

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ



ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ

ΔΠΜΣ: «Εφαρμοσμένες Μαθηματικές Επιστήμες»

Αιτιώδης συμπεραματολογία: Αιτιώδες μοντέλο Rubin και Δομικά μοντέλα εξισώσεων

Φοιτήτρια:
ΠΛΙΑΤΣΙΚΑ
ΑΝΤΩΝΙΝΑ

Επιβλέπων:
ΚΟΥΚΟΥΒΙΝΟΣ
ΧΡΗΣΤΟΣ
ΚΑΘΗΓΗΤΗΣ Ε.Μ.Π.

4 Απριλίου 2014

Αφιερωμένο στον άνθρωπο
που στέκεται δίπλα μου
τα τελευταία 11 χρόνια
της ζωής μου.

Ευχαριστίες

Η παρούσα εργασία είναι αποτέλεσμα εκτενούς αναζήτησης και έρευνας. Θα ήθελα να απευθύνω θερμές ευχαριστίες στον επιβλέποντα καθηγητή μου κ. Χρήστο Κουκουβίνο, Καθηγητή Ε.Μ.Π. για την εμπιστοσύνη και την ανάθεση της διπλωματικής εργασίας. Ευχαριστώ επίσης την υποψήφια διδάκτωρα Χριστίνα Παρπούλα για την καθοδήγηση και την άμεση και ουσιαστική βοήθεια που μου παρείχε κατά τη διάρκεια της εκπόνησης αυτής της διπλωματικής εργασίας. Αισθάνομαι ιδιαίτερη ανάγκη να ευχαριστήσω τη συμφοιτήτριά μου Αικατερίνη Γιαννακοπούλου για τη σημαντική βοήθεια, ανταλλαγή ιδεών, απόψεων και κουράγιου σε όλο αυτό το διάστημα. Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου που με την ανιδιοτελή τους υποστήριξη και με στερήσεις, ενθαρρύνει τις προσπάθειές μου καθ' όλη τη διάρκεια των μαθητικών αλλά και ακαδημαϊκών μου σπουδών. Χρωστάω σε όλους ένα μεγάλο ευχαριστώ.

Περίληψη

Προβλήματα που περιλαμβάνουν την αιτιώδη συμπερασματολογία έχουν απασχολήσει τη Στατιστική από την αρχή της εμφάνισής τους. Σκοπός αυτής της μελέτης είναι να παρουσιάσει τα δυο πλαίσια της αιτιώδους συμπερασματολογίας, μια προσέγγιση στατιστικής ανάλυσης που βασίζεται στο πλαίσιο των ενδεχόμενων αποτελεσμάτων και μιας στατιστικής τεχνικής για τον έλεγχο και την εκτίμηση των αιτιωδών σχέσεων χρησιμοποιώντας τον συνδυασμό στατιστικών δεδομένων και ποιοτικών υποθέσεων.

Παρουσιάζουμε το πρώτο παλίσιο της αιτιώδους συμπερασματολογίας που αναφέρεται κοινώς ως ‘Rubin αιτιώδες μοντέλο’ (RCM). Το πλαίσιο έχει δυο βασικά μέρη. Το πρώτο μέρος του (RCM) ορίζει τις αιτιώδεις επιδράσεις μέσω των ενδεχόμενων αποτελεσμάτων και το δεύτερο αφορά το μηχανισμό ανάθεσης. Θέματα που εξετάζονται περιλαμβάνουν, τη σπουδαιότητα της τυχαιοποίησης ως μη συγχυτική μέθοδο της ανάθεσης, μεθόδους για τη διαχείριση της μη συμμόρφωσης και των ελλειπών δεδομένων και μεθόδους για τον περιορισμό της μεροληψίας στην ανάλυση των παρατηρήσεων μελέτης. Το δεύτερο πλαίσιο της αιτιώδους συμπερασματολογίας αναφέρεται ως ‘Δομικά μοντέλα εξισώσεων’ (SEM). Παρουσιάζουμε τα εργαλεία για την ανάγνωση και εξηγούμε υποθέσεις που ενσωματώνονται σε ένα SEM. Αποφασίζουμε μαθηματικώς αν οι αιτιώδεις σχέσεις μπορούν να εκτιμηθούν από τα δεδομένα και αν όχι, ποιες επιπρόσθετες υποθέσεις, μετρήσεις ή πειράματα τις καθιστούν εκτιμήσιμες. Επίσης παραθέτουμε μεθόδους αναγνώρισης και δημιουργίας ενός ισοδύναμου μοντέλου. Τέλος παρουσιάζουμε μια γενίκευση των SEM σε κατηγορικά δεδομένα και μη γραμμικές συσχετίσεις, περιλαμβάνοντας μια λύση στο πρόβλημα της μεσολάβησης.

Abstract

Problems involving causal inference have concerned Statistics since its earliest days. The purpose of this study is, to introduce two common frameworks for causal inference an approach to the statistical analysis of cause and effect based on the framework of potential outcomes, and a statistical technique for testing and estimating causal relations using a combination of statistical data and qualitative causal assumptions.

We present the first framework for causal inference that is commonly referred as “Rubin’s Causal Model” (RCM). The framework has two essential parts. The first part of the RCM defines causal effects through potential outcomes and the second part concerns the assignment mechanism. Topics considered include, the importance of randomization as an unconfounded method of assignment, methods for handling non compliance and missing data and methods for limiting bias in the analysis of observational data. The second framework for causal inference advert as “Structural Equation Modeling” (SEM). We introduce tools of reading and explicating the causal assumptions embodied in SEM models and methods for identifying the testable implications of the assumptions in SEM. Deciding mathematically if the causal relationships are estimable from the data and if not, what additional assumptions, measurements or experiments would render them estimable. Also we display methods of recognizing and generating an equivalent models. In last we present a generalization of SEM to categorical data and nonlinear interactions, including a solution to the so-called “mediation problem”.

Περιεχόμενα

	Σελίδα
1 Αιτιώδης συμπερασματολογία	10
1.1 Βασικές έννοιες	10
1.1.1 Πως ορίζεται η αιτία, η επίδραση και η αιτιώδης σχέση . .	10
1.2 Αιτιότητα	14
1.2.1 Ιστορική Αναδρομή	14
1.3 Ορισμοί και έννοιες αιτιότητας	16
1.4 Αιτιώδης συμπερασματολογία στη Στατιστική	19
1.4.1 Εισαγωγή	19
1.4.2 Η βασική διάκριση και οι διακλαδώσεις της	20
1.5 Το μοντέλο για τη συμπερασματολογία των συσχετίσεων	22
1.6 Το μοντέλο του <i>Rubin</i> για την αιτιώδη συμπερασματολογία . . .	24
1.6.1 Το Θεμελιώδες πρόβλημα της αιτιώδους συμπερασματολογίας	27
1.7 Δομικά μοντέλα εξισώσεων	29
1.7.1 Βασικές γνώσεις γραφημάτων	32
1.7.2 Η μέθοδος του <i>Wright</i> για την ανάλυση μονοπατιών . . .	34
1.8 Σύνδεση του αιτιώδους μοντέλου του <i>Rubin</i> με τα δομικά μοντέλα εξισώσεων	35
2 Αιτιώδες μοντέλο Rubin	37
2.1 Εισαγωγή	37
2.1.1 Μονάδες, θεραπείες και ενδεχόμενα αποτελέσματα	38
2.1.2 Αιτιώδεις επιδράσεις: η αντιγραφή και η υπόθεση <i>SUTVA</i>	40
2.1.3 Συμμεταβλητές	43
2.1.4 Ιστορική αναφορά στα ενδεχόμενα αποτελέσματα	44
2.2 Μηχανισμός ανάθεσης	46
2.2.1 Διευκρινίζοντας τη κρισιμότητα του μηχανισμού ανάθεσης	46
2.2.2 Το παράδοξο του <i>Lord</i>	49
2.2.3 Είδη μηχανισμών ανάθεσης	51
2.3 Μορφές αιτιώδους συμπερασματολογίας που βασίζονται στην α- νάθεση	53

2.3.1	Συμπερασματολογία βασισμένη στη τυχαιοποίηση κατά <i>Fisher</i>	54
2.3.2	Συμπερασματολογία βασισμένη στη τυχαιοποίηση κατά <i>Neyman</i>	56
2.3.3	Συμπερασματολογία βασισμένη σε ένα μοντέλο	57
2.3.4	Ο ρόλος των συµµεταβλητών στα τυχαιοποιηµένα πειράµατα	59
2.3.5	Πιο σύνθετα τυχαιοποιηµένα πειράµατα	60
2.4	Κίνδυνοι µη συγχυτικών µηχανισµών ανάθεσης	61
2.4.1	Μη-συµµόρφωση	61
2.4.2	Ελλιπή δεδοµένα	65
2.5	Η αιτιώδης συµπερασματολογία στις µελέτες παρατήρησης	67
2.5.1	Μέθοδοι προσαρµογής για συγχυτικούς παράγοντες	67
3	Δομικά μοντέλα εξισώσεων	69
3.1	Εισαγωγή	69
3.2	Η λογική των <i>SEM</i>	70
3.3	Η αιτιώδης ερµηνεία των δοµικών µοντέλων εξισώσεων	73
3.3.1	Ντετερµινιστικό αιτιώδες µοντέλο	76
3.3.2	Αιτιώδεις υποθέσεις σε µη παραµετρικά µοντέλα	76
3.3.3	Παρουσιάζοντας τις παρεµβάσεις και τις αιτιώδεις επιδράσεις	78
3.3.4	Πληθυσµιακό αιτιώδες µοντέλο	82
3.3.5	Τα µη παραµετρικά <i>SEM</i> ως <i>DAGs</i>	84
3.3.6	<i>Counterfactual</i> ανάλυση στα δοµικά µοντέλα εξισώσεων	86
3.3.7	Σύνδεση µε το <i>framework</i> του ενδεχόµενου αποτελέσµατος	94
3.4	Συµπεράσµατα δοµικών εξισώσεων προς έλεγχο	95
3.4.1	Το κριτήριο του <i>d</i> -διαχωρισµού	97
3.4.2	Οι κανόνες του <i>do</i> -λογισµού	100
3.4.3	Ισοδύναµα µοντέλα	101
3.5	Η χρήση γραφηµάτων στο ‘ <i>Back – Door</i> ’ κριτήριο	102
3.5.1	Αναγνώριση παραµέτρων και αιτιωδών επιδράσεων	103
3.5.2	Η σηµασία των συµµεταβλητών	104
3.5.3	Παραµετρικός προσδιορισµός στα γραµµικά <i>SEM</i>	106
3.5.4	Αναγνωρίζοντας τις βοηθητικές µεταβλητές	107
3.6	Μεσολάβηση: άµεσες και έµµεσες επιδράσεις	108

3.6.1	Διάσπαση	108
3.6.2	Ο Τύπος της μεσολάβησης: Μία απλή λύση για το ακαν- θώδες πρόβλημα	111
	Βιβλιογραφία	117

Κεφάλαιο 1

Αιτιώδης συμπερασματολογία

1.1 Βασικές έννοιες

1.1.1 Πως ορίζεται η αιτία, η επίδραση και η αιτιώδης σχέση

Οι περισσότεροι άνθρωποι αναγνωρίζουν ενστικτωδώς τις αιτιώδεις σχέσεις στην καθημερινή τους ζωή. Ο αριθμός των ωρών που ξοδεύει κάποιος για μελέτη, αποτελούν μια αιτία για τους βαθμούς του σε ένα διαγώνισμα, ή η ποσότητα της τροφής που τρώει ένα άτομο είναι μια αιτία για το βάρος του. Μπορούν ακόμη να αποδειχτούν πιο περίπλοκες αιτιώδεις σχέσεις, παρατηρώντας ότι ένας χαμηλός βαθμός σε ένα διαγώνισμα είναι αποθαρρυντικός, με αποτέλεσμα τη μείωση της μελέτης στο επόμενο διαγώνισμα, με συνέπεια τους ακόμα πιο χαμηλούς βαθμούς. Εδώ η ίδια μεταβλητή (χαμηλός βαθμός) μπορεί να είναι και μια αιτία και μια επίδραση, και μπορεί να υπάρξει μια αμοιβαία σχέση μεταξύ δύο μεταβλητών (χαμηλοί βαθμοί και έλλειψη μελέτης) που προκαλούν η μια την άλλη.

Παρά την ενστικτώδη οικειότητά μας σχετικά με τις αιτιώδεις σχέσεις, ένας ακριβής ορισμός της αιτίας και της επίδρασης διαφεύγει για αιώνες από τις θεωρίες των φιλόσοφων. Πράγματι οι ορισμοί των όρων όπως η αιτία και η επίδραση εξαρτώνται μερικώς η μια από την άλλη, και από την αιτιώδη σχέση στην οποία και οι δύο ενσωματώνονται. Ο φιλόσοφος του 17ου αιώνα *John Locke* είχε δηλώσει:

‘Αυτό που παράγει οποιαδήποτε απλή ή σύνθετη ιδέα, ορίζεται από τη γενική ονομασία αιτία, και αυτό που παράγεται είναι η επίδραση’.

Και επίσης,

‘Μια αιτία είναι αυτή που παράγει οτιδήποτε άλλο, είτε αυτό είναι μια απλή ιδέα, είτε μια υπόσταση, είτε ένας τρόπος, και έτσι ξεκινά να είναι μια επίδραση, της οποίας η αρχή της οφείλεται σε ένα άλλο γεγονός’.

