. Η μεταβλητή που περιγράφει το χρόνο επιλέγεται και με το βελάκι πρέπει να τοποθετηθεί δεξιά, όπως επίσης και η μεταβλητή που περιγράφει αν έγινε το γεγονός ή όχι. Η μεταβλητή *death* εισάγεται στο παράθυρο *status* και στη συνέχεια πατώντας την επιλογή *Define event* πρέπει να δηλωθεί στο πρόγραμμα ότι το γεγονός συμβαίνει εκεί που η μεταβλητή παίρνει την τιμή 1. Όλες οι άλλες τιμές που μπορεί να έχουν καταχωρηθεί στο φύλλο δεδομένων θεωρούνται από το πρόγραμμα ότι δηλώνουν λογοκρίμενες παρατηρήσεις.



Στη συνέχεια από το μενού *options* επιλέγονται τα στατιστικά και τα διαγράμματα που θέλουμε να παράξει το πρόγραμμα.



Τα αποτελέσματα παρουσιάζονται στο παράθυρο output . Πρώτα ένας πίνακας επιβίωσης που περιλαμβάνει το χρόνο και τη δείκτρια συνάρτηση μαζί με την εκτίμηση της πιθανότητας επιβίωσης στα χρονικά σημεία όπου συμβαίνει ένας γεγονός. Δίπλα παρουσιάζεται το τυπικό σφάλμα της εκτίμησης αυτής, ενώ υπάρχουν και δύο στήλες ακόμα με τον αριθμό των γεγονότων που παρατηρούνται και τον αριθμό των ατόμων σε κίνδυνο σε κάθε χρονικό σημείο. Ο δεύτερος

Survival Table

	Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases
			Estimate	Std. Error		
1	6.000	1	.900	.095	1	9
2	8.000	1	.800	.126	2	8
3	12.000	0	.	.	2	7
4	14.000	1	.688	.151	3	6
5	16.000	1	.	.	4	5
6	16.000	1	.457	.166	5	4
7	16.000	0	.	.	5	3
8	19.000	1	.305	.167	6	2
9	21.000	0	.	.	6	1
10	24.000	1	.000	.000	7	0

πίνακας παρουσιάζει το μέσο και διάμεσο εκτιμώμενο χρόνο επιβίωσης μαζί με το τυπικό σφάλμα και ένα 95% διάστημα εμπιστοσύνης. Τέλος, το SPSS δίνει ένα διάγραμμα με την καμπύλη

Means and Medians for Survival Time

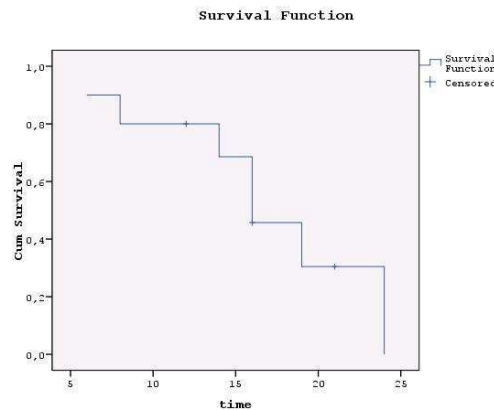
		Mean ^a		Median			
Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
16.867	2.153	12.647	21.086	16.000	2.181	11.725	20.275

a. Estimation is limited to the largest survival time if it is censored.

επιβίωσης. Στα σημεία όπου συμβαίνει μια λογοκρισία αυτή σημειώνεται με ένα σταυρό στην καμπύλη επιβίωσης.

8.4.2.1 Η συνάρτηση κινδύνου

Η συνάρτηση επιβίωσης που παρουσιάστηκε στην προηγούμενη ενότητα, μπορεί να χρησιμοποιηθεί για την εκτίμηση του αθροιστικού κινδύνου, από τον τύπο $H(t) = -\ln[S(t)]$. Έχει όμως αποδειχθεί από τους Nelson (1972) και Aalen (1978) ότι ο εκτιμητής για τη συνάρτηση



κινδύνου $h(t)$ έχει καλύτερες ιδιότητες μικρού δείγματος. Ο εκτιμητής που έχει καθιερωθεί να λέγεται εκτιμητής Nelson-Aalen για την αθροιστική συνάρτηση κινδύνου δίνεται από :

$$\hat{H}(t) = \begin{cases} 0, & \text{αν } t < t_1 \\ \sum_{t_i \leq t} \frac{d_i}{N_i}, & \text{αν } t_1 \leq t \end{cases} \quad (8.15)$$

ενώ η διακύμανση δίνεται από :

$$s_H^2(t) = \sum_{t_i \leq t} \frac{d_i}{N_i^2} \quad (8.16)$$

Από τον εκτιμητή Nelson-Aalen προκύπτει εύκολα ότι $\hat{S}(t) = \exp[-\hat{H}(t)]$.

Η θεμελίωση του εκτιμητή της συνάρτησης κινδύνου από τον Aalen στηρίχθηκε πάνω σε counting processes και martingales . Ο Aalen ήταν ο πρώτος που προσέφερε μια αυστηρή μαθηματική θεμελίωση σε θέματα ανάλυσης επιβίωσης και όλες οι εξελίξεις πάνω στον τομέα στηρίχθηκαν στην προσέγγιση του για την απόδειξη του εκτιμητή της συνάρτησης κινδύνου. Η ανάλυση επιβίωσης με τη χρήση martingales παρουσιάζεται με λεπτομέρεια στα βιβλία των Andersen, Borgan, Gill, Keiding (1993) καθώς και στο βιβλίο των Fleming, Harrington (1991) . Ο εκτιμητής των Nelson-Aalen χρησιμοποιείται για δύο πρακτικούς λόγους επίσης. Ο πρώτος είναι για την αξιολόγηση διαφορετικών παραμετρικών μοντέλων. Για παράδειγμα, ένα διάγραμμα του $\hat{H}(t)$ έναντι του χρόνου t θα πρέπει να είναι κατά προσέγγιση γραμμικό αν η εκθετική κατανομή με παράμετρο λ εφαρμόζει καλά στα δεδομένα. Ο δεύτερος λόγος χρήσης του εκτιμητή είναι στον υπολογισμό της $h(t)$. Αν από τα δεδομένα υπολογιστεί η αθροιστική συνάρτηση κινδύνου $\hat{H}(t)$ τότε η κλίση του εκτιμητή αυτού μπορεί να δώσει μια εκτίμηση για την $h(t)$. Καλύτερες εκτιμήσεις προκύπτουν αν εξομαλυνθεί η $\hat{H}(t)$ με την χρήση παραμετρικών kernels.