Έπειτα από αυτές τις δηλώσεις, άλλοι φιλόσοφοι και επιστήμονες μας έδωσαν χρήσιμους ορισμούς αυτών των τριών βασικών εννοιών (αιτία, επίδραση και αιτιώδεις σχέσεις), που

είναι πιο συγκεκριμένες και εξηγούν ορθότερα πως δουλεύουν τα πειράματα. Δεν θα υπερασπιστούμε οποιονδήποτε από αυτούς τους ορισμούς ως σωστότερο, δεδομένου ότι έχει αποφευχθεί από τους φιλόσοφους για χιλιετίες. Ισχυριζόμαστε όμως ότι αυτές οι έννοιες μας βοηθούν να αποσαφηνίσουμε την επιστημονική πρακτική της εξέτασης των αιτιών.

Η Αιτία

Ας θεωρήσουμε την αιτία μιας φωτιάς σε ένα δάσος. Γνωρίζουμε ότι μια φωτιά μπορεί να προκληθεί με αρκετούς διαφορετικούς τρόπους—ένα σπύρτο πετάγεται από ένα αυτοκίνητο, μια αστραπή, ή μια φωτιά που σιγοκαίει σε κάμπιγκ προσκόπων. Καμιά από αυτές τις αιτίες δεν είναι απαραίτητα, γιατί μια φωτιά στο δάσος μπορεί να ξεκινήσει ακόμα και χωρίς τη παρουσία ενός σπύρτου. Επίσης, καμιά από αυτές δεν επαρκεί για να ξεσπάσει μια φωτιά. Ένα σπύρτο πρέπει να παραμείνει αναμένο για αρκετή ώρα για να ξεσπάσει πυρκαγιά, και πρέπει να έρθει σε επαφή με καύσιμο υλικό, όπως τα ξερά φύλλα, να υπάρχει οξυγόνο για την καύση, καθώς επίσης η έλλειψη υγρασίας. Έτσι το σπύρτο αποτελεί μέρος μιας αλληλουχίας συνθηκών εκ των οποίων όλες είναι απαραίτητες για να προκαλέσουν φωτιά, αν και μερικές από αυτές τις συνθήκες λαμβάνονται σαν δεδομένες, όπως η διαθεσιμότητα οξυγόνου. Ένα αναμένο σπύρτο είναι, όπως αποκαλεί ο Mackie [35, p.62] μια *inipus* συνθήκη – ‘ένα ανεπαρκές αλλά μη περιττό μέρος ενός μη απαραίτητου αλλά επαρκούς κατάστασης’.

1. Είναι ανεπαρκές επειδή ένα σπύρτο δεν μπορεί να αρχίσει μια πυρκαγιά χωρίς τις υπόλοιπες προϋποθέσεις.
 2. Είναι μη περιττό μόνο εάν προστίθεται κάποιο άλλο προωθητικό φωτιάς που είναι διαφορετικό από τους άλλους παράγοντες που συμβάλλουν για τη δημιουργία της (π.χ. το οξυγόνο, τα ξερά φύλλα)
- . Εξάλλου, θα ήταν πιο δύσκολο να ειπωθεί ότι το σπύρτο προκάλεσε την πυρκαγιά εάν κάποιος άλλος προσπαθούσε ταυτόχρονα να την ανάψει με έναν αναπτήρα. Είναι μέρος μιας επαρκούς κατάστασης για να αρχίσει μια πυρκαγιά με το συνδυασμό όλων των υπολοίπων συστατικών. Αλλά το σπύρτο δεν είναι απαραίτητο σε όλη τη κατάσταση διότι υπάρχουν άλλα σύνολα καταστάσεων που μπορούν να ξεκινήσουν μια πυρκαγιά.

Ένα ερευνητικό παράδειγμα μιας *inipus* κατάστασης, αφορά μια νέα πιθανή θεραπεία για τον καρκίνο. Στα τέλη της δεκαετίας του 1990, μια ομάδα ερευνητών στην Βοστώνη που επικεφαλής της ήταν ο Δρ. *Judah Folkman* ανέφερε ότι ένα νέο φάρμακο με την ονομασία *Endostatin* συρρίκνωσε τους όγκους με το να περιορίζει τον ανεφοδιασμό τους με αίμα

[14]. Άλλοι ερευνητές δεν θα μπορούσαν να αναπαράγουν την επίδραση του φαρμάκου, ακόμα και αν χρησιμοποιούσαν τα φάρμακα που στάλθηκαν σε αυτούς από το εργαστήριο του Δρ *Folkman*. Οι επιστήμονες τελικά αναπαρήγαγαν τις επιδράσεις μετά το ταξίδι στο εργαστήριο του *Folkman*, όπου έμαθαν πως να κατασκευάζουν, να μεταφέρουν, να αποθηκεύουν και να χειρίζονται κατάλληλα το φάρμακο και πως να το χορηγούν στους ασθενείς με ένεση, στο κατάλληλο βάθος και γωνία. Το φάρμακο *Endostatin* είναι μια *inuis* συνθήκη. Ήταν ανεπαρκής αιτία από μόνη της, αλλά η αποτελεσματικότητά της ήταν απαραίτητη για να ενσωματωθεί σε ένα μεγαλύτερο σύνολο καταστάσεων, που δεν ήταν πλήρως κατανοητές από τους αρχικούς ερευνητές.

Οι περισσότερες αιτίες καλούνται ακριβέστερα *inuis* συνθήκες. Πολλοί παράγοντες απαιτούνται συνήθως προκειμένου να λάβει χώρα μια επίδραση, και σπάνια είμαστε σε θέση να τις γνωρίζουμε όλες αυτές και τον τρόπο που σχετίζονται μεταξύ τους. Αυτός είναι ένας λόγος που οι αιτιώδεις σχέσεις που συζητάμε δεν είναι ντετερμινιστικές αλλά μόνο αυξάνουν την πιθανότητα ότι μια επίδραση θα προκύψει.

Η επίδραση

Μπορούμε να κατανοήσουμε καλύτερα τι είναι μια επίδραση μέσω ενός *counterfactual* μοντέλου που βασίζεται στο φιλόσοφο του 18ου αιώνα *David Hume* [31, p.556]. Ένα *counterfactual* είναι κάτι που είναι αντίθετο προς ένα γεγονός. Σε ένα πείραμα, παρατηρούμε τι συνέβη όταν οι άνθρωποι λαμβάνουν μια θεραπεία. Το *counterfactual* είναι γνώση του τι θα μπορούσε να συμβεί σε αυτούς τους ίδιους ανθρώπους αν αυτοί ταυτόχρονα δεν είχαν λάβει αυτή την θεραπεία. Μια επίδραση είναι η διαφορά μεταξύ του τι συνέβη με αυτό που θα μπορούσε να συμβεί.

Δεν μπορούμε να παρατηρήσουμε στη πραγματικότητα το *counterfactual*. Εξετάστε την φαινυλκετονουρία (*PKU*), μια γενετικά μεταβολική ασθένεια που προκαλεί τη διανοητική καθυστέρηση, εκτός αν αντιμετωπιστεί κατά τη διάρκεια των πρώτων εβδομάδων της εμφάνισης της. Η *PKU* είναι η απουσία ενός ενζύμου που θα απέτρεπε τη συγκέντρωση της φαινυλαλανίνης, μιας τοξικής ουσία στο νευρικό σύστημα. Όταν μια αυστηρή δίαιτα φαινυλαλανίνης ξεκινήσει νωρίς και διατηρηθεί, επιβραδύνεται η πρόληψη. Σε αυτό το παράδειγμα, αιτία θα μπορούσε να θεωρηθεί μια ελλοχεύουσα γενετική ατέλεια, ως ενζυματική αναταραχή ή ως δίαιτας. Κάθε ένα συνεπάγεται ένα διαφορετικό *counterfactual*. Για παράδειγμα, αν πούμε ότι μια αυστηρή δίαιτα φαινυλαλανίνης προκάλεσε μια μείωση στην *PKU*, που βασίζεται στη πνευματική καθυστέρηση των νηπίων που έχουν αυξημένα επίπεδα φαι-

νυλκετονουρίας όταν γεννηθούν, τότε το *counterfactual* είναι οτιδήποτε θα είχε συμβεί αν αυτά τα ίδια νήπια δεν είχαν λάβει την ίδια αυστηρή διατροφή φαινυλαλανίνης. Η ίδια λογική ισχύει για τη γενετική ή ενζυματική μορφή της αιτίας. Όμως είναι αδύνατο για αυτά τα ίδια νήπια ταυτόχρονα να ακολουθούν μια συγκεκριμένη διαίτα και από την άλλη να μην την ακολουθούν, να έχουν και να μην έχουν τη γενετική διαταραχή, ή την ενζυμική ανεπάρκεια.

Έτσι ένας κεντρικός στόχος για όλη την έρευνα που προκαλεί την αιτία, είναι η δημιουργία λογικών προσεγγίσεων σε αυτό το αδύνατο φυσικά *counterfactual*. Για παράδειγμα, αν ήταν ηθικό θα μπορούσαμε να αντιπαραβάλουμε τα νήπια με φαινυλκετονουρία στα οποία δόθηκε η διατροφή με άλλα νήπια με φαινυλκετονουρία που δεν δόθηκε η διατροφή, αλλά που είχαν τα ίδια χαρακτηριστικά από πολλές απόψεις με εκείνα που ήταν παρόμοια σε γένος, ηλικία, κοινωνικοοικονομική θέση και κατάσταση υγείας. Ή θα μπορούσαμε (αν ήταν ηθικό) να αντιπαραθέσουμε τα νήπια που δεν ήταν σε διατροφή για τους πρώτους 3 μήνες της ζωής τους με αυτά τα νήπια που ξεκίνησαν την διατροφή μετά τον 4ο μήνα της ζωής τους. Καμία από αυτές τις προσεγγίσεις δεν αποτελεί αληθές *counterfactual*. Στην πρώτη περίπτωση, τα μεμονωμένα νήπια στην συνθήκη θεραπείας είναι διαφορετικά από αυτά της συνθήκης σύγκρισης, στην δεύτερη περίπτωση, οι ταυτότητες είναι ίδιες, αλλά ο χρόνος που πέρασε και ορισμένες αλλαγές ακόμα και αν η θεραπεία υποβλήθηκε στα νήπια (συμπεριλαμβανομένου της μόνιμης ζημιάς που έγινε από την φαινυλαλανίνη κατά τη διάρκεια των 3 μηνών ζωής). Έτσι οι δύο κεντρικοί στόχοι στο πειραματικό σχέδιο δημιούργησαν μια υψηλής ποιότητας αλλά απαραίτητως μια ημιτελής πηγή *counterfactual* συμπερασματολογίας και κατανόησης γιατί αυτή η πηγή διαφέρει από την συνθήκη επεξεργασίας.

Αυτός ο αντίθετος συλλογισμός είναι θεμελιωδώς ποιοτικός γιατί η αιτιώδη συμπερασματολογία, ακόμα και στα πειράματα, είναι θεμελιωδώς ποιοτική. Ωστόσο, μερικά από αυτά τα σημεία έχουν τυποποιηθεί από τους στατιστικολόγους σε μια ειδική περίπτωση που ορισμένες φορές ονομάζεται μοντέλο αιτιότητας του *Rubin* [23][44][46][47][51]. Θα αναφερθούμε στο μοντέλο του *Rubin* και στην συνέχεια του κεφαλαίου αυτού, αλλά αναλυτικά θα συζητηθεί στο κεφάλαιο 2.

Η αιτιώδης σχέση

Πώς γνωρίζουμε αν η αιτία με την επίδραση σχετίζονται; Σε μια κλασική ανάλυση που θεμελιώθηκε από τον φιλόσοφο του 19ου αιώνα *John Stuart Mill*, μια αιτιώδης σχέση υπάρχει εάν:

1. η αιτία προηγείται της επίδρασης

2. η αιτία σχετιζόταν με την επίδραση και
 3. δεν μπορούμε να βρούμε καμία εύλογη εναλλακτική εξήγηση για την επίδραση εκτός από την αιτία
- . Αυτά τα τρία χαρακτηριστικά αντανακλούν τι συμβαίνει στα πειράματα όπου:
1. χειριζόμαστε τη θεωρούμενη αιτία και παρατηρούμε μια έκβαση
 2. στη συνέχεια βλέπουμε εάν μια μεταβολή στην αιτία συσχετίζεται με τη μεταβολή στην επίδραση και
 3. χρησιμοποιούμε διάφορες μεθόδους κατά τη διάρκεια του πειράματος για να ελαττώσουμε την ευλογοφάνεια των άλλων εξηγήσεων για την επίδραση, μαζί με τις βοηθητικές μεθόδους για να ερευνήσουμε την ευλογοφάνεια εκείνων που δεν μπορούμε να αποκλείσουμε.

Ως εκτούτου τα πειράματα είναι κατάλληλα για τη μελέτη των αιτιώδων σχέσεων. Καμία άλλη επιστημονική μέθοδος τακτικά δεν ταιριάζει με τα χαρακτηριστικά των αιτιώδων σχέσεων τόσο καλά. Η ανάλυση του *Mill* δείχνει επίσης την αδυναμία άλλων μεθόδων. Σε πολλές μελέτες συσχέτισης, παραδείγματος χάριν, είναι αδύνατο να γνωρίζουμε ποιές από τις δύο μεταβλητές ήρθαν πρώτες, έτσι ισχυριζόμαστε ότι μια αιτιώδη σχέση μεταξύ τους είναι αβέβαιη. Κατανοώντας αυτή τη λογική των αιτιώδων σχέσεων και πως οι βασικοί όροι της, όπως είναι η αιτία και η επίδραση ορίζονται ώστε να βοηθήσουν τους ερευνητές να εξετάσουν τις μελέτες πρόκλησης αιτίας.