Ο Petterson (1977) πρότεινε έναν άλλο εκτιμητή για την αθροιστική συνάρτηση κινδύνου

που προκύπτει από τον εκτιμητή κατά Kaplan-Meier ως εξής:

$$\hat{H}(t) = -\ln \hat{S}(t) = -\ln \prod_{t_i \leq t} \frac{1 - d_i}{N_i} = -\ln \sum_{t_i \leq t} \frac{1 - d_i}{N_i} = \sum_{t_i \leq t} \ln \frac{1 - d_i}{N_i} \quad (8.17)$$

Με τη χρήση του αναπτύγματος Taylor μπορούμε να δείξουμε ότι

$$d_i/N_i \leq -\ln(1 - d_i)/N_i$$

για κάθε χρόνο t . Αυτό σημαίνει πως ο εκτιμητής των Nelson-Aalen θα είναι πάντα μεγαλύτερος ή ίσος από αυτόν που προκύπτει από τον εκτιμητή Kaplan-Meier. Αν το μέγεθος όμως των ατόμων σε κίνδυνο είναι μεγάλο σε σχέση με τον αριθμό των γεγονότων τότε $d_i/N_i \cong -\ln((1 - d_i)/N_i)$ και στην πράξη δεν θα παρατηρείται κάποια διαφορά ανάμεσα στους εκτιμητές.

Το SPSS δίνει το διάγραμμα της αθροιστικής συνάρτησης κινδύνου όπως υπολογίζεται από τον τύπο του Petterson αν επιλεγεί από το μενού options το διάγραμμα Hazard. Τα βήματα είναι παρόμοια με αυτά που είδαμε για την εκτίμηση της συνάρτησης επιβίωσης.

8.5 Έλεγχοι υποθέσεων

Η εκτίμηση της συνάρτησης επιβίωσης αποτελεί το πρώτο μέρος μιας ανάλυσης δεδομένων επιβίωσης. Το επόμενο βήμα συνήθως είναι η σύγκριση δύο ομάδων ασθενών που διαφέρουν κατά ένα χαρακτηριστικό. Παράδειγμα αποτελεί η σύγκριση δύο ομάδων ασθενών από μια κλινική δοκιμή στην οποία, η μια ομάδα έχει λάβει ένα νέο πειραματικό σκεύασμα και η δεύτερη τη καθιερωμένη θεραπεία. Χρειάζεται συνεπώς ένας στατιστικός έλεγχος που θα ελέγχει αν οι δύο ομάδες διαφέρουν ως προς τις πιθανότητες επιβίωσης. Στην ανάλυση επιβίωσης, λόγω της ειδικής φύσης των δεδομένων χρησιμοποιούνται μη παραμετρικοί έλεγχοι. Ο πιο συνηθισμένος έλεγχος είναι ο έλεγχος log-rank.

Ο έλεγχος log-rank συγκρίνει τις εκτιμήσεις δύο συναρτήσεων κινδύνου σε κάθε χρόνο που παρατηρείται ένα γεγονός. Αν υποθέσουμε ότι συγκρίνουμε την συνάρτηση κινδύνου μιας ομάδας ασθενών που συμμετέχουν σε μια κλινική μελέτη και λαμβάνουν το φάρμακο (ομάδα Α) με την αντίστοιχη ομάδα που παίρνει ψευδοφάρμακο (ομάδα Β), οι υποθέσεις του ελέγχου διατυπώνονται ως εξής:

$$H_0 : h_A(t) = h_B(t), \text{ για κάθε } t$$

$$H_A : h_A(t) \neq h_B(t), \text{ για κάποιο } t$$

Σε κάθε χρονική στιγμή i ο αριθμός των παρατηρούμενων γεγονότων στην ομάδα k , D_{ik} συγκρίνεται με τον αριθμό των αναμενόμενων γεγονότων για το ίδιο χρονικό διάστημα στην ίδια ομάδα E_{ik} , κάτω από την αρχική υπόθεση.

Έστω $t_{(1)} < t_{(2)} < \dots < t_{(D)}$ οι διατεταγμένοι χρόνοι των γεγονότων σε όλο το δείγμα. Στον χρόνο $t_{(i)}$ παρατηρούνται d_{ik} γεγονότα στην k ομάδα, σε ένα πλήθος N_{ik} ασθενών σε κίνδυνο. Και έστω $d_i = \sum_{j=1}^k d_{ij}$ το πλήθος των θανάτων στον χρόνο $t_{(i)}$ αθροιστικά για όλες τις ομάδες, και $N_i = \sum_{j=1}^k N_{ik}$ το πλήθος των ατόμων σε κίνδυνο σε όλο το δείγμα. Αν η αρχική υπόθεση είναι σωστή, τότε ο αναμενόμενος αριθμός γεγονότων σε κάθε υποομάδα δεν θα διαφέρει, αλλά θα είναι ο εκτίμητης του κινδύνου σε όλο το δείγμα $\frac{d_i}{N_i}$ πολλαπλασιασμένος με τον αριθμό των ατόμων σε κίνδυνο από την κάθε υπο-ομάδα: $E_{ik} = N_{ik} \frac{d_i}{N_i}$. Ο έλεγχος log-rank για τη σύγκριση δύο ομάδων χρησιμοποιεί την ελεγχουσυνάρτηση:

$$Z = \frac{\sum_{i=1}^D d_{i1} - N_{i1}(d_i/N_i)}{\sqrt{\sum_{i=1}^D V_i}} \quad (8.18)$$

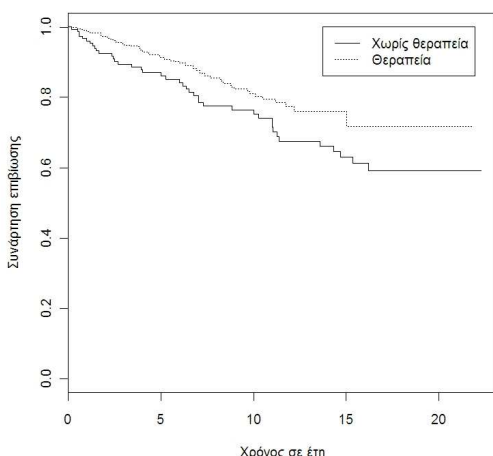
όπου

$$V_i = d_i \left(\frac{N_{i1}}{N_i} \right) \left(1 - \frac{N_{i1}}{N_i} \right) \left(\frac{N_i - d_i}{N_i - 1} \right)$$

είναι η διακύμανση. Η αρχική υπόθεση απορρίπτεται αν $|Z| > Z_{\alpha/2}$. Σε πολλά στατιστικά πακέτα χρησιμοποιείται το τετράγωνο του Z και ο έλεγχος γίνεται με κρίσιμες τιμές αυτές που προκύπτουν από την \mathcal{Q}_2 κατανομή με ένα βαθμό ελευθερίας. Φυσικά, ο έλεγχος μπορεί να γενικευτεί για περισσότερες από δύο ομάδες.

ΠΑΡΑΔΕΙΓΜΑ Το αρχείο ενός νοσοκομείου περιλαμβάνει 690 ασθενείς με καρκίνο του μαστού που μετά από χειρουργείο βρέθηκαν με αρνητικούς λεμφαδένες και μικρό μέγεθος όγκου (<20χιλιοστά). Από αυτές τις γυναίκες 537 έλαβαν ορμονική θεραπεία, ενώ οι υπόλοιπες 153 δεν πήραν. Το διάγραμμα 8.5 παρουσιάζει τις καμπύλες επιβίωσης για τις δύο διαφορετικές ομάδες. Φαίνεται πως οι πιθανότητες επιβίωσης αυξάνουν στην συγκεκριμένη ομάδα ασθενών αν αυτοί λαμβάνουν ορμονική θεραπεία. Αυτό επιβεβαιώνεται και από τον έλεγχο που παρουσιάζεται στον πίνακα 8.3.