1.2 Αιτιότητα

1.2.1 Ιστορική Αναδρομή

Οι στατιστικολόγοι (επιστήμονες - μελετητές - ερευνητές) ενδιαφέρονται για την ερμηνεία των αναλύσεων τους. Αυτός ήταν και ένας λόγος που έδειχναν πάντα ιδιαίτερο ενδιαφέρον στην αιτιότητα, ακόμα και αν ήταν αρχικά φειδωλοί με την χρήση της.

Ο *Thus Yule* [66] ειδικότερα, τόνισε στο πλαίσιο μιας χρονοσειράς τη διάκριση μεταξύ της συσχέτισης και της αιτιότητας.

Ο *Fisher* [12],[13] έδειξε ότι η τυχαιοποίηση θα μπορούσε να αποδώσει αιτιώδη συμπερασματολογία για επιδράσεις θεραπειών, στις οποίες η αβεβαιότητα θα μπορούσε να αξιολογηθεί πιθανολογικά με βάση την τυχαιοποίηση χωρίς ειδικές υποθέσεις για τη δομή της μη ελέγξιμης μεταβλητότητας.

Ο *Cochran* [7] απέδωσε μια διεισδυτική εκδοχή από πολλές πτυχές της ανάλυσης μελετών παρατήρησης και παράλληλα επισήμανε την ανάγκη για επέκταση της ανάλυσης μονοπατιών του *Sewall – Wright* για την αντιμετώπιση θεμάτων με πιθανή αιτιότητα. Επιπρόσθετα ο *Cochran* παρέθεσε την απάντηση του *Fisher*, σε μια ερώτηση όπου ο *Cochran* τον ρωτούσε πως να κάνει τις μελέτες παρατήρησης πιο σαφείς - περιεκτικές, πιθανές ώστε να δίνουν απαντήσεις αιτιότητας. Η απάντηση ήταν : ‘Κάντε πιο αναλυτικές τις θεωρίες σας’. Αυτό μπορεί να επιτευχθεί με διάφορους τρόπους, για παράδειγμα, από το να συγκεντρώσουμε αποδείξεις διαφορετικών τύπων ή να εξασφαλίσει κάπως παρόμοιες αποδείξεις κάτω από ένα ευρύ φάσμα συνθηκών.

Ο *Hill* [22] έδωσε κατευθυντήριες γραμμές. Δεν αναφέρει ρητά τι εννοούσε με τον όρο αιτιότητα , παρόλα αυτά φαίνεται πολύ πιθανό ότι ήταν αυτό που αναφέρεται στη συνέχεια πρώτο-επίπεδο αιτιότητας. Ωστόσο οι κατευθυντήριες γραμμές του διατυπώνονται σε ένα επιδημιολογικό πλαίσιο. Τόνισε, ιδιαίτερα ότι πράγματι αποτελούν κατευθύνσεις και όχι κριτήρια.

Ο *Box* [4] τόνισε ότι πρέπει να δοθεί οπωσδήποτε προσοχή στην επίδραση μιας αιτιώδους ερμηνείας σε εξισώσεις παλινδρόμησης προσαρμοσμένα σε παρατηρήσιμα δεδομένα. Αν και η εξήγησή του τέθηκε πάνω σε πλαίσιο χημικής μηχανικής, το επιχείρημα ήταν και πάλι της ευρείας εφαρμογής.

Ο *Rubin* [44] σε μια σημαντική δημοσίευση του, προσάρμοσε έννοιες της αιτιότητας από σχεδιασμούς πειραμάτων (*Designs Of Experiments, DOE*) σε μελέτες παρατήρησης διαμέσου μιας αναπαράστασης παρόμοιας με αυτή του *Fisher*, χωρίς το ουσιώδες στοιχείο της φυσικής τυχαιοποίησης που είχε δοθεί από το *Neyman* [38]. Μεταγενέστερα, ο *Rubin* ανέπτυξε και εφάρμοσε αυτές τις ιδέες ιδιαίτερα στα πλαίσια κοινωνικών επιστημών. Η μελέτη του καθώς και πολλές άλλες προηγούμενες μελέτες προσεγγίζονται καλύτερα από μια δημοσίευση ανασκόπησης του *Holland* [23].

Οι *Cox & Snell* [8, p.84,85] σε μια στοιχειώδη έκθεση παλινδρόμησης, περιέγραψαν πέντε διαφορετικές ερμηνείες από εξισώσεις παλινδρόμησης και συντελεστές. Η μια ήταν για να εξετάσει την επίδραση από τις επιβαλλόμενες αλλαγές σε μια ή περισσότερες μεταβλητές και την απαιτούμενη προσοχή, ειδικά σε μελέτες παρατήρησης, προσδιορίζοντας οτιδήποτε κρίθηκε και δόθηκε σε αυτό έμφαση, σύμφωνα με τις επιβαλλόμενες αλλαγές.

Ο *Robins* σε μια μεγάλη σειρά δημοσιεύσεων για τις επιδράσεις, διερευνά έννοιες της αιτιότητας σε κλινικές δοκιμές και σε επιδημιολογικά σύνολα. Για προβλήματα, όπου οι θεραπείες ή οι παρεμβάσεις εφαρμόζονται διαδοχικά, μπορούμε να εμβαθύνουμε περισσότερο διευρύνοντας τη δημοσίευση των *Van der Laan* και *Robins* [63].

Το παραπάνω έργο θεωρείται πολύ σημαντικό στη στατιστική, με βασική ιδέα την ανάλυση παλινδρόμησης, η οποία λαμβάνει χώρα υπό την έννοια της μελέτης της εξάρτησης δυο ή περισσότερων επεξηγηματικών μεταβλητών. Το κεντρικό ζήτημα γενικότερα είναι:

- Η επιλογή μιας κατάλληλης μορφής σχέσης παλινδρόμησης
- Ο προσδιορισμός ποιων επεξηγηματικών μεταβλητών μπορούμε να συμπεριλάβουμε στη συσχέτιση, συμπληρωματικά με εκείνες που κατέχουν πιθανή αιτιώδη ερμηνεία
- Η εξέταση πιθανής αλληλεπίδρασης επιδράσεων, οι οποίες μπορεί να είναι βασικές για την ορθή ερμηνεία
- Ο συνδυασμός γεγονότων που προέρχονται από διαφορετικές μελέτες

Υπάρχουν περιπτώσεις όπου η αιτιότητα παρουσιάζεται ξεκάθαρα. Η επίδραση μπορεί να είναι μεγάλη και συνέπεια μεγάλης σύγχυσης του συστήματος. Ωστόσο επικεντρωνόμαστε περισσότερο σε καταστάσεις όπου η σύσταση της αιτιότητας είναι περισσότερο ευαίσθητη, είτε γιατί η επίδραση κάτω από μια μελέτη μπορεί να είναι μικρή είτε λόγω του πιθανού ανταγωνισμού των αιτιολογήσεων του μοντέλου. Υπάρχουν αρκετά παραδείγματα όπου η επιτυχής αναζήτηση για μια αιτιώδη επίδραση περιλαμβάνουν μια αλυσίδα διαφορετικών τύπων μελέτης. Μπορεί να ξεκινήσει από τη παρατήρηση αρκετών ασυνήθιστων γεγονότων, να συνεχιστεί με αναδρομικές και μελλοντικές μελέτες και στοιχεία από άλλες πηγές (μελέτες σε πειραματόζωα στα πλαίσια ανθρώπινης υγείας).

1.3 Ορισμοί και έννοιες αιτιότητας

Στην ενότητα αυτή παρουσιάζουμε τρεις διαφορετικές έννοιες – πτυχές της αιτιότητας. Είναι αρκετά σημαντικό να διακρίνουμε την αιτιότητα από την ιδιότητα που έχει σε διάφορους επιστημονικούς τομείς, για παράδειγμα στον τομέα της Φυσικής, της Βιολογίας, της Κοινωνιολογίας, από την ερμηνεία της στα στατιστικά μοντέλα. Η πρώτη πτυχή της αιτιότητας ονομάζεται **μηδενικό επίπεδο** αιτιότητας (*zero – level*), και χρησιμοποιείται συχνά στις στατιστικές μελέτες. Αυτή είναι μια στατιστική συσχέτιση, δηλαδή μη ανεξάρτητη, ξεκάθαρα

καθιερωμένη από την αιτία στην απόκριση, που δεν μπορεί να αφαιρεθεί θέτοντας συνθήκες με ‘επιτρεπτά’ εναλλακτικά χαρακτηριστικά. Το μηδενικό επίπεδο αιτιότητας είναι συνώνυμο της σταθερής συσχέτισης. Περιλαμβάνει κλασική αιτιότητα, ισχυρή εξάρτηση, *granger* αιτιότητα (μια πλευρά αιτιότητας ειδική για την ανάλυση χρονοσειρών) και αποτελέσματα από τις μετα – αναλύσεις (μια στατιστική μελέτη που συνδυάζει πληροφορίες από διάφορες άλλες μελέτες, για το προσδιορισμό μιας γενικής επίδρασης θεραπείας και είναι στατιστικά σημαντική).

Το μηδενικό επίπεδο αιτιότητας μελετήθηκε από τον *Good* [18][19] και συνολικά αναπτύχθηκε από τον *Suppes* [61], και σε μια γενική διατύπωση, στοχαστικής διαδικασίας από τον *Schweder* [53] και τον *Aalen* [1]. Στη συνέχεια, εισάγουμε μια διαφορετική πτυχή της αιτιότητας που ονομάζεται πρώτο επίπεδο αιτιότητας (*first level*). Αυτή η πτυχή, που είναι μια πιο γενική προσέγγιση της αιτιότητας, φαίνεται να είναι πιο άμεση σε σχετικές εφαρμογές που απασχολούν τους στατιστικολόγους. Αποτελεί συνώνυμο της αιτιότητας στο πλαίσιο δυο ενδεχόμενων αποτελεσμάτων και περιλαμβάνει την αιτιότητα ως επίδραση μιας παρέμβασης και ένα στοχαστικό μοντέλο για τις ιδιότητές της.

Για το σκοπό αυτό, αντιμετωπίζοντας δύο ή περισσότερες παρεμβάσεις σε ένα σύστημα, σκοπός είναι η σύγκριση των αποτελεσμάτων που θα προέκυπταν από διαφορετικές παρεμβάσεις. Για παράδειγμα, έστω δύο ιατρικές παρεμβάσεις τη C_1 και τη C_0 , μια νέα θεραπεία και μια θεραπεία ελέγχου, με μόνο μια από τις δυο να μπορεί να χρησιμοποιηθεί σε ένα συγκεκριμένο ασθενή. Στόχος είναι η σύγκριση του παρατηρήσιμου αποτελέσματος, αν έχει χρησιμοποιηθεί η C_1 θεραπεία, με το αποτέλεσμα που θα μπορούσε να παρατηρηθεί αν είχε χρησιμοποιηθεί η C_0 . Η ένδειξη μιας συστηματικής διαφοράς, θα ήταν μια απόδειξη, ότι η χρήση της θεραπείας C_1 αντί της χρήσης της C_0 , προκαλεί μια διαφορά στο αποτέλεσμα. Αυτή η αντίληψη μπορεί να θέτει σαν στόχο τον προβληματισμό για τη λήψη μιας απόφασης, αν και δεν είναι απαραίτητο. Για παράδειγμα, κατά την εξέταση εάν ένα ανώμαλο γονίδιο μπορεί να προκαλέσει κάποια πάθηση, η παρέμβαση μεταξύ μιας μη φυσιολογικής και μιας φυσιολογικής εκδοχής του γονιδίου (ομαλό και ανώμαλο) είναι υποθετική, και δεν λαμβάνεται τυπικά καμία άμεση λήψη απόφασης για αντιμετώπιση. Αυτή η ερμηνεία της αιτιότητας είναι σαφώς συγκριτική.

Μιά από τις λεπτές πτυχές αυτής της διατύπωσης είναι ότι μπορεί άμεσα να διατυπωθεί για κάθε άτομο, αλλά η επαλήθευσή της και συχνά η πραγματική ερμηνεία περιλαμβάνει συνολικά ή στατιστικά ζητήματα, δηλαδή συνεπάγονται μια επίδραση πάνω σε κάποιο σύνολο ατόμων. Σε αυτή την περίπτωση, ο ρητός καθορισμός του πληθυσμού των ατόμων αναφοράς μπορεί να είναι σημαντικός.