Οι υπολογισμοί αυτοί γίνονται αυτόματα από το SPSS κάτω από το μενού Analyze > Survival > Kaplan-Meier. Για άλλη μια φορά πρέπει να δηλωθεί η μεταβλητή που δηλώνει το χρόνο και η δείκτρια μεταβλητή για το γεγονός. Στο ίδιο παράθυρο πρέπει να καθοριστεί ο παράγοντας με βάση τον οποίο θέλουμε να γίνει η σύγκριση των συναρτήσεων επιβίωσης. Εν προκειμένω, επιλέγεται η μεταβλητή horm, η οποία παίρνει την τιμή 1 αν μια ασθενής έχει λάβει ορμονοθεραπεία και 0 αν δεν έχει λάβει. Στη συνέχεια, επιλέγεται το compare factor και



Σχήμα 8.8: Καμπύλες επιβίωσης για ασθενείς με καρκίνο του μαστού, με μέγεθος όγκου μικρότερο των 20 χιλιοστών και αρνητικούς λεμφαδένες που έλαβαν ορμονική θεραπεία (...) έναντι των αντίστοιχων που δεν έλαβαν (—). log-rank p-value=0.023

Πίνακας 8.3: Πίνακας παρατηρούμενων και αναμενόμενων αριθμών θανάτων ασθενών με καρκίνο του μαστού.

	# ασθενών	D_k	E_k	$(D_k - E_k)^2 / V$
Θεραπεία	537	65	75.3	5.16
Όχι Θεραπεία	153	41	30.7	5.16

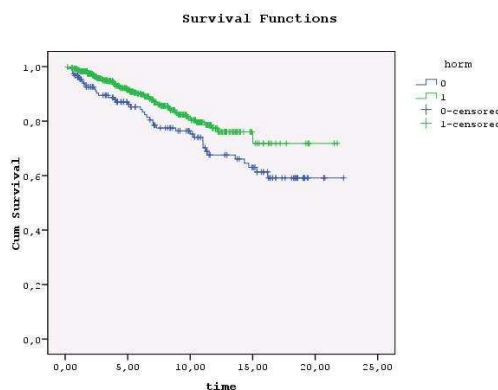
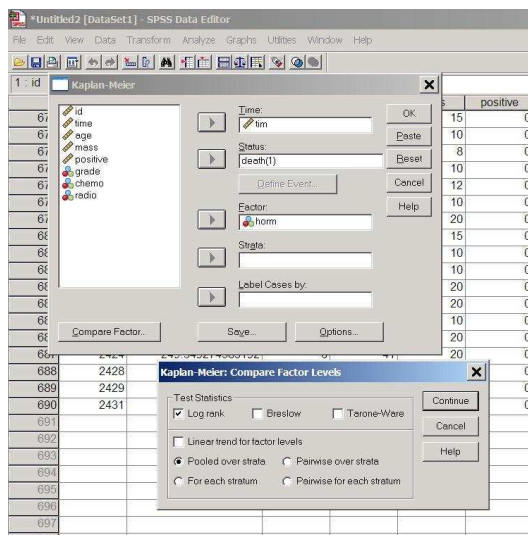
στο υπομενού που προκύπτει δηλώνεται ο στατιστικός έλεγχος log-rank .

Το αποτέλεσμα που προκύπτει είναι ένα διάγραμμα με τις δύο καμπύλες επιβίωσης για τις διαφορετικές ομάδες ασθενών, καθώς και δυο πίνακες με στοιχεία για την ανάλυση.

Ο πρώτος πίνακας περιλαμβάνει περιγραφικά στοιχεία για τις ομάδες. Συνολικά βλέπουμε πως 153 άτομα δεν έλαβαν θεραπεία, από τα οποία 41 απεβίωσαν, ενώ από τα 537 άτομα που έλαβαν θεραπεία τα 65 απεβίωσαν. Στη συνέχεια παρουσιάζεται ο πίνακας με την τιμή που παίρνει η ελεγχοσυνάρτηση, 5.156, τους βαθμούς ελευθερίας και το p-value που αντιστοιχεί.

8.6 Μοντέλο αναλογικών κινδύνων

Στην ανάλυση επιβίωσης, πέρα από την περιγραφή της θνησιμότητας που μπορεί να αναλυθεί με καμπύλες επιβίωσης και κινδύνου, πολύ συχνά ερευνάται πως οι πιθανότητες επιβίωσης



επιηρεάζονται, συσχετίζονται ή μεταβάλλονται ανάλογα με κάποιες πληροφορίες που μπορεί να υπάρχουν για συγκεκριμένους ασθενείς. Οι πληροφορίες αυτές μπορεί να είναι κλινικά δεδομένα που να περιγράφουν την ασθένεια, για παράδειγμα το μέγεθος του όγκου, ή ο αριθμός των διηθημένων λεμφαδένων, καθώς επίσης και τον τύπο της θεραπείας ή φαρμάκων που μπορεί να έλαβε ένας ασθενής, όπως χημειοθεραπεία ή ακτινοθεραπεία. Αρκετές φορές οι πληροφορίες μπορεί να περιγράφουν ποιοτικά χαρακτηριστικά όπως αν ο ασθενής είναι καπνιστής ή όχι.

Το μοντέλο αναλογικού κινδύνου του Cox (proportional hazards model) αποτελεί μια μέθοδο παλινδρόμησης προσαρμοσμένη ειδικά σε δεδομένα που περιέχουν λογοκριμένες παρατηρήσεις. Με βάση αυτό το μοντέλο μπορούν να εξεταστούν σχέσεις ανάμεσα στον κίνδυνο θανάτου και τις πληροφορίες που υπάρχουν για το δείγμα. Το μοντέλο γράφεται ως

$$h(t|X) = h_0(t) \exp(X\beta) \tag{8.19}$$

όπου X είναι ένας $n \times p$ πίνακας που περιέχει σε στήλες τις p ανεξάρτητες μεταβλητές πάνω σε n παρατηρήσεις, β είναι ένα διάνυσμα διαστάσεων $p \times 1$ ενώ $h_0(t)$ είναι μια συνάρτηση βασικού

Case Processing Summary

horm	Total N	N of Events	Censored	
			N	Percent
0	153	41	112	73.2%
1	537	65	472	87.9%
Overall	690	106	584	84.6%

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	5.156	1	.023

Test of equality of survival distributions for the different levels of horm.

κινδύνου (baseline hazard function) η οποία δε χρειάζεται να καθοριστεί πλήρως. Η συνάρτηση βασικού κινδύνου δηλώνει τον κίνδυνο που έχει οποιοδήποτε άτομο από τα δεδομένα, να υποστεί το προκαθορισμένο γεγονός σε ένα χρόνο t ανεξάρτητα από τις τιμές που μπορούν να παίρνουν οι μεταβλητές X .

Το μοντέλο του Cox για να είναι έγκυρο πρέπει να υπακούει σε μια βασική συνθήκη, την υπόθεση της αναλογικότητας, για αυτό ονομάζεται και μοντέλο αναλογικού κινδύνου. Αν θεωρήσουμε δύο ασθενείς, με τιμές συμμεταβλητών X και X^* , η αναλογία των κινδύνων τους γράφεται:

$$\frac{h(t|X)}{h(t|X^*)} = \frac{h_0(t) \exp(X\beta)}{h_0(t) \exp(X^*\beta)} = \exp [(X - X^*)\beta]$$

Συνεπώς, ο λόγος των κινδύνων είναι σταθερός και δεν εξαρτάται από το χρόνο. Η ποσότητα (8.6), (hazard ratio) είναι ο κίνδυνος ενός ατόμου με χαρακτηριστικά X να πάθει το προκαθορισμένο γεγονός σε σχέση με ένα άτομο με χαρακτηριστικά X^* .

8.6.1 Εκτίμηση των παραμέτρων

Η καινοτομία που εισήγαγε ο Cox ήταν η εκτίμηση των παραμέτρων του μοντέλου, μέσω της συνάρτησης μερικής πιθανοφάνειας (partial likelihood function). Έστω ότι έχουμε δεδομένα της μορφής $(T_i, d_i, X_i), i = 1, 2, \dots, n$, όπου T_i είναι οι παρατηρούμενοι χρόνοι επιβίωσης, d_i είναι μια δείκτρια συνάρτηση που παίρνει την τιμή 1 όταν συμβεί το γεγονός, 0 αλλιώς, και ότι η λογοκρισία είναι μη πληροφοριακή. Έστω $t_{(1)}, t_{(2)}, \dots, t_{(D)}$ οι διατεταγμένοι χρόνοι θανάτου (και για απλότητα δεν υπάρχουν δύο ίδιοι χρόνοι) και X_i είναι η i -οστή γραμμή του πίνακα συμμεταβλητών που σχετίζεται με τον χρόνο t_i . Ορίζουμε το σύνολο των ατόμων σε κίνδυνο (risk set) στον χρόνο t_i , $R(t_i)$ ως το πλήθος όλων των ατόμων που βρίσκονται σε κίνδυνο και υπό μελέτη, σε χρόνο ακριβώς πριν το t_i . Η συνάρτηση πιθανοφάνειας γράφεται ως:

$$L(\beta) = \prod_{i=1}^D \frac{\exp(X_i\beta)}{\sum_{j \in R(t_i)} \exp(X_j\beta)} \tag{8.20}$$

Η συνάρτηση περιλαμβάνει στον αριθμητή την πληροφορία μόνο τα άτομα από που παθαίνουν το προκαθορισμένο γεγονός, ενώ στον παρονομαστή την πληροφορία όλων των ατόμων που δεν έχουν ακόμα υποστεί το γεγονός, ακόμα και αυτά που πιθανόν να λογοκριθούν στο μετέπειτα χρόνο. Με άλλα λόγια, μπορούμε να δούμε ότι στην χρονική στιγμή t_i που συμβαίνει το γεγονός, το άτομο με χαρακτηριστικά X_i συγκρίνεται με όλα τα υπόλοιπα άτομα που είναι ακόμα υπό παρακολούθηση. Θα μπορούσαμε να δούμε το μοντέλο σαν μια σειρά από μελέτες ασθενών-μαρτύρων που συμβαίνουν κάθε στιγμή που παρατηρείται ένα γεγονός. Η συνάρτηση μερικής πιθανοφάνειας συνδυάζει όλη αυτή την πληροφορία με τον ίδιο τρόπο που συνδυάζονται δεδομένα από πίνακες χρησιμοποιώντας τη μέθοδο των Mantel-Haenzel.

Ένα από τα προτερήματα της σχέσης 8.20 είναι ότι η εκτίμηση των παραμέτρων δεν εξαρτάται από την συνάρτηση βασικού κινδύνου. Έστω $l(\beta) = \ln L(\beta)$, τότε η (8.20) γράφεται:

$$l(\beta) = \sum_{i=1}^D X_i \beta - \sum_{i=1}^D \ln \left[\sum_{j \in R(t_i)} \exp(X_j \beta) \right] \quad (8.21)$$

Η συνάρτηση μερικής λογαριθμικής πιθανοφάνειας μεγιστοποιείται σαν να ήταν πλήρης πιθανοφάνεια χρησιμοποιώντας τη μέθοδο Newton-Raphson. Οι εξισώσεις των scores βρίσκονται αν πάρουμε την πρώτη μερική παράγωγο της 8.21 ως προς τα β . Έστω $U(\beta) = \partial l(\beta) / \partial \beta$ η συνάρτηση των scores τότε

$$U(\beta) = \sum_{i=1}^D X_i \beta - \sum_{i=1}^D \frac{\sum_{j \in R(t_i)} X_j \exp(X_j \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} \quad (8.22)$$

Ο πίνακας πληροφορίας δίνεται από τον αρνητικό πίνακα της δεύτερης μερικής παραγώγου της (8.21) με το kl στοιχείο του πίνακα να υπολογίζεται ως:

$$I_{kl} = \frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_l} = \sum_{i=1}^D \frac{\sum_{j \in R(t_i)} X_{jk} X_{jl} \exp(X_j \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} - \sum_{i=1}^D \left[\frac{\sum_{j \in R(t_i)} X_{jk} \exp(X_j \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} \right] \left[\frac{\sum_{j \in R(t_i)} X_{jl} \exp(X_j \beta)}{\sum_{j \in R(t_i)} \exp(X_j \beta)} \right]$$

8.6.2 Έλεγχοι υποθέσεων

Το μοντέλο αναλογικού κινδύνου αποτελεί μια ειδική περίπτωση ενός γενικευμένου γραμμικού μοντέλου. Συνεπώς, οι στατιστικοί έλεγχοι που χρησιμοποιούνται σε μοντέλα λογιστικής παλινδρόμησης κ.α. μπορούν να χρησιμοποιηθούν και εδώ με ανάλογο τρόπο. Μετά την εκτίμηση των παραμέτρων ενδιαφέρει να ελεγχθεί η υπόθεση $H_0 : \beta = \beta_0$ όπου β είναι οι εκτιμήσεις για τις άγνωστες παραμέτρους β . Οι τρεις έλεγχοι που μπορούν να χρησιμοποιηθούν είναι ο έλεγχος των scores, ο έλεγχος των λόγων πιθανοφανειών και ο έλεγχος Wald.

Ο έλεγχος που χρησιμοποιεί τις πρώτες παραγώγους της συνάρτησης πιθανοφάνειας χρησιμοποιεί την ελεγχουσυνάρτηση

$$X_{sc}^2 = U(\beta_0)'I^{-1}(\beta_0)U(\beta_0)$$

ακολουθεί την συνάρτηση X^2 με p βαθμούς ελευθερίας, όπου p είναι ο αριθμός των εκτιμηθέντων παραμέτρων. Ο έλεγχος αυτός βασίζεται στις ασυμπτωτικές ιδιότητες των συναρτήσεων των scores (κάτω από την H_0 ακολουθούν πολυμεταβλητή κανονική κατανομή με μέσο 0 και διακύμανση-συνδιακύμανση $I(\beta)$).