Τέλος παρουσιάζουμε την τρίτη πτυχή της αιτιότητας, το δεύτερο επίπεδο αιτιότητας (*second level*). Σε ένα επιστημονικό πλαίσιο, ως υποθέσουμε ότι ο προσεκτικός σχεδιασμός και η ανάλυση έχουν αποδείξει ένα πρότυπο από εξαρτήσεις ή συσχετίσεις ή έχουν εξασφαλιστεί λογικές αποδείξεις για το πρώτο ή το μηδενικό επίπεδο αιτιότητας. Το ερώτημα που προκύπτει είναι να εξηγήσουμε πως αυτές οι εξαρτήσεις και οι συσχετίσεις προέκυψαν, ή ποιά βαθύτερη παραγόμενη διαδικασία ενεπλάκη, ή ποια βαθύτερη δομή παρατηρείται. Αυτό συνεπάγεται σε συγχωνευμένη πληροφορία από πολλές διαφορετικές πηγές. Για παράδειγμα στα πλαίσια της Φυσικής, οι αποδεδειγμένες συνδέσεις με βασικές αρχές της Κλασικής ή της Κβαντομηχανικής και μεταξύ παρατηρήσεων σε επίπεδο εργαστηρίου. Μια μεθοδολογική διάκριση μεταξύ της επιδημιολογικής και κοινωνιολογικής έρευνας, είναι ότι στην πρώτη η πιθανή αιτιολογική επίδραση, από συγκεκριμένους παράγοντες κινδύνου, συχνά σχετίζεται ενδεχομένως με συστάσεις δημόσιας υγείας. Ενδιαφέρον παρουσιάζει το κοινωνιολογικό έργο που συχνά όλη η διαδικασία εξαπλώνεται συνδέοντας γονικά την κοινωνικοοικονομική τάξη με τα ατομικά χαρακτηριστικά καθημερινότητας. Το δεύτερο επίπεδο αιτιότητας χαρακτηρίζεται από τις πολυπαραγοντικές του εξαρτήσεις που επεξηγούν μια διαδικασία. Περιλαμβάνει γραφικά μοντέλα, μοντέλα συσχετίσεων λογαριθμικά γραμμικά μοντέλα και όλων των αιτιών δομικά μοντέλα εξισώσεων.

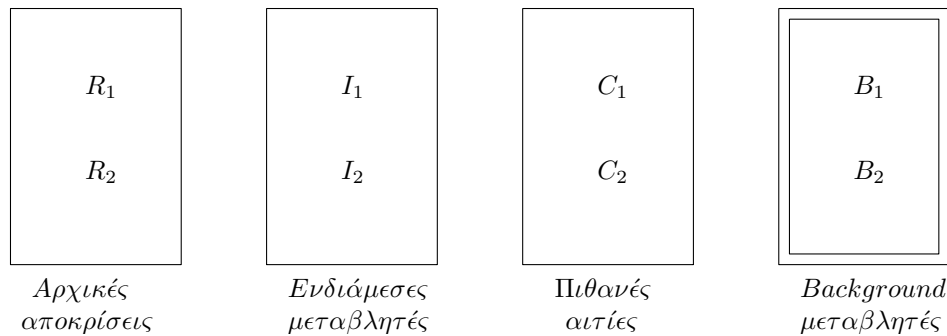
Η χρήση των παραπάνω όρων ποικίλλουν ουσιαστικά μεταξύ ατομικών και ομαδικών πεδίων. Παρόλα αυτά η αρχή στην οποία βασίζεται η διαδικασία ή οι διαδικασίες με χρήση αποδείξεων μοιάζει να ανταποκρίνεται ευρέως αλλά όχι με σε αποκλειστική χρήση στις Φυσικές Επιστήμες. Η πτυχή του πρώτου επιπέδου παρόλα αυτά, συνδέεται πιο συχνά με την Στατιστική, ειδικότερα σε πεδία Επιδημιολογίας με σύνδεση σε εφαρμοσμένο στόχο.

Για λόγους απλούστευσης όσων αφορά την ερμηνεία, είναι βολικό αρκετά συχνά να χρησιμοποιείται η ακόλουθη ορολογία. Έστω C μια υποψήφια αιτία, αν είναι λογικό να εξεταστεί η C ως πιθανή αιτία της R , για παράδειγμα υπό την έννοια του πρώτου επιπέδου αιτιότητας. Έστω ότι η C θεωρείται ενδεχόμενη αιτία, αν υπάρχει απόδειξη μιας αιτιώδους επίδρασης, για παράδειγμα οι θεωρητικές απαντήσεις σε εναλλακτικά επίπεδα, έστω για παράδειγμα οι C_1 και C_0 είναι συστηματικά διαφορετικές. Παραλείπουμε το χαρακτηριστικό ενδεχόμενο, όταν τα στοιχεία αποδεικνύουν ότι δεν υπάρχει εναλλακτική εξήγηση και ειδικότερα όταν η αναπτυσσόμενη διαδικασία είναι καλώς κατανοητή. Χρησιμοποιείται αυτή η προσεκτική προσέγγιση όχι για να αποθαρρυνθεί η έρευνα για αιτιότητα, αλλά για να αποκλειστεί η ενδεχόμενη πιθανότητα ότι οι πραγματικές συσχετίσεις μπορούν να θεωρηθούν αιτιώδεις, απλά και μόνο ονομάζοντάς τες έτσι.

Έχει επισημανθεί μια πιθανή σύνδεση των εννοιών εκ πρώτης όψης από τον *Suppes*

[61], για τις πραγματικές και τις ψευδές αιτίες. Η πρώτη από αυτές τις συνδέσεις αντιστοιχεί ευρέως στο τι ονομάζουμε πιθανές και ενδεχόμενες αιτίες. Ενώ η τρίτη από τις έννοιες του *Suppes* έχει να κάνει με μεταβλητές των οποίων η πιθανή αιτιώδη επίδραση εξηγείται διαμέσου επιτρεπτών μεταβλητών.

Οι τρεις πτυχές της αιτιότητας που είδαμε προηγουμένως, προβάλλονται σε ένα πλαίσιο πιθανοθεωρητικών μοντέλων. Για το λόγο αυτό αρκεί να θεωρήσουμε ένα σύστημα με τέσσερις μεταβλητές που υπολογίζονται για κάθε άτομο, η απόκριση R , η ενδιάμεση μεταβλητή I , η ενδεχόμενη αιτιώδη μεταβλητή C και μια *background* μεταβλητή B . Δίνεται το Σχήμα 1.1 με τις τέσσερις προαναφερθείσες μεταβλητές όπου θεωρούμε ότι κάθε μία από αυτές τις μεταβλητές έχει 2 συνιστώσες.



Σχήμα 1.1: Γραφική παράσταση τεσσάρων τύπων μεταβλητών. Στη στατιστική ανάλυση οι *background* μεταβλητές B συνήθως θεωρούνται δεσμευμένες για τια αντίστοιχες παρατηρήσιμες τιμές τους

Στη συνέχεια του κεφαλαίου αυτού θα αναπτύξουμε τα δυο βασικά *frameworks* της αιτιώδους συμπερασματολογίας, το *Neyman – Rubin* μοντέλο και τα δομικά μοντέλα εξισώσεων, που αφορούν τα δύο επίπεδα αιτιότητας.

1.4 Αιτιώδης συμπερασματολογία στη Στατιστική

1.4.1 Εισαγωγή

Τα προβλήματα που αφορούν την αιτιώδη συμπερασματολογία, επιμένουν να υπάρχουν στις βάσεις της Στατιστικής Επιστήμης, από τις απαρχές της. Η συσχέτιση δεν συνεπάγεται αιτιότητα, ακόμα και αν τα αιτιώδη συμπεράσματα εξάγονται από προσεκτικά σχεδιασμένα πειράματα. **Πως μπορούμε από ένα στατιστικό μοντέλο να ερμηνεύσουμε την αιτιότητα;** Η αντίδραση αρκετών στατιστικών ήταν αρνητική, όταν αντιμετώπισαν

τη πιθανότητα ότι ο τομέας τους θα συμβάλλει σε θέματα σχετικά με την αιτιότητα.

‘Αυτή η συσχέτιση δεν είναι αιτιότητα, και είναι ίσως το πρώτο πράγμα που πρέπει να ειπωθεί’.[3]

Πιθανότατα, αυτή ήταν μια πράξη διαφυγής, και αποτέλεσε απάντηση σε όλους όσους δήλωναν απρόσμενα το εξής:

‘Αν η Στατιστική δεν μπορεί να συνδέσει την αιτία με την επίδραση, τότε μπορεί να προστεθεί το πρόβλημα αυτό στον τομέα της Ρητορικής.’[56]

Προκειμένου να διαψεύσουμε τις παραπάνω δηλώσεις και να διερενήσουμε τις αιτιώδεις σχέσεις θα πρέπει να σχεδιάσουμε ορθά ένα τυχαίο πείραμα. Τα τυχαιοποιημένα πειράματα αρχικά χρησιμοποιήθηκαν στη γεωργία και στη συνέχεια επεκτάθηκαν και σε άλλα πεδία, διότι μπορούσαν να ελέγχουν τις εξωτερικές μεταβολές χωρίς την απαίτηση ενός εργαστηρίου. Το κύριο χαρακτηριστικό τους είναι ξεκάθαρο και πολύ σημαντικό: συγκρίνει τις διαφορετικές θεραπείες που ανατίθενται σε κάποιες πειραματικές μονάδες, με τυχαίο τρόπο, π.χ. με κλήρωση ή τη χρήση ενός πίνακα τυχαίων αριθμών. Η τυχαία ανάθεση παράγει δυο ή περισσότερα σύνολα τα οποία είναι πιθανολογικός παρόμοια μεταξύ τους. Οπότε, οποιαδήποτε διαφορά των αποτελεσμάτων που παρατηρείται σε αυτά τα σύνολα στο τέλος της μελέτης, οφείλεται στη θεραπεία, και όχι στις διαφορές μεταξύ των ομάδων που προϋπήρχαν από την αρχή του πειράματος. Επιπλέον όταν γίνονται συγκεκριμένες υποθέσεις, το τυχαιοποιημένο πείραμα αποδίδει μια εκτίμηση του μεγέθους της επίδρασης της θεραπείας σε σχέση τις εκτιμήσεις των πιθανοτήτων που η πραγματική επίδραση εμπίπτει σε ένα διάστημα εμπιστοσύνης.

1.4.2 Η βασική διάκριση και οι διακλαδώσεις της

Ο στόχος της στατιστικής ανάλυσης, που χαρακτηρίζεται από την παλινδρόμηση, την εκτίμηση και από τις υποθετικές τεχνικές ελέγχου, είναι να εκτιμήσει τις παραμέτρους μιας κατανομής από δείγματα που ελήφθησαν από αυτήν. Με τη βοήθεια αυτών των παραμέτρων μπορεί κανείς να συμπεράνει συσχετίσεις μεταξύ των μεταβλητών, να εκτιμήσει πιθανότητες παρελθοντικών ή μελλοντικών γεγονότων, καθώς και να προσαρμόσει αυτές τις πιθανότητες λαμβάνοντας υπόψη νέες αποδείξεις ή νέες μετρήσεις. Εφόσον οι συνθήκες του πειράματος παραμένουν ίδιες, οι παραπάνω δυνατότητες της στατιστικής ανάλυσης, μπορούν να διαχειριστούν. Στόχος της αιτιώδους ανάλυσης είναι να συμπεράνει όχι μόνο πεποιθήσεις ή

πιθανότητες κάτω από στατιστικές συνθήκες, αλλά και τη δυναμική των πεποιθήσεων κάτω από την αλλαγή συνθηκών, για παράδειγμα μεταβολές που προκαλούνται από αγωγές ή από εξωτερικές παρεμβάσεις.

Η διάκριση συνεπάγεται ότι οι αιτιώδεις έννοιες και οι έννοιες συσχέτισεων **δεν αναμειγνύονται**. Δεν υπάρχει κανένα στοιχείο σε μια από κοινού κατανομή ενός συμπτώματος και μιας αθένειας που μπορεί να μας εξασφαλίσει πως η θεραπεία του συμπτώματος θα θεραπεύσει η όχι την ασθένεια. Γενικότερα δεν υπάρχει κανένα στοιχείο σε μια συνάρτηση κατανομής που να μας βεβαιώνει πως αυτή η κατανομή θα μπορούσε να διαφέρει αν οι εξωτερικές συνθήκες έπρεπε να αλλάξουν. Για παράδειγμα από μια διαδικασία παρατήρησης σε μια πειραματική διάταξη – διότι οι νόμοι της θεωρίας πιθανοτήτων δεν υπαγορεύουν πως μια ιδιότητα μιας κατανομής θα πρέπει να αλλάζει όταν μια άλλη ιδιότητα έχει μεταβληθεί. Η πληροφορία αυτή θα πρέπει να παρέχεται από αιτιώδεις υποθέσεις που προσδιορίζουν τις σχέσεις που παραμένουν αμετάβλητες όταν οι εξωτερικές συνθήκες αλλάζουν.

Οι παραπάνω σχέψεις υπαινίσσονται την έκφραση *‘Η συσχέτιση δεν συνεπάγει την αιτιότητα’*. Η διατύπωση αυτή μπορεί να μεταφραστεί σε μια χρήσιμη αρχή: Κανείς δεν μπορεί να τεκμηριώσει αιτιώδεις αξιώσεις μόνο από συσχέτισεις, ακόμα και σε πληθυσμιακό επίπεδο. Πίσω από κάθε αιτιώδες συμπέρασμα πρέπει να βρίσκεται αιτιώδη υπόθεση που δεν είναι ελεγχόμενη από μελέτες παρατήρησης.

Η διαχωριστική γραμμή που διευκρινίζει τη διάκριση μεταξύ των εννοιών συσχέτισης και αιτιότητας, και εφαρμόζεται με απλό τρόπο, μπορεί να διατυπωθεί ως ακολούθως. Μια έννοια συσχέτισης είναι κάθε σχέση που μπορεί να οριστεί σε όρους μιας κοινής κατανομής παρατηρήσιμων μεταβλητών, και μια αιτιώδης έννοια είναι κάθε σχέση που δεν μπορεί να οριστεί μόνο από μια κατανομή. Οι έννοιες συσχέτισης περιλαμβάνουν μεθόδους όπως η συσχέτιση, η παλινδρόμηση, η εξάρτηση, η υπό συνθήκη ανεξαρτησία, η πιθανοφάνεια, τα αποτελέσματα ροπής, η αναλογία ρίσκου, η αναλογία πιθανοτήτων, η περιθωριοποίηση και οι ‘έλεγχοι για’ (*controlling for*). Οι έννοιες αιτιότητας περιλαμβάνουν την τυχαιοποίηση, την επιρροή, την επίδραση, τους συγχυτικούς παράγοντες, τη ψευδή συσχέτιση, την σταθερότητα, τις βοηθητικές μεταβλητές και τη παρέμβαση.

Η διαχωριστική αυτή γραμμή, είναι εξαιρετικά χρήσιμη στη διάκριση της αιτιώδους ανάλυσης γιατί βοηθά τους ερευνητές να εξιχνιάσουν τις υποθέσεις που χρειάζονται για να τεκμηριώσουν διάφορα είδη επιστημονικών αξιώσεων. Κάθε αξίωση που επικαλείται αιτιώδεις έννοιες πρέπει να βασίζεται σε προϋποθέσεις που επικαλούνται τέτοιες έννοιες. Δεν μπορεί να συνταχθεί από αυτό ή ακόμα να οριστεί μόνο με όρους στατιστικών συσχέτισεων.

Η αρχή που διατυπώσαμε, έχει σοβαρές συνέπειες που γενικά δεν είναι αναγνωρίσιμες από

την τυπική βιβλιογραφία της Στατιστικής. Μερικοί ερευνητές, για παράδειγμα, είναι ακόμα πεπεισμένοι ότι οι συγχυτικοί παράγοντες έχουν ανακαλυφθεί από την τυπική Στατιστική και αυτό μπορεί να φανεί από έναν ορισμό συσχετίσεων που λέει:

‘Ο U είναι ένας ενδεχόμενος συγχυτικός παράγοντας για την εξέταση της επίδρασης της θεραπείας X στο αποτέλεσμα Y , όταν οι X και U και αντίστοιχα οι U και Y δεν είναι ανεξάρτητοι’.

Είναι φανερό ότι, αυτός ο ορισμός και όλες οι εκδοχές του πρέπει να αποτυγχάνουν λόγω της διαχωριστικής γραμμής που έχει οριστεί. Εάν οι συγχυτικοί παράγοντες είναι προσδιορισμοί σε όρους των στατιστικών συσχετίσεων, θα ήμασταν ικανοί να τους προσδιορίσουμε από χαρακτηριστικά μη πειραματικών δεδομένων να προσαρμόσουμε εκείνους τους συγχυτικούς παράγοντες και να αποκτήσουμε αμερόληπτους εκτιμητές αιτιωδών επιδράσεων. Αυτό θα παραβίαζε το κανόνα: *‘Πίσω από κάθε αιτιώδες συμπέρασμα πρέπει να υπάρχει κάποια αιτιώδη υπόθεση που δεν έχει ελεγχθεί σε μελέτες παρατήρησης’*. Ως εκ τούτου ο ορισμός πρέπει να είναι ψευδής. Επιπλέον η απογοήτευση της γενιάς των ερευνητών της επιδημιολογίας και των κοινωνικών επιστημών όσον αφορά τη σύγχυση δεν μπορεί να διερευνηθεί ή να διορθωθεί από στατιστικές μεθόδους. Μόνο όταν κάποιος κάνει ορισμένες επικριτικές υποθέσεις σχετικά με τις αιτιώδεις σχέσεις, πριν μια προσαρμογή, μπορεί με ασφάλεια να διορθώσει τη μεροληψία σύγχυσης.

Στη συνέχεια αναλύονται και συγκρίνονται τα στατιστικά μοντέλα που είναι κατάλληλα για τη συμπερασματολογία συσχετίσεων και αιτιώδη συμπερασματολογία.

1.5 Το μοντέλο για τη συμπερασματολογία των συσχετίσεων

Το κατάλληλο μοντέλο για τη συμπερασματολογία συσχετίσεων είναι το κλασικό στατιστικό μοντέλο που σχετίζει δύο μεταβλητές πάνω σ’ ένα πληθυσμό. Για λόγους σαφήνειας και σύγκρισης με το μοντέλο της αιτιώδους συμπερασματολογίας, το οποίο περιγράφεται στη συνέχεια, θα εξετάσουμε εδώ την συσχέτιση εν συντομία.

Το μοντέλο ξεκινά με έναν πληθυσμό U από μονάδες. Μια μονάδα του U θα υποδηλώνεται με u . Οι μονάδες είναι τα βασικά αντικείμενα μελέτης σε μια έρευνα. Παραδείγματα μονάδων είναι τα ανθρώπινα θέματα ή θέματα που σχετίζονται με ανθρώπους, ο εργαστηριακός εξοπλισμός, τα νοικοκυριά, τα οικοπέδα. Μια μεταβλητή είναι απλά μια πραγματική

συνάρτηση που είναι καθορισμένη σε κάθε μονάδα της U . Η τιμή μιας μεταβλητής για μια δοθείσα μονάδα u είναι ο αριθμός που αποδίδεται από κάποια διαδικασία μέτρησης στο U . Ένα πλήθος από μονάδες και μεταβλητές που ορίζονται σε αυτές τις μονάδες είναι τα βασικά στοιχεία των μοντέλων για την σχέση και την αιτιότητα όπως παρουσιάζονται εδώ. Αντιστοιχούν στις μαθηματικές έννοιες ενός συνόλου και οι πραγματικές συναρτήσεις ορίζονται στα στοιχεία αυτού του συνόλου.

Ας υποθέσουμε ότι για κάθε μονάδα u στο U , υπάρχει μια τιμή $Y(u)$, που σχετίζεται. Ας υποθέσουμε ακόμα ότι η Y είναι μια μεταβλητή στατιστικού ενδιαφέροντος, υπό την έννοια ότι κάποιος επιθυμεί να καταλάβει γιατί οι τιμές της Y μεταβάλλονται με την πάροδο των μονάδων στη U . Η Y είναι η μεταβλητή απόκρισης εξαιτίας της κατάστασής της ‘μεταβλητή που πρέπει να εξηγηθεί’. Κατασκευάζοντας συμπερασματολογίες συσχετίσεων, κάποιος είναι ικανοποιημένος με την ανακάλυψη πως οι τιμές της Y ενώνονται με τις τιμές άλλων μεταβλητών που ορίζονται από τις μονάδες του U . Διακρίνουμε την A από την Y , με το να αποκαλούμε την A χαρακτηριστικό της μονάδας της u . Λογικά, ωστόσο, οι A και Y είναι επί ίσοις όροις, καθώς και οι δύο είναι απλά μεταβλητές που καθορίζονται στην U .

Όλες οι πιθανότητες, οι κατανομές και οι αναμενόμενες τιμές που εμπλέκονται με τις μεταβλητές υπολογίζονται πάνω στη U . Μια πιθανότητα δεν θα σημαίνει τίποτα λιγότερο ή περισσότερο από μια αναλογία από μονάδες στη U . Οι υπό συνθήκες αναμενόμενες τιμές είναι μέσα από υποσύνολα μονάδων όπου τα υποσύνολα ορίζονται από τη συνθηκολόγηση των τιμών και των μεταβλητών. Είναι υπό αυτή την έννοια τα μοντέλα που περιγράφονται εδώ από τα μοντέλα πληθυσμού.

Ο ρόλος του χρόνου χρειάζεται να τονιστεί εδώ. Οι πληθυσμοί των μονάδων υπάρχουν μέσα σε ένα χρονικό πλαίσιο κάποιου είδους και οι μετρήσεις των χαρακτηριστικών των μονάδων που οι μεταβλητές αναπαριστούν πρέπει επίσης να είναι κατασκευασμένες σε ειδικότερους χρόνους. Για την συμπερασματολογία συσχετίσεων ωστόσο, ο ρόλος του χρόνου είναι απλά για να επηρεάζει τον ορισμό του πληθυσμού των μονάδων ή να καθορίζει μια λειτουργική έννοια μιας ειδικότερης μεταβλητής. Όπως θα δούμε, στην αιτιώδη συμπερασματολογία ο ρόλος του χρόνου έχει μεγαλύτερη σημασία.

Η πιο λεπτομερής πληροφορία, που μπορεί να έχουμε στο μοντέλο που περιγράφεται είναι οι τιμές της $Y(u)$ και $A(u)$, καθώς και ότι όλες οι u ανήκουν στη U . Η από κοινού κατανομή της Y και της A πάνω στην U προσδιορίζεται από την $Pr(Y = y, A = a) =$ ποσοστό του u στη U , όπου $Y(u) = y$ και $A(u) = a$.

Οι παράμετροι συσχέτισης καθορίζονται από την κοινή τους κατανομή. Για παράδειγμα η δεσμευμένη κατανομή της Y δεδομένης της A προσδιορίζεται από την $Pr(Y = y|A =$

$a) = Pr(Y = y, A = a)/Pr(A = a)$. Αυτή η δεσμευμένη κατανομή περιγράφει πως η κατανομή των τιμών της Y , μεταβάλλεται πάνω στη U καθώς κυμαίνεται η A . Μια τυπική παράμετρος συσχέτισης είναι η παλινδρόμηση της Y στην A , έτσι ώστε η δεσμευμένη αναμονή $E(Y|A = a)$.

Η συμπερασματολογία συσχέτισεων αποτελείται από τη κατασκευή στατιστικών συμπερασμάτων (εκτιμήσεις, δοκιμές, *posterior* κατανομές) για παραμέτρους συσχέτισεων όπου η Y και η A σχετίζονται με τη βάση των δεδομένων που συγκεντρώνονται για την Y και την A από τις μονάδες του U . Με αυτή την έννοια, η συμπερασματολογία συσχέτισεων είναι απλά περιγραφική στατιστική.

1.6 Το μοντέλο του *Rubin* για την αιτιώδη συμπερασματολογία

Επειδή ο πειραματισμός είναι ένα ισχυρά επιστημονικό και στατιστικό εργαλείο και είναι αυτό που συχνά εισάγει τη σαφήνεια σε ειδικές περιπτώσεις αιτιότητας, θα επιστήσουμε τη προσοχή μας στη γλώσσα και το πλαίσιο των πειραμάτων για το μοντέλο για την αιτιώδη συμπερασματολογία. Όχι γιατί θεωρείται πως ένα πείραμα είναι ένα ορθό περιβάλλον για την τεκμηρίωση της αιτιότητας, αλλά σίγουρα είναι το πιο απλοποιημένο περιβάλλον.

Σκοπός μας είναι η κατασκευή ενός μοντέλου, που είναι περίπλοκο ώστε να επιτρέπει να τυποποιηθούν βασικές διαισθήσεις σχετικά με την αιτία και την επίδραση. Το σημείο εκκίνησης είναι η ανάλυση αιτιωδών επιδράσεων που έχουν δοθεί από τον *Rubin* [44],[47]. Θα είναι επαρκής, ωστόσο η ενασχόλησή μας με μια απλοποιημένη, πληθυσμιακού επιπέδου έκδοση του μοντέλου του *Rubin*. Αυτό το απλοποιημένο μοντέλο χρησιμοποιήθηκε από τον *Holland* και τον *Rubin* [25] για την αιτιώδη συμπερασματολογία σε εκ των υστέρων, σε ελεγχόμενες περιπτώσεις μελετών που χρησιμοποιούνται στην ιατρική έρευνα αλλά και για να αναλυθεί το παράδοξο της ανάλυσης συνδιακύμανσης του *Lord* [26]. Γίνεται αναφορά σ'αυτό ως 'το μοντέλο του *Rubin*' ακόμα και αν ο ίδιος υποστήριζε ότι οι ιδέες προϋπήρχαν από τον *Fisher*. Πιστεύεται ότι ο *Rubin* [44] ήταν το πρόσφορο έδαφος όπου αυτές οι ιδέες εφαρμόστηκαν για πρώτη φορά στη μελέτη της αιτιότητας.

Το μοντέλο αυτό ξεκινά επίσης με ένα πληθυσμό μονάδων U . Οι μονάδες στο μοντέλο για αιτιώδη συμπερασματολογία είναι αντικείμενα μελέτης στις οποίες οι αιτίες ή οι θεραπείες μπορούν να δράσουν. Οι όροι αιτία και θεραπεία θα εφαρμόζονται εναλλακτικά, και η έννοια που φέρουν είναι ένα σημαντικό μέρος του μοντέλου αυτού. Είναι σημαντικό να τονιστεί

ότι χρησιμοποιώντας τους όρους της αιτίας και της θεραπείας εναλλακτικά, δεν περιορίζεται η ανάπτυξη του θέματος, για τις δραστηριότητες εντός μιας ελεγχόμενης τυχαιοποιημένης μελέτης. Δίνεται έμφαση σε μια ιδέα που θεωρείται ότι λαμβάνει ανεπαρκή προσοχή σε γενικά ζητήματα αιτιότητας. Αυτό είναι το γεγονός ότι η επίδραση μιας αιτίας είναι πάντα σχετική με μια άλλη αιτία. Για παράδειγμα, η φράση ‘Η A προκαλεί τη B ’ σχεδόν πάντα σημαίνει ότι η A προκαλεί τη B , σχετικά με μια άλλη αιτία που περιλαμβάνει τη συνθήκη ‘όχι A ’. Η ορολογία γίνεται μάλλον κάπως βεβιασμένα αν προσπαθήσουμε να χρησιμοποιήσουμε μόνο τη συνήθη αιτιώδη γλώσσα, οπότε είναι πιο απλό αν χρησιμοποιήσουμε τη γλώσσα των πειραμάτων – θεραπεία (δηλαδή μια αιτία) έναντι ενός ελέγχου (δηλαδή μια άλλη αιτία).