Παρόμοια, ο έλεγχος του λόγου πιθανοφανειών ελέγχει την αρχική υπόθεση με την συνάρτηση

$$X_{LR}^2 = 2 \left[l(\hat{\beta}) - l(\beta_0) \right]$$

που ακολουθεί X^2 κατανομή με p βαθμούς ελευθερίας.

Τέλος, ο έλεγχος του Wald βασίζεται στην ασυμπτωτική κανονικότητα των εκτιμητών μέγιστης πιθανοφάνειας και χρησιμοποιεί την συνάρτηση

$$X_W^2 = (\hat{\beta} - \beta_0)'I(\hat{\beta})(\hat{\beta} - \beta_0)$$

που ακολουθεί την κατανομή X^2 με p βαθμούς ελευθερίας. Ο έλεγχος του Wald χρησιμοποιείται τις περισσότερες φορές για τον έλεγχο της υπόθεσης ότι μια μόνο παράμετρος παίρνει μια συγκεκριμένη τιμή, δηλαδή $H_0 : \beta_l = \beta_{l0}$. Για μια μόνο μεταβλητή ο έλεγχος γίνεται ο γνωστός z έλεγχος με συνάρτηση $\hat{\beta}/se(\hat{\beta})$.

Οι τρεις αυτοί έλεγχοι αν και ισοδύναμοι ασυμπτωτικά, μπορεί να έχουν διαφορές σε πεπερασμένα δείγματα. Γενικά, ο έλεγχος πιθανοφανειών θεωρείται ο πιο αξιόπιστος και ο έλεγχος του Wald ο λιγότερο. Και οι τρεις αυτοί έλεγχοι μπορούν να χρησιμοποιηθούν τόσο για να ελέγξουν την υπόθεση ότι μια από τις παραμέτρους, έστω β_1 είναι ίση με μία συγκεκριμένη τιμή (συνήθως 0), ή ένα υποσύνολο των παραμέτρων, π.χ. $\beta = (\beta_1, \beta_2) = \beta_0$ ή τέλος, το σύνολο των εκτιμηθέντων παραμέτρων ισούτε με μια συγκεκριμένη τιμή, όπως περιγράφεται στην προηγούμενη παράγραφο.

ΠΑΡΑΔΕΙΓΜΑ Το μοντέλο αναλογικού κινδύνου εφαρμόστηκε στα δεδομένα του νοσοκομείου που παρουσιάστηκαν στην προηγούμενη ενότητα, για να διαπιστωθεί πως επηρεάζουν από κοινού τα χαρακτηριστικά των ασθενών και η θεραπεία που ακολούθησαν την πιθανότητα θανάτου. Στην ομάδα που μας αφορά μετρήθηκαν, μεταξύ άλλων παραγόντων που δεν συμπεριλήφθηκαν στην ανάλυση, η ηλικία κατά την επέμβαση, το μέγεθος του όγκου, ο αριθμός των διηθημένων λεμφαδένων, η διαφοροποίηση του όγκου, καθώς επίσης και το είδος της θεραπείας που έλαβαν οι ασθενείς, αν δηλαδή υποβλήθηκαν σε χημειοθεραπεία, ακτινοθεραπεία και /ή

ορμονοθεραπεία. Ο πίνακας 3 παρουσιάζει τις εκτιμήσεις των σχετικών κινδύνων μαζί με τα τυπικά σφάλματα (standard errors) και το p -value για κάθε έναν από τους παράγοντες αυτούς, όπως προκύπτουν με τη χρήση του στατιστικού πακέτου P .

Πίνακας 8.4: Αποτελέσματα από παλινδρόμηση κατά Cox. Η ηλικία είναι σε χρόνια, το μέγεθος του όγκου μετράται σε χιλιοστά, λεμφαδένες είναι το πλήθος των διηθημένων λεμφαδένων, Γραδωτή διαφοροποίηση του όγκου, χημειοθεραπεία ακτινοβολία και ορμόνες παίρνουν τις τιμές 0 (όχι) και 1 (ναι).

	$\hat{\beta}$	Τυπικό σφάλμα	Wald	p
Ηλικία	0.017	0.003	4.80	<0.001
Μέγεθος όγκου	0.011	0.002	4.53	<0.001
Λεμφαδένες	0.032	0.003	8.70	<0.001
Διαφοροποίηση	0.373	0.070	5.29	<0.001
Χημειοθεραπεία	0.006	0.099	0.06	0.951
Ακτινοβολία	-0.208	0.095	-2.17	0.029
Ορμόνες	-0.623	0.087	-7.11	<0.001

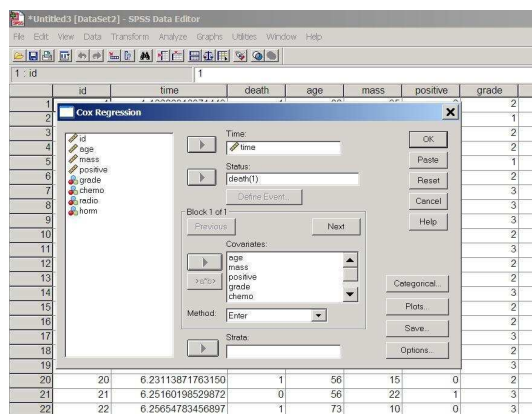
Από τον πίνακα προκύπτουν τα ακόλουθα συμπεράσματα:

- Κάθε επιπλέον χρόνος στην ηλικία μιας ασθενούς αυξάνει τον λογαριθμικό κίνδυνο κατά 0.017, ενώ η διαφορά κινδύνου για τη νεώτερη γυναίκα υπό μελέτη (23 ετών) με τη μεγαλύτερη (98 ετών) είναι 1.275 (δηλαδή 0.017 επί 75 χρόνια). Αν αυτός ο κίνδυνος μεταφραστεί από λογαριθμικό σε φυσιολογικό, προκύπτει πως ο σχετικός κίνδυνος μιας γυναίκας 98 ετών έναντι μιας 23 ετών είναι 3.58 ($=\exp(1.275)$), δηλαδή έχει τρεις και μισή φορές μεγαλύτερο κίνδυνο να πεθάνει από καρκίνο.
- Ο αριθμός των διηθημένων λεμφαδένων είναι ένας πολύ σημαντικός παράγοντας. Κάθε επιπλέον διηθημένος λεμφαδένας συνδέεται με αύξηση του κινδύνου κατά 0.03 στο λογαριθμικό σχετικό κίνδυνο. Αν και αυτό δεν φαίνεται τόσο θεαματικό αρχικά, θα πρέπει να τονιστεί ότι ασθενείς με 10 διηθημένους λεμφαδένες έχουν αυξημένο κίνδυνο κατά τρεις μονάδες σχέση με αυτούς με αρνητικούς. Επιπλέον η σχέση αυτή αποδεικνύεται πως είναι πολύ ισχυρή στατιστικά, όπως δείχνει η μεγάλη τιμή που παίρνει το Wald statistic .
- Η διαφοροποίηση του όγκου έχει θεαματική επίδραση στον κίνδυνο, καθώς το ρίσκο αυξάνει κατά 0.373 (ή σε κανονική κλίμακα 1.45, δηλαδή αύξηση 45%) από όγκου χαμηλής

διαφοροποίησης σε όγκο μεσαίας. Η διαφορά στον κίνδυνο γίνεται πολύ μεγαλύτερη αν πάμε από έναν όγκο χαμηλής διαφοροποίησης σε έναν υψηλής (σχετικός κίνδυνος 2.1 ή αύξησή κατά 110%).