Η βασική έννοια, ωστόσο είναι το ενδεχόμενο (ανεξάρτητα του τι μπορεί να επιτευχθεί με πρακτική εφαρμογή ή όχι) για την έκθεση ή όχι κάθε μονάδας στη δράση μιας αιτίας. Για την αιτιώδη συμπερασματολογία είναι κρίσιμο για κάθε μονάδα που είναι ενδεχομένως εκτεθειμένη σε κάθε μία από αυτές τις αιτίες. Για παράδειγμα, η εκπαίδευση που λαμβάνει ένας μαθητής μπορεί να αποτελεί μια αιτία, σύμφωνα με τη λογική μας, στη προετοιμασία του μαθητή για ένα τεστ, ενώ η εθνικότητα του μαθητή ή το φύλο του δεν είναι.

Για λόγους απλότητας, θα πρέπει να θεωρήσουμε ότι υπάρχουν μόνο δύο αιτίες ή επίπεδα θεραπείας, που συμβολίζονται με t (για τη θεραπεία) και με c (για τον έλεγχο). Έστω S μια μεταβλητή που συμβολίζει την αιτία στην οποία κάθε μονάδα εκτίθεται στη U έτσι ώστε $S = t$, και υποδηλώνει ότι η μονάδα είναι εκτεθειμένη στην t , και η $S = c$ υποδηλώνει έκθεση στη c . Σε μια ελεγχόμενη μελέτη, η S είναι κατασκευασμένη από τον πειραματιστή. Σε μια μη ελεγχόμενη μελέτη, η S προσδιορίζεται σε κάποιο βαθμό από τους παράγοντες πέρα από τον έλεγχο του πειραματιστή. Σε άλλη περίπτωση, το κρίσιμο χαρακτηριστικό της έννοιας της αιτίας σε αυτό το μοντέλο είναι ότι η τιμή της $S(u)$ για κάθε μονάδα θα μπορούσε να είναι διαφορετική.

Η μεταβλητή S είναι ανάλογη της μεταβλητής A του μοντέλου της συμπερασματολογίας συσχετίσεων. Η βασική διαφορά είναι ότι η $S(u)$ δηλώνει την έκθεση της u σε μια συγκεκριμένη αιτία, ενώ η $A(u)$ μπορεί να δηλώνει μια ιδιότητα ή ένα χαρακτηριστικό του u . Σε αυτή την περίπτωση η τιμή της $A(u)$ δεν θα μπορούσε να είναι διαφορετική.

Ο ρόλος του χρόνου τώρα γίνεται σημαντικός εξαιτίας του γεγονότος ότι όταν μια μονάδα είναι εκτεθειμένη σε μια αιτία αυτό πρέπει να συμβεί σε κάποιο ορισμένο χρόνο ή μέσα σε μια ορισμένη χρονική περίοδο. Οι μεταβλητές τώρα χωρίζονται σε δύο τάξεις: της προ-έκθεσης, αυτές των οποίων οι τιμές είναι καθορισμένες πριν την έκθεση στην αιτία, και της μετά-έκθεσης, αυτές των οποίων οι τιμές είναι καθορισμένες μετά την έκθεση στην αιτία.

Ο ρόλος μιας μεταβλητής απόκρισης Y είναι να μετρά την επίδραση μιας αιτίας και επιπλέ-

ον οι μεταβλητές απόκρισης πρέπει να ανήκουν στη μετά-έκθεση τάξη. Αυτό συνεπάγεται ένα άλλο κρίσιμο στοιχείο του μοντέλου. Οι τιμές της μετά-έκθεσης μεταβλητών είναι ενδεχομένως επηρεασμένες από μια ειδική αιτία, t ή c , στην οποία η μονάδα είναι εκτεθειμένη. Αυτό δεν είναι τίποτα λιγότερο από τη δήλωση ότι οι αιτίες έχουν επιδράσεις, που είναι η καρδιά της έννοιας της αιτιότητας. Για το μοντέλο που αναπαριστά πιστά αυτή την κατάσταση, χρειαζόμαστε όχι μόνο μια μεταβλητή Y , να αναπαριστά μια απόκριση, αλλά δύο μεταβλητές Y_t και Y_c , να αναπαριστούν δύο ενδεχόμενες αποκρίσεις. Η ερμηνεία αυτών των δύο τιμών, $Y_t(u)$ και $Y_c(u)$ για μια δοθείσα μονάδα u , είναι ότι η $Y_t(u)$ είναι η τιμή της απόκρισης που θα μπορούσε να παρατηρηθεί αν η μονάδα ήταν εκτεθειμένη στην t , και η $Y_c(u)$ είναι η τιμή που θα μπορούσε να παρατηρηθεί στην ίδια μονάδα αν ήταν εκτεθειμένη στην c .

Ο συμβολισμός $Y_t(u)$ και $Y_c(u)$ μερικές φορές, προκαλεί σύγχυση γιατί μια μεταβλητή συνήθως αναπαριστά μια μέτρηση κάποιου είδους και μια μέτρηση συνήθως θεωρείται ως αποτέλεσμα μιας διαδικασίας που εφαρμόζεται σε μια μονάδα. Αυτό δεν είναι πραγματικά ορθό. Για τις μετέπειτα - έκθεσης μεταβλητές ή μέτρηση εφαρμόζεται στο ζεύγος (u, t) (δηλαδή u μετά την έκθεση στο t) ή στο (u, c) (δηλαδή u μετά την έκθεση στην c). Ένας συμβολισμός που εκφράζει πιο κοντά αυτή την από κοινού εξάρτηση της Y στη u και την εκτεθειμένη αιτία είναι η $Y_t(u) = Y(u, t)$ και $Y_c(u) = Y(u, c)$. Ωστόσο θα χρησιμοποιήσουμε τον συμβολισμό Y_t και Y_c γιατί μας οδηγεί σε απλουστευμένες εκφράσεις.

Η επίδραση της αιτίας t στη u όπως μετράται από την Y και είναι σχετική με την αιτία c , είναι η διαφορά μεταξύ της $Y_t(u)$ και $Y_c(u)$. Στο μοντέλο, αυτό θα αναπαριστάται από την αλγεβρική διαφορά

$$Y_t(u) - Y_c(u) \tag{1.1}$$

Η διαφορά της εξίσωσης (1.1) ονομάζεται αιτιώδης επίδραση της t (σχετική με τη c) στη u (όπως μετρήθηκε από την Y). Η έκφραση (1.1) είναι ο τρόπος που το μοντέλο για αιτιώδη συμπερασματολογία εκκράζει τις πιο βασικές από όλες τις δηλώσεις για αιτιότητα (αιτιώδεις δηλώσεις). Εν ολίγοις η θεραπεία t προκαλεί την επίδραση $Y_t(u) - Y_c(u)$ στη μονάδα U (σχετική με τη θεραπεία της c) ή πιο απλά ότι

$$\eta \ t \text{ προκαλεί την επίδραση } Y_t(u) - Y_c(u) \tag{1.2}$$

Η αιτιώδης συμπερασματολογία τελικά ενδιαφέρεται για τις επιδράσεις των αιτιών σε συγκεκριμένες μονάδες, δηλαδή, με εξακρίβωση της τιμής της αιτιώδους επιδράσεως στην εξίσωση (1.2). Από εδώ προκύπτει και το Θεμελιώδες Πρόβλημα της Αιτιώδους Συμπερασματολογί-

ας.

1.6.1 Το Θεμελιώδες πρόβλημα της αιτιώδους συμπερασματολογίας

Είναι αδύνατο να παρατηρηθεί η τιμή της $Y_t(u)$ και της $Y_c(u)$ στην ίδια μονάδα και ως εκ τούτου, είναι απίθανο να παρατηρηθεί η επίδραση t στη u .

Η έμφαση στη λέξη παρατηρώ. Η αδυναμία της παρατήρησης και των δυο $Y_t(u)$, $Y_c(u)$, είναι από μόνη της εμφανής σε μερικά παραδείγματα και λιγότερο ξεκάθαρη σε άλλα. Για παράδειγμα, αν η μονάδα u είναι μια συγκεκριμένη τετάρτου βαθμού, η t αναπαριστά ένα νέο ετήσιο πρόγραμμα μελέτης της αριθμητικής, η c αναπαριστά ένα πρότυπο αριθμητικό πρόγραμμα και η Y είναι η βαθμολογία σε ένα διαγώνισμα στο τέλος του χρόνου. Τότε είναι προφανές ότι μπορεί να παρατηρηθεί είτε η $Y_t(u)$ είτε η $Y_c(u)$, αλλά όχι και οι δύο. Ποτέ δεν θα παρατηρηθεί η επίδραση της t ήταν στην u . Από την άλλη, αν η u είναι ένα δωμάτιο σε ένα σπίτι, η t δηλώνει ότι ανοίγω το διακόπτη του λαμπτήρα σε αυτό το δωμάτιο, η c ότι δεν τον ανοίγω και η Y αντιστοιχεί στο αν το φως είναι ανοιχτό ή όχι, για ένα μικρό χρονικό διάστημα μετά την εφαρμογή της t ή c , τότε τείνουμε να πιστεύουμε ότι είναι εφικτή η εύρεση των τιμών και των δύο $Y_t(u)$, $Y_c(u)$ με το απλό χτύπημα του διακόπτη. Αν για παράδειγμα το φως είχε ανοιγοκλείσει χωρίς κάποιο συγκεκριμένο λόγο ενώ σκεφτόμαστε την αρχή του πειράματος, είναι αμφίβολο πως θα μπορούσαμε να γνωρίσουμε τις τιμές της $Y_t(u)$ και $Y_c(u)$ μετά το άνοιγμα του διακόπτη.

Η έμμεση απειλή του Θεμελιώδους προβλήματος της αιτιώδους συμπερασματολογίας είναι ότι η αιτιώδης συμπερασματολογία είναι αδύνατη. Αλλά δεν θα πρέπει να καταλήξουμε σε αυτό το συμπέρασμα τόσο γρήγορα. Υποστηρίζοντας ότι η ταυτόχρονη παρατήρηση, της $Y_t(u)$ και $Y_c(u)$ είναι αδύνατη, δεν εννοούμε ότι οι γνώσεις μας που είναι σχετικές με αυτές τις τιμές, είναι εντελώς απύσυχες. Αυτό θα εξαρτηθεί από την εξεταζόμενη κατάσταση. Υπάρχουν δύο γενικές λύσεις στο Θεμελιώδες πρόβλημα, στις οποίες για λόγους ευκολίας, τις ονομάζουμε, επιστημονική λύση και την στατιστική λύση. Η επιστημονική λύση έγγειται στην εκμετάλλευση διάφορων ομοιογενών ή αναλλοίωτων υποθέσεων. Για παράδειγμα με την προσεκτική μελέτη της συμπεριφοράς ενός μέρους του εργαστηριακού εξοπλισμού, ένας επιστήμονας μπορεί να καταλήξει στο να πιστεύει ότι η τιμή της $Y_c(u)$ μετρημένη σε ένα προγενέστερο χρόνο είναι ίση με την τιμή της $Y_c(u)$ στο τρέχον πείραμα. Το μόνο που χρειάζεται να κάνει τώρα είναι να εκθέσει το u με το t και να μετρήσει την $Y_t(u)$ κι έτσι έχει ξεπεράσει το θεμελιώδες.

Η στατιστική λύση είναι διαφορετική και κάνει χρήση του πληθυσμού U με ένα τυπικό στατιστικό τρόπο. Η μέση αιτιώδης επίδραση, T , του t , (σχετική με τη c) πάνω στη U είναι η αναμενόμενη τιμή της διαφοράς $Y_t(u) - Y_c(u)$ πάνω στις μονάδες u στην U , δηλαδή:

$$E(Y_t - Y_c) = T \quad (1.3)$$

η T που ορίζεται στην (1.3), είναι η μέση αιτιώδης επίδραση. Από τις ιδιότητες που ισχύουν για τη μέση τιμή, η εξίσωση (1.3) μπορεί επίσης να εκφραστεί ως:

$$T = E(Y_t) - E(Y_c) \quad (1.4)$$

Αν και δεν είναι προφανές, η εξίσωση (1.4) αποκαλύπτει ότι η πληροφορία σε διαφορετικές μονάδες που μπορεί να παρατηρηθεί, μπορεί να χρησιμοποιηθεί για να αποκτηθεί η γνώση για την T . Για παράδειγμα, αν μερικές μονάδες εκτίθενται στην t , μπορούν να χρησιμοποιηθούν για να δώσουν πληροφορία για την $E(Y_t)$. Ο τύπος 1.4 χρησιμοποιείται για να αποκτήσουμε γνώση για την T . Ο ακριβής τρόπος που οι μονάδες θα επιλέγονται για έκθεση στο t ή στο c είναι πολύ σημαντικός και περιλαμβάνει όλες τις συνήθειες εκτιμήσεις για τους στατιστικούς σχεδιασμούς σε πειράματα. Το σημαντικό σημείο εδώ είναι ότι η στατιστική λύση αντικαθιστά την αδύνατον - να παρατηρηθεί αιτιώδη επίδραση της t σε μια συγκεκριμένη μονάδα με την πιθανή - να εκτιμηθεί αιτιώδη μέση επίδραση της t σε ένα πληθυσμό από μονάδες.