- Το μέγεθος του όγκου παρουσιάζει αυξημένο κίνδυνο κατά 0.011 για κάθε επιπλέον χιλιοστό.
- Η ορμονική θεραπεία μειώνει τον κίνδυνο κατά 0.623, όπως υποδηλώνει το αρνητικό πρόσημο μπροστά από την εκτίμηση. Αν αυτό μετατραπεί στον γνωστό σχετικό κίνδυνο, φαίνεται πως η ορμονική θεραπεία μειώνει τον κίνδυνο κατά 47% ($= \exp(-0.623) = 0.53$).
- Η ακτινοβολία δρα προφυλακτικά μειώνοντας τον κίνδυνο κατά 0.208, εφόσον δρα ως τοπική θεραπεία.
- Τέλος, για την χημειοθεραπεία, η σχέση που προκύπτει δεν είναι στατιστικά σημαντική (p -value = 0.951), συνεπώς οποιαδήποτε συμπερασματολογία δεν θα ήταν αξιόπιστη και ο συντελεστής b μπορεί να θεωρηθεί ίσος με το μηδέν. Αν και φαίνεται λοιπόν πως οι γυναίκες που λαμβάνουν χημειοθεραπεία έχουν αυξημένο κίνδυνο, δεν έχουμε αρκετές ενδείξεις για να υποστηρίξουμε κάτι τέτοιο.

Η ανάλυση αυτή γίνεται από το SPSS από το μενού Analyze > Survival > Cox Regression . Η μεταβλητή που δηλώνει το χρόνο και το γεγονός πρέπει να μπουνε στο δεξιό μέρος στο παράθυρο που προκύπτει. Όλη η πληροφορία που θέλουμε να συμπεριληφθεί στο μοντέλο δηλώνεται ως covariate , ενώ υπάρχουν και επιλογές για την εμφάνιση διαγραμμάτων. Τα αποτελέσματα



συνοψίζονται σε ένα πίνακα παρόμοιο με τον πίνακα που παρουσιάστηκε προηγουμένως.

	B	SE	Wald	df	Sig.	Exp(B)
age	.017	.004	23.080	1	.000	1.018
mass	.012	.003	20.604	1	.000	1.012
positive	.033	.004	75.753	1	.000	1.033
grade	.373	.070	28.085	1	.000	1.452
chemo	.006	.100	.004	1	.951	1.006
radio	-.209	.096	4.750	1	.029	.811
horn	-.623	.088	50.605	1	.000	.538

Ο πρώτος έλεγχος υποθέσεων που μπορεί να γίνει είναι να δούμε ποιές από τις μεταβλητές που υπάρχουν στο μοντέλο δεν υποδυκνείουν μια στατιστικά σημαντική σχέση. Ο πίνακας περιλαμβάνει την στήλη με τις τιμές της στατιστικής συνάρτησης κατά Wald και δίπλα τις κρίσιμες τιμές που αντιστοιχούν σε αυτές p-value. Από τον έλεγχο αυτό μπορούμε να συμπεράνουμε πως δεν έχουμε αρκετά στοιχεία για να δείξουμε πως η χημειοθεραπεία μειώνει (ή αυξάνει) τον κίνδυνο θανάτου από καρκίνο του μαστού σε αυτά τα δεδομένα, όπως φαίνεται από την κρίσιμη τιμή του ελέγχου (0.95).

Ας υποθέσουμε πως θέλουμε να δούμε πως επηρεάζεται ο κίνδυνος θανάτου των ασθενών, ανεξάρτητα από την θεραπεία που έχουν ακολουθήσει, μόνο με βάση την πληροφορία που υπάρχει στα κλινικά χαρακτηριστικά τους, δηλαδή την ηλικία, το μέγεθος του όγκου, τον αριθμό των διηθυσμένων λεμφαδένων και τη διαφοροποίηση του όγκου. Ο πίνακας που ακολουθεί παρουσιάζει τις εκτιμήσεις των παραμέτρων αυτού του μοντέλου:

Πίνακας 8.5: Αποτελέσματα από παλινδρόμηση κατά Cox. Η ηλικία είναι σε χρόνια, το μέγεθος του όγκου μετράται σε χιλιοστά, λεμφαδένες είναι το πλήθος των διηθυσμένων λεμφαδένων, Γραδε η διαφοροποίηση του όγκου

	$\hat{\beta}$	Τυπικό σφάλμα	Wald	p
Ηλικία	0.013	0.003	3.89	<0.001
Μέγεθος όγκου	0.013	0.002	5.15	<0.001
Λεμφαδένες	0.035	0.003	9.87	<0.001
Διαφοροποίηση	0.345	0.070	4.91	<0.001

Το δεύτερο μοντέλο έχει μικρότερο αριθμό παραμέτρων και όπως είναι αναμενόμενο, οι εκτιμήσεις για τους συντελεστές είναι διαφοροποιημένες από τις αντίστοιχες του μοντέλου με όλη την πληροφορία. Σύμφωνα με τον έλεγχο Wald όλοι οι συντελεστές είναι στατιστικά σημαντική. Το ερώτημα που θα πρέπει να απαντηθεί είναι πιο μοντέλο είναι καλύτερο, ποιο μοντέλο δηλαδή προσαρμόζει καλύτερα στα δεδομένα. Για να απαντηθεί το ερώτημα μπορεί να χρησιμοποιηθεί ο έλεγχος πιθανοφανειών. Το πλήρες μοντέλο είναι αυτό που περιλαμβάνει όλη την πληροφορία,

μερικό μοντέλο είναι αυτό με τις μεταβλητές που αφορούν μόνο τα κλινικά χαρακτηριστικά των ασθενών, ενώ μηδενικό μοντέλο (null) θεωρείται εκείνο το οποίο δεν περιλαμβάνει καμία ανεξάρτητη μεταβλητή. Για κάθε ένα από τα μοντέλα αυτά υπολογίζεται η τιμή της μερικής πιθανοφάνειας, σύμφωνα με τον τύπο (8.21). Οι τιμές που προκύπτουν είναι -4123.519 για το πλήρες, -4150.684 για το μερικό και -4255.621 για το μηδενικό. Για να συγκριθεί το μηδενικό μοντέλο με το μερικό χρησιμοποιείται ο έλεγχος πιθανοφανειών. Η διαφορά των πιθανοφανειών μεταξύ των μοντέλων είναι $-4150.684 - (-4255.621) = 104.937$. Η διαφορά αυτή πολλαπλασιάζεται με 2 έτσι ώστε η τιμή του ελέγχου είναι 209.874 . Η τιμή αυτή συγκρίνεται με την αντίστοιχη τιμή που δίνει μια X^2 κατανομή με 4 βαθμούς ελευθερίας, αφού στο μερικό μοντέλο έχουμε προσθέσει 4 παραμέτρους. Το αποτέλεσμα του ελέγχου δίνει μια κρίσιμη τιμή πολύ κοντά στο 0, συνεπώς η υπόθεση ότι τα δύο μοντέλα είναι ισοδύναμα απορρίπτεται. Το ενδιάμεσο μοντέλο θεωρείται ότι προσαρμόζει καλύτερα στα δεδομένα από ότι το μηδενικό.