Η χρησιμότητα και της στατιστικής και της επιστημονικής λύσης στο Θεμελιώδες πρόβλημα της αιτιώδους συμπερασματολογίας εξαρτάται από την αλήθεια των διαφορετικών συνόλων υποθέσεων που δεν έχουν εξεταστεί.

Είναι χρήσιμο να υπάρχει ένας συμβολισμός που να εκφράζει το γεγονός ότι μια μεταβλητή του δείκτη αιτιότητας S καθορίζει ποια τιμή, Y_t ή Y_c , παρατηρείται για μια δεδομένη μονάδα. Αν ισχύει ότι $S(u) = t$, τότε η $Y_t(u)$ παρατηρείται, και αν $S(u) = c$ τότε η $Y_c(u)$ παρατηρείται. Επιπλέον η παρατηρήσιμη απάντηση στην μονάδα u είναι η $Y_{S(u)}(u)$. Η παρατηρήσιμη μεταβλητή απόκρισης, είναι ωστόσο η Y_S . Ως εκ τούτου, ακόμα και αν το μοντέλο περιέχει τρεις μεταβλητές, S, Y_t, Y_c , η διαδικασία της παρατήρησης εμπλέκει μόνο δύο, τις S και Y_S . Η διάκριση μεταξύ (α) της διαδικασίας μέτρησης της U , που παράγει την μεταβλητή απόκρισης και (β) των εκδοχών της μεταβλητής της απόκρισης Y_t και Y_c , που αντιστοιχεί σε ποια αιτία, η μονάδα είναι εκτεθειμένη (και από την άποψη ποιες αιτιώδεις επιδράσεις ορίζονται) και (γ) της μεταβλητής παρατηρήσιμης απάντησης Y_S είναι πολύ σημαντική και συχνά, δεν είναι 'κατασκευασμένη' από συζητήσεις αιτιότητας. Αυτές οι διακρίσεις ποτέ δεν

προκύπτουν από τη μελέτη μιας απλής συσχέτισης, αλλά είναι κρίσιμες στην ανάλυση μιας αιτιότητας.

Είναι χρήσιμο να επανεξεταστεί το μοντέλο για την συμπερασματολογία συσχετίσεων και το μοντέλο *Rubin*, δίπλα δίπλα για να δωθεί έμφαση στις διαφορές. Και οι δύο εμπλέκουν έναν πληθυσμό από μονάδες τον U , και οι δύο εμπλέκουν δύο παρατηρήσιμες μεταβλητές: (A, Y) για την συσχέτιση και (S, Y_S) για την αιτιότητα. Ωστόσο, αυτό είναι που έχουν και οι δύο πλευρές σε ομοιότητα. Ενώ η A και η U είναι απλώς μεταβλητές που ορίζονται στις μονάδες της U , S και Y_S , προϋποθέτουν μια πιο πολύπλοκη δομή προκειμένου να εφαρμοστούν σε πραγματικές καταστάσεις. Δύο ή περισσότερες αιτίες (ή θεραπείες) πρέπει να είναι ικανές για να εκτεθούν σε όλες τις μονάδες και η απάντηση Y πρέπει να είναι μια προ-έκθεσης μεταβλητή προκειμένου η παρατηρήσιμη απάντηση Y_S να ορισθεί. Η συμπερασματολογία συσχετίσεων εμπλέκει την ένωση ή τις δεσμευμένες κατανομές των τιμών της U και A , και η αιτιώδη συμπερασματολογία στρέφεται κυρίως προς τις τιμές $Y_t(u) - Y_c(u)$ σε μεμονωμένες μονάδες. Οι αιτιώδεις συμπερασματολογίες προέρχονται από τις παρατηρήσιμες τιμές των S και Y_S και από τις υποθέσεις που ταυτοποιούν ή δείχνουν το θεμελιώδες πρόβλημα της αιτιώδους συμπερασματολογίας αλλά συνήθως είναι μη ελεγχόμενες. Οι αιτιώδεις συμπερασματολογίες δεν εμπλέκονται κατά ανάγκη σε στατιστικές συμπερασματολογίες, αλλά οι συμπερασματολογίες συσχετίσεων πάντα το κάνουν.

1.7 Δομικά μοντέλα εξισώσεων

Τα δομικά μοντέλα εξισώσεων (*structural equations models, SEM*) είναι ένα από τα σημαντικότερα εργαλεία της αιτιώδους ανάλυσης στις Κοινωνικές επιστήμες και στις επιστήμες της Συμπεριφοράς. Ωστόσο η μεγαλύτερη ανάπτυξή τους οφείλεται στους επιστήμονες που ασχολούνται με τη Στατιστική. Στα *SEM* οι σχέσεις μεταξύ ενός συνόλου παρατηρήσιμων μεταβλητών μπορούν να εκφραστούν από μια γραμμική εξίσωση. Κάθε εξίσωση περιγράφει την εξάρτηση μιας μεταβλητής σε όρους μιας άλλης, και κάθε εξίσωση περιέχει ένα στοχαστικό όρο σφάλματος που αντιπροσωπεύει την επιρροή των μη παρατηρήσιμων όρων. Οι υποθέσεις αναξαρτησίας σε ζεύγη όρων σφαλμάτων επεξηγούνται επίσης από το μοντέλο.

Ένα χαρακτηριστικό των *SEM* που τα κάνει ακόμη πιο επιτυχή είναι η απλή τους αιτιώδη αναπαράσταση. Συγκεκριμένα η γραμμική εξίσωση $Y = \beta X + e$, κωδικοποιεί δυο ξεχωριστές υποθέσεις:

1. τη πιθανή ύπαρξη μιας (άμεσης) αιτιώδους επιρροής της μεταβλητής X στην Y

2. την απουσία μιας (άμεσης) αιτιώδους επιρροής στη Y από οποιαδήποτε άλλη μεταβλητή που δεν εμφανίζεται στο δεξί μέλος της εξίσωσης.

Η παράμετρος β ποσοτικοποιεί την (άμεση) αιτιώδη επίδραση της X στην Y . Δηλαδή η εξίσωση υποστηρίζει ότι μια αύξηση της μονάδας X έχει ως αποτέλεσμα β μονάδες αύξησης της Y , θεωρώντας πως οτιδήποτε άλλο παραμένει σταθερό.

Παραθέτουμε ένα παράδειγμα του Pearl [41]. Το μοντέλο αυτό ερευνά τη σχέση μεταξύ του καπνίσματος (X) και του καρκίνου των πνευμόνων (Y), λαμβάνοντας υπόψη το ποσοστό πίσσας (Z), στο πνεύμονα ενός ατόμου, και επιτρέποντας για τους μη παρατηρήσιμους παράγοντες να επηρεάζουν και το κάπνισμα (X) και το καρκίνο (Y). Η κατάσταση αυτή περιγράφεται με τις παρακάτω εξισώσεις.

$$X = \epsilon_1$$

$$Z = aX + \epsilon_2$$

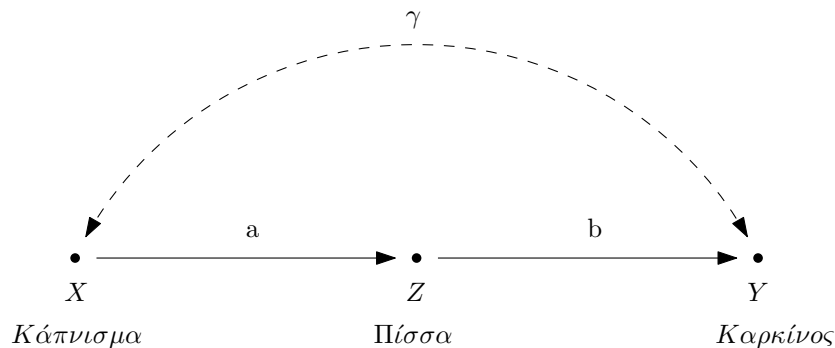
$$Y = bZ + \epsilon_3$$

$$Cov(\epsilon_1, \epsilon_2) = Cov(\epsilon_2, \epsilon_3)$$

$$Cov(\epsilon_1, \epsilon_3) = \gamma$$

Η πρώτη εξίσωση ισχυρίζεται ότι το επίπεδο του καπνίσματος ενός ατόμου εξαρτάται μόνο από παράγοντες που δεν περιλαμβάνονται στο μοντέλο, ενώ η δεύτερη ότι η ποσότητα πίσσας που έχει αποθηκευτεί στους πνεύμονες εξαρτάται από το επίπεδο του καπνίσματος καθώς επίσης και από εξωτερικούς παράγοντες. Τέλος το επίπεδο του καρκίνου εξαρτάται από τη ποσότητα της πίσσας στους πνεύμονες και εξωτερικούς παράγοντες. Οι υπόλοιπες εξισώσεις δηλώνουν – δείχνουν ότι οι εξωτερικοί παράγοντες που βοηθούν στην αποθήκευση της πίσσας στους πνεύμονες, είναι ανεξάρτητοι από τους εξωτερικούς παράγοντες που επηρεάζουν άλλες μεταβλητές. Όμως οι εξωτερικοί παράγοντες που έχουν επηρεάσει το κάπνισμα και το καρκίνο, ενδεχομένως να συσχετίζονται. Όλες οι παραπάνω πληροφορίες που περιέχονται στις εξισώσεις, μπορούν να εκφραστούν με μια αναπαράσταση που καλείται αιτιώδες γράφημα και παρουσιάζεται στο Σχήμα 1.2

Ένα δομικό μοντέλο εξισώσεων M για ένα διάνυσμα $Y = [Y_1, \dots, Y_n]'$ ορίζεται από ένα σύνολο γραμμικών εξισώσεων της μορφής $Y_j = \sum_i c_{ij} Y_i + e_j$ για $j = 1, \dots, n$ ή σε μορφή πίνακα $\mathbf{Y} = \mathbf{C}\mathbf{Y} + \epsilon$, όπου $C = [c_{ij}]$ και $\epsilon = [e_1, \dots, e_n]'$. Ο όρος e_j σε κάθε εξίσωση αντιστοιχεί σε ένα στοχαστικό σφάλμα που υποθέτεται ότι έχει κανονική κατανομή



Σχήμα 1.2: Παράδειγμα καπνίσματος και καρκίνου.

με μηδενική μέση τιμή. Το μοντέλο επίσης καθορίζει τις ανεξάρτητες υποθέσεις για εκείνους τους όρους σφαλμάτων, των η ένδειξη καταχώρησης στο πίνακα $\Psi = [\phi_{ij}] = Cov(e_i, e_j)$ έχει μηδενική τιμή. Οι δομικές εξισώσεις που κωδικοποιούνται στο μοντέλο M αποτελούνται από:

- ένα σύνολο μεταβλητών που παραλείπονται στο δεξί μέλος κάθε εξίσωσης (δηλαδή τις μηδενικές καταχωρήσεις του πίνακα), και
- τα ζεύγη των ανεξάρτητων όρων σφαλμάτων (δηλαδή τις καταχωρήσεις του πίνακα Ψ).

Η διαδικασία της ανάλυσης δεδομένων χρησιμοποιώντας τα SEM αποτελείται από τέσσερα βήματα [29]:

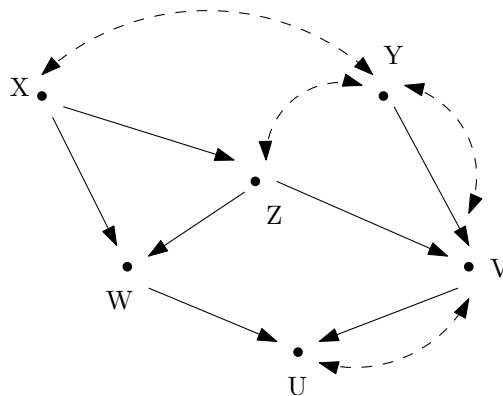
1. Προσδιορισμός: Περιγραφή δομής ενός μοντέλου. Δηλαδή οι ποσοτικές σχέσεις μεταξύ των μεταβλητών καθορίζονται με γραμμικές εξισώσεις. Ποσοτικές πληροφορίες δεν καθορίζονται γενικότερα και αναπαριστούνται από παραμέτρους.
2. Αναγνώριση: Ανάλυση προκειμένου να αποφασιστεί αν υπάρχει μοναδική εκτίμηση παραμέτρων που καθιστούν το μοντέλο συγκρίσιμο με τα παρατηρήσιμα δεδομένα. Η αναγνώριση ενός μοντέλου SEM .
3. Εκτίμηση: Πραγματική εκτίμηση των παραμέτρων από τις στατιστικές πληροφορίες στις παρατηρήσιμες μεταβλητές.
4. Εκτίμηση προσαρμογής: Αξιολόγηση της ποιότητας του μοντέλου ως περιγραφή των δεδομένων.