Αντίστοιχα συγκρίνεται το ενδιάμεσο μοντέλο με το πλήρες. Η τιμή του στατιστικού είναι 54.32 που πρέπει να συγκριθεί με την τιμή της X^2 κατανομής με τρεις βαθμούς ελευθερίας, αφού στο πλήρες μοντέλο έχουν προστεθεί οι τρεις μεταβλητές που σχετίζονται με την θεραπεία. Και εδώ προκύπτει ότι το μοντέλο με τις περισσότερες παραμέτρους είναι προτιμότερο.

8.6.3 Εκτίμηση της βασικής συνάρτησης κινδύνου και της συνάρτησης επιβίωσης

Η εκτίμηση της συνάρτησης επιβίωσης $S(t|X)$ από το μοντέλο, απαιτεί την εκτίμηση της βασικής συνάρτησης κινδύνου. Ο Breslow προτείνει ως εκτιμητή της βασικής συνάρτησης κινδύνου τη συνάρτηση:

$$\hat{h}_0(t) = \frac{d_i}{\sum_{j \in R(t_i)} \exp(\mathbf{X}_j \hat{\beta})} \quad (8.23)$$

Κατά συνέπεια, η βασική αθροιστική συνάρτηση κινδύνου δίνεται από τον τύπο

$$\hat{H}_0(t) = \sum_{t_i \leq t} \hat{h}_0(t)$$

ενώ η βασική συνάρτηση επιβίωσης $S_0(t)$ εκτιμάται από $\hat{S}_0(t) = \exp[-\hat{H}_0(t)]$. Η βασική συνάρτηση επιβίωσης δίνει τις πιθανότητες επιβίωσης για ένα άτομο όταν όλες οι τιμές των συμμεταβλητών X είναι 0.

Αν όλες αυτές οι ποσότητες έχουν υπολογιστεί, τότε η πιθανότητα επιβίωσης ενός ατόμου με συμμεταβλητές $X = X_0$ υπολογίζεται ως:

$$\hat{S}(t|X_0) = \hat{S}_0(t)^{\exp(X_0 \hat{\beta})} \quad (8.24)$$

ΠΑΡΑΔΕΙΓΜΑ Στα δεδομένα των ασθενών από το νοσοκομείο καταλήξαμε ότι το μοντέλο που περιλαμβάνει τις μεταβλητές για τα κλινικά χαρακτηριστικά και την θεραπευτική αγωγή είναι το καταλληλότερο για την περιγραφή του κινδύνου. Με αφετηρία τις εκτιμήσεις από το πλήρες μοντέλο προκύπτει η συνάρτηση κινδύνου και η αθροιστική συνάρτηση κινδύνου.

Αν θεωρηθεί ένας υποθετικός ασθενής 30 ετών, με μέγεθος όγκου 10 χιλιοστά, που δεν έχει διηθημένους λεμφαδένες, διαφοροποίηση όγκου 2, που δεν έχει λάβει χημειοθεραπεία, έχει κάνει ακτινοβολία και παίρνει ορμονική θεραπεία, εφαρμόζεται ο τύπος 8.24:

$$\hat{S}(t|X) = \hat{S}_0(t)^{\exp(30*0.017+10*0.011+0*0.032+2*0.373+0*0.006-1*0.208-1*0.623)}$$

ενώ αντίστοιχα υπολογίζεται η πιθανότητα επιβίωσης ενός ασθενούς με ίδια κλινικά χαρακτηριστικά, που δεν έχει λάβει όμως ορμονική θεραπεία.

8.7 Επεκτάσεις - Ειδικά θέματα

Το μοντέλο αναλογικού κινδύνου αποτελεί ένα πολύ χρήσιμο εργαλείο στην ανάλυση επιβίωσης, κάτω όμως από την βασική υπόθεση ότι ισχύει η αναλογικότητα του κινδύνου. Στην πράξη, δεν είναι λίγες οι φορές που αυτή η υπόθεση καταρρέει και τότε οι εκτιμήσεις που δίνει το μοντέλο δεν είναι αξιόπιστες. Υπάρχουν τέσσερις βασικοί λόγοι για τους οποίους μπορεί να μην ισχύει η υπόθεση της αναλογικότητας.

Ο πρώτος και πιο συχνός λόγος είναι η εφαρμογή του μοντέλου σε μελέτες με πολύ μακρά παρακολούθηση των ασθενών. Έστω μια μελέτη που διαρκεί τρία χρόνια και κατά την οποία κάποια από τα άτομα που εισέρχονται υποβάλλονται σε μια φαρμακευτική αγωγή κατά τους 6 πρώτους μήνες, ενώ κάποια άλλα άτομα δεν λαμβάνουν την αγωγή αυτή. Το φάρμακο δρα προστατευτικά για μια χρονική περίοδο, μειώνοντας τον κίνδυνο υποτροπής στα άτομα που το έλαβαν. Μετά το πέρας όμως των έξι μηνών, είναι λογικό να θεωρήσουμε ότι η επίδραση του φαρμάκου σταδιακά εξασθενεί, έως ότου εξαφανιστεί τελείως κάποια στιγμή αργότερα. Το μοντέλο αναλογικού κινδύνου θεωρεί ότι η αναλογία κινδύνου μεταξύ αυτών που δεν έλαβαν και έλαβαν το φάρμακο παραμένει σταθερή. Με άλλα λόγια, το μοντέλο υποθέτει ότι η επίδραση του φαρμάκου παραμένει η ίδια όσο κι αν περάσει ο χρόνος. Κάτι τέτοιο φυσικά, είναι αντίθετο με την ιατρική επιχειρηματολογία.

Για να ελεγχθεί αυτή η περίπτωση ο πιο πρακτικός τρόπος είναι η ανάλυση των δεδομένων πρώτα με καμπύλες Kaplan-Meier . Αν οι καμπύλες επιβίωσης των ατόμων που δεν έλαβαν με αυτών που έλαβαν το φάρμακο, τέμνονται τότε η υπόθεση της αναλογικότητας δεν ισχύει. Παρόλα αυτά, αυτός ο έλεγχος μπορεί να γίνει μόνο σε περιπτώσεις κατηγορικών μεταβλητών. Ο Therneau προτείνει έναν στατιστικό έλεγχο για να αξιολογηθεί αν η υπόθεση της αναλογικότητας

ισχύει, βασισμένο πάνω στα martingale κατάλοιπα. Ο έλεγχος αυτός είναι διαθέσιμος μόνο σε εξειδικευμένα στατιστικά πακέτα όπως το S-plus, R.