1.7.1 Βασικές γνώσεις γραφημάτων

Στο σημείο αυτό είναι σημαντικό να αναφέρουμε ότι ένα μονοπάτι μεταξύ των μεταβλητών X και Y σε ένα αιτιώδες γράφημα αποτελείται από μια ακολουθία ακμών $\langle e_1, e_2, \dots, e_n \rangle$ τέτοιο ώστε η ακμή e_1 να προσπίπτει στη X , η e_n να προσπίπτει στη Y και κάθε ζεύγος διαδοχικών ακμών στη σειρά να έχει κοινή μεταβλητή. Οι μεταβλητές X και Y καλούνται σημεία ακρότατα του μονοπατιού, και οποιαδήποτε άλλη μεταβλητή που εμφανίζεται σε κάποια ακμή e_i είναι μια ενδιάμεση μεταβλητή του μονοπατιού. Λέμε ότι το μονοπάτι δείχνει το ακρότατο σημείο X (Y) εάν η ακμή e_1 (e_n) έχει ένα βέλος που δείχνει στο X (Y).

Για παράδειγμα τα επόμενα μονοπάτια είναι ορισμένα μεταξύ των μεταβλητών X και U στο αιτιώδες γράφημα του Σχήματος 1.3.

- $X \rightarrow Z \rightarrow V \rightarrow U$
- $X \rightarrow Z \rightarrow V \leftrightarrow U$
- $X \leftrightarrow Y \leftrightarrow Z \rightarrow W \rightarrow U$
- $X \rightarrow Z \rightarrow V \leftrightarrow Y \leftrightarrow Z \rightarrow W \rightarrow U$



Σχήμα 1.3: Αιτιώδες γράφημα

Αξίζει να σημειωθεί ότι μόνο το τρίτο μονοπάτι δείχνει στη μεταβλητή X , άλλα όλα τα υπόλοιπα δείχνουν στη U .

Ένα μονοπάτι $p = \langle e_1, e_2, \dots, e_n \rangle$ μεταξύ των μεταβλητών X και Y , είναι έγκυρο αν η μεταβλητή X εμφανίζεται μόνο στην ακμή e_1 , η μεταβλητή Y μόνο εμφανίζεται στην ακμή e_n , και οποιαδήποτε ενδιάμεση μεταβλητή εμφανίζεται σε ακριβώς δυο μεταβλητές στο μονοπάτι.

Από τα παραπάνω παραδείγματα μόνο τα πρώτα τρία είναι έγκυρα. Το τελευταίο είναι μη έγκυρο γιατί η μεταβλητή Z εμφανίζεται σε παραπάνω από δυο ακμές.

Η ιδιαίτερη περίπτωση των μονοπατιών αποτελείται μόνο από άμεσες ακμές. Το πρώτο παράδειγμα ανταποκρίνεται σε μια αλυσίδα από τη X στη U .

Επίσης θα χρησιμοποιήσουμε οικογενειακούς όρους προκειμένου να αναφερθούμε σε συγκεκριμένες σχέσεις μεταβλητών. Συγκεκριμένα η ακμή $X \rightarrow Y$ παρουσιάζεται στο αιτιώδες διάγραμμα, οπότε λέμε ότι η μεταβλητή X είναι γονέας της Y . Παρόμοια, εάν υπάρχει μια αλυσίδα από την X στη Y , τότε η X θεωρείται πρόγονος της Y , και η Y απόγονος της X . Είναι ξεκάθαρο πως σε ένα αναδρομικό – επαναληπτικό μοντέλο [15], δεν μπορεί να υπάρξει μια κατάσταση όπου η X είναι ταυτόχρονα και πρόγονος και απόγονος της Y . Στο αιτιώδες γράφημα του Σχήματος 1.3, οι W και V είναι γονείς της μεταβλητής U , και η X είναι πρόγονος και των δυο U και V .

Δεδομένου ένα μονοπατιού p μεταξύ των μεταβλητών X και Y , και μια ενδιάμεση μεταβλητή Z στο p , συμβολίζουμε με $p[X..Z]$ το μονοπάτι που αποτελείται από τις ακμές του p που εμφανίζονται μεταξύ των X και Z . Η μεταβλητή Z είναι ένας κόμβος σύγκρουσης στο μονοπάτι p μεταξύ των X και Y , αν τα μονοπάτια $p[X..Z]$ και $p[Z..Y]$ δείχνουν στη Z . Ένα μονοπάτι το οποίο δεν περιέχει κανένα κόμβο σύγκρουσης, θεωρείται ότι είναι *unblocked*. Στη συνέχεια θεωρούμε κάποιες πραγματικότητες για τα *unblocked* μονοπάτια.

Το βάθος (*depth*) ενός κόμβου Y σε ένα αιτιώδες γράφημα ορίζεται ως το μήκος (δηλαδή τον αριθμό των ακμών) της μεγαλύτερης αλυσίδα από οποιοδήποτε πρόγονο του Y στο Y . Οι κόμβοι που δεν έχουν προγόνους έχουν βάθος 0.

Λήμμα 1. Έστω οι κόμβοι X και Y σε ένα αιτιώδες γράφημα ενός αναδρομικού μοντέλου για το οποίο ισχύει $depth(X) \geq depth(Y)$. Τότε οποιοδήποτε μονοπάτι μεταξύ των X και Y , το οποίο περιλαμβάνει ένα κόμβο Z για τον οποίο ισχύει $depth(Z) \geq depth(Y)$, πρέπει να διαθέτει ένα κόμβο σύγκρουσης.

Απόδειξη. Θεωρούμε ένα μονοπάτι p μεταξύ των X και Y , και τον κόμβο Z που ικανοποιεί τα κριτήρια του Λήμματος. Παρατηρούμε ότι το Z δε μπορεί να είναι πρόγονος ούτε του X ούτε του Y , διαφορετικά θα ίσχυε $depth(Z) < depth(X)$ ή $depth(Z) < depth(Y)$.

Ας θεωρήσουμε το τμήμα του μονοπατιού p μεταξύ των Z και Y . Εάν αυτό το τμήμα έχει τη μορφή $Z \rightarrow \dots Y$ τότε θα πρέπει να περιέχει ένα κόμβο σύγκρουσης εφόσον δεν είναι ένα κατευθυνόμενο μονοπάτι από το Z στο Y . Παρόμοια εάν το τμήμα μονοπατιού του p μεταξύ των X και Z είναι της μορφής $X \dots \leftarrow Z$, τότε θα πρέπει να περιέχει ένα κόμβο

σύγκρουσης.

Σε όλες τις υπόλοιπες περιπτώσεις ο Z είναι ένας μόνος που εμποδίζει το μονοπάτι. \square

Εάν ένα μονοπάτι p μεταξύ των X και Y , και ένα μονοπάτι q είναι ένα μονοπάτι μεταξύ των X και Z , τότε με το συμβολισμό $p \oplus q$ ορίζει το μονοπάτι που λαμβάνεται από το συνδυασμό των διαδοχικών ακμών αντίστοιχα στο p και q .

Λήμμα 2. Το μονοπάτι p θεωρείται ένα *unblocked* μονοπάτι μεταξύ των X και Y , και το q είναι ένα *unblocked* μονοπάτι μεταξύ των Y και Z . Τότε το μονοπάτι $p \oplus q$ είναι ένα έγκυρο *unblocked* μονοπάτι μεταξύ των X και Z αν και μόνο εάν:

1. τα p και q δεν έχουν καμία κοινή ενδιάμεση μεταβλητή
2. είτε το p αποτελεί μια αλυσίδα από το Y στο X , είτε το q αποτελεί μια αλυσίδα από το Y στο Z

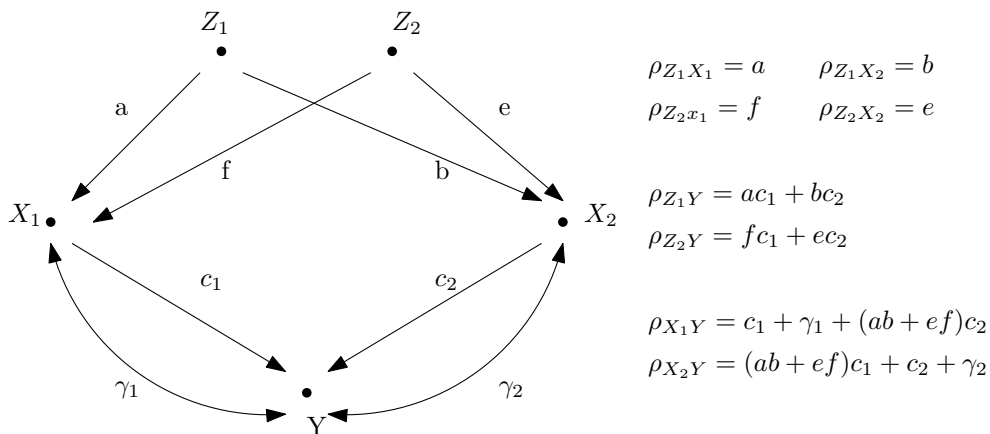
1.7.2 Η μέθοδος του *Wright* για την ανάλυση μονοπατιών

Η μέθοδος για της ανάλυσης μονοπατιών [65] για την αναγνώριση, βασίζεται στη παραγοντοποίηση των συσχετίσεων μεταξύ των παρατηρήσιμων μεταβλητών, σε πολώνυμα σχετικά με τις παραμέτρους του μοντέλου. Πιο συγκεκριμένα για τις μεταβλητές X και Y σε ένα αναδρομικό μοντέλο, ο συντελεστής μονοπατιού του X και Y , που συμβολίζεται με ρ_{XY} , μπορεί να εκφραστεί ως:

$$\rho_{XY} = \sum_{paths p_i} T(p_i) \quad (1.5)$$

όπου ο όρος $T(p_i)$ αντιπροσωπεύει το αποτέλεσμα των παραμέτρων των ακμών κατά μήκος του μονοπατιού p , και το άθροισμα κυμαίνεται έναντι όλων των *unblocked* μονοπατιών μεταξύ των X και Y . Για να ισχύει αυτή η ισότητα, οι μεταβλητές του μοντέλου πρέπει να είναι τυποποιημένες (δηλαδή διακύμανση ίση με 1) και να έχουν μηδενική μέση τιμή. Στη εξίσωση (1.5) αναφερόμαστε ως τη παραγοντοποίηση του *Wright* για τη διακύμανση ρ_{XY} . Το Σχήμα 1.4 παρουσιάζει ένα απλό μοντέλο και τις παραγοντοποιήσεις των συσχετίσεων για κάθε ζεύγος μεταβλητών.

Το σύνολο των μεταβλητών που λαμβάνεται από τη παραγοντοποίηση του *Wright* συνοψίζει όλες τις στατιστικές πληροφορίες που κωδικοποιούνται σε ένα μοντέλο. Οπότε



Σχήμα 1.4: Οι εξισώσεις του Wright

οποιαδήποτε ερώτηση σχετικά με την αναγνώριση, μπορεί να αποφασιστεί με τη μελέτη των λύσεων για αυτό το σύστημα εξισώσεων. Ωστόσο, από τη στιγμή που αυτό το σύστημα δεν είναι ένα γραμμικό σύστημα, είναι πολύ δύσκολο να αναλυθεί η αναγνώριση μεγάλων μοντέλων από την άμεση μελέτη των λύσεων για αυτές τις ερωτήσεις.

1.8 Σύνδεση του αιτιώδους μοντέλου του Rubin με τα δομικά μοντέλα εξισώσεων

Κατά την άποψη του Pearl [41] το RCM απαντάται στα δομικά μοντέλα εξισώσεων, που χρησιμοποιούνται στην Οικονομετρία και τις Κοινωνικές επιστήμες, με τη παραμετρική τους μορφή. Αυτή η άποψη που παρουσιάζεται επίσημα από τον Pearl [41] αμφισβητείται από τον Heckman [21]. Η βασική σύνδεση μεταξύ των RCM και τα SEM, στηρίζεται στην ερμηνεία της μεταβλητής του ενδεχόμενου αποτελέσματος $Y_x(u)$, να αποτελεί τη λύση της μεταβλητής Y σε ένα τροποποιημένο δομικό μοντέλο, στο οποίο η εξωτερική παρέμβαση $X = x$ προσομοιώνεται από την αντικατάσταση της εξίσωσης που προσδιορίζει τη X , με τη σταθερά $X = x$.

Η μεταβλητή u , η οποία στο RCM συμβολίζει το τύπο κάθε πειραματικής μονάδας, στα SEM αντιπροσωπεύεται με ένα διάνυσμα εξωγενών μεταβλητών (συνήθως μη παρατηρήσιμων) που χαρακτηρίζει κάθε μονάδα. Με αυτή την αναπαράσταση κάθε θεώρημα στο RCM μπορεί να αποδειχθεί ότι είναι ένα θεώρημα στα δομικά μοντέλα και αντίστροφα. Η ερμηνεία αυτή έχει οδηγήσει σε πλήρη axiomatization του RCM και στηρίζεται στα παράγωγα των

Shpitser – Pearl, μια πλήρη λύση για την αναγνώριση των αιτιωδών επιδράσεων χρησιμοποιώντας γραφήματα. Πλήρη λύση σημαίνει ότι για κάθε υποσύνολο X μεταβλητών και ένα σύμβολο A αιτιωδών υποθέσεων που κωδικοποιούνται σε ένα γράφημα G , είναι πιθανό να προσδιοριστεί αλγοριθμικά, εάν η αιτιώδης επίδραση $P(Y_x = y)$, μπορεί να εκτιμηθεί με συνέπεια – σταθερά από μη πειραματικά δεδομένα, και στη περίπτωση αυτή ποια