Στις περιπτώσεις αυτές, για να αντιμετωπισθεί το πρόβλημα της μη αναλογικότητας των κινδύνων, η πιο απλή προσέγγιση είναι να εφαρμοστεί ένα μοντέλο για τα πρώτα χρόνια της παρακολούθησης και ένα δεύτερο μοντέλο για αργότερα. Φυσικά τίθεται το θέμα πως θα επιλεγεί το διάστημα του χρόνου στο οποίο εφαρμόζεται το κάθε μοντέλο. Η τακτική αυτή μπορεί να χρησιμοποιηθεί μόνο προσεγγιστικά και με προσοχή. Μια εναλλακτική προσέγγιση είναι να επεκταθεί το μοντέλο έτσι ώστε να περιλαμβάνει χρονο-μεταβλητές εκτιμήσεις για τις παραμέτρους β . Αυτό γίνεται αν συμπεριληφθεί στην εξίσωση κινδύνου η αλληλεπίδραση μιας μεταβλητής με κάποιες συναρτήσεις του χρόνου:

$$h(t|X) = h_0(t) \exp(X\beta_1 + X\beta_2 f(t))$$

όπου $f(t)$ μπορεί να είναι μια οποιαδήποτε συνάρτηση του χρόνου που καθορίζεται από τον ερευνητή. Προσοχή χρειάζεται στα μοντέλα αυτά στον καθορισμό της συνάρτησης πιθανοφάνειας καθώς επίσης και στην ερμηνεία των αποτελεσμάτων. Στα στατιστικά πακέτα SPSS, S-PLUS, SAS, R, STATA υπάρχουν έτοιμες ρουτίνες που μπορούν να υπολογίσουν αυτά τα μοντέλα.

Ένας δεύτερος λόγος που μπορεί να οδηγήσει στην κατάρρευση της υπόθεσης της αναλογικότητας είναι η παρουσία συστάδων στα δεδομένα (clusters). Συστάδες δημιουργούνται σε μεγάλες πολυκεντρικές κλινικές δοκιμές, όπου οι ασθενείς που συμμετέχουν στη μελέτη προέρχονται από διαφορετικές κλινικές. Η παρουσία αυτής της ετερογένειας μπορεί να αντιμετωπιστεί είτε με στρωματοποιημένα μοντέλα (που παρουσιάζονται στη συνέχεια), είτε με μικτά μοντέλα που στην ανάλυση επιβίωσης καλούνται μοντέλα ευπάθειας (frailty models).

8.7.1 Στρωματοποιημένα μοντέλα

Μια επέκταση του μοντέλου αναλογικού κινδύνου περιλαμβάνει την ανάλυση κατά στρώματα. Αν τα δεδομένα προέρχονται από στρωματοποιημένη δειγματοληψία, όπως για παράδειγμα δεδομένα από μια πολυκεντρική κλινική δοκιμή, τα στρώματα χωρίζουν τους ασθενείς σε διακριτές ομάδες ασθενών που προέρχονται από την ίδια κλινική. Τότε, με το στρωματοποιημένο μοντέλο αναλογικού κινδύνου ορίζεται διαφορετική βασική συνάρτηση κινδύνου για κάθε στρώμα, αλλά οι συντελεστές β των μεταβλητών παραμένουν ίδιοι για όλα τα στρώματα. Έστω n_1 ασθενείς που προέρχονται από το ίδιο νοσοκομείο, n_2 από ένα δεύτερο και n_k οι ασθενείς από το k κέντρο. Τότε ο κίνδυνος για ένα ασθενή i που ανήκει στο j ίδρυμα είναι $h(t|X) = h_{0j}(t) \exp(X\beta)$.

Στην περίπτωση της στρωματοποιημένης ανάλυσης, η συνάρτηση μερικής πιθανοφάνειας γράφεται ως το άθροισμα των επιμέρους συναρτήσεων πιθανοφάνειας για κάθε στρώμα $l_j(\beta), j =$

$1, 2, \dots, k$. Όπου $l_j(\beta)$ είναι η σχέση (8.21) αλλά ορισμένη μόνο στα άτομα που ανήκουν στο j στρώμα. Αντίστοιχα προκύπτουν και οι συναρτήσεις των scores και του πίνακα πληροφορίας. Τα περισσότερα στατιστικά πακέτα περιλαμβάνουν ρουτίνες για στρωματοποιημένα μοντέλα αναλογικού κινδύνου.

8.8 Άλλες προσεγγίσεις

Εκτός από το μοντέλο αναλογικού κινδύνου υπάρχουν και μια σειρά από άλλα μοντέλα τα οποία χρησιμοποιούνται στην ανάλυση επιβίωσης. Τα περισσότερα από αυτά, αν και έχουν αρκετές χρήσιμες ιδιότητες, δεν είναι ιδιαίτερα δημοφιλή, εκτός από εξειδικευμένες περιπτώσεις και εφαρμογές.

Στην ανάλυση αξιοπιστίας, τις περισσότερες φορές χρησιμοποιούνται παραμετρικά μοντέλα, στα οποία υπάρχει η βασική υπόθεση ότι ο χρόνος ακολουθεί μια συγκεκριμένη κατανομή. Σε αυτή την κατηγορία μοντέλων ανήκουν και τα accelerated failure time models. Σε μερικές περιπτώσεις, αν είναι γνωστή η κατανομή που ακολουθεί ο χρόνος, μπορεί να έχουν πιο αξιόπιστα αποτελέσματα, ιδίως όσον αφορά τις προβλέψεις. Παρόλα αυτά, στην βιοστατιστική πολύ σπάνια κάποιο τέτοιο μοντέλο συμπεριφέρεται καλύτερα από το μοντέλο αναλογικού κινδύνου.

Μια άλλη μορφή μικτών μοντέλων αποτελούν τα cure rate μοντέλα, σύμφωνα με τα οποία, κάποια από τα άτομα που συμμετέχουν στη μελέτη μπορεί να αποκτήσουν ανοσία στην ασθένεια και συνεπώς η πιθανότητα να υποτροπιάσουν είναι πάντα 0. Τα μοντέλα αυτά δεν έχουν εφαρμογές σε μελέτες που αφορούν καρδιαγγειακές παθήσεις, καρκίνο, AIDS και άλλες ασθένειες που θεωρούνται ανίατες. Σε διαφορετικές περιπτώσεις μπορούν να φανούν ιδιαίτερα χρήσιμα, αν και δεν υπάρχει κανένα στατιστικό πρόγραμμα που να περιλαμβάνει ρουτίνες για την εκτίμησή τους.

Τέλος, επιγραμματικά αναφέρουμε πως υπάρχουν προσεγγίσεις στην ανάλυση επιβίωσης που χρησιμοποιούν μοντέλα κατά Bayes, νευρωνικά δίκτυα, το μοντέλο αθροιστικού κινδύνου του Aalen και άλλα, τα οποία όμως δεν είναι εύκολα εφαρμόσιμα από μη ειδικούς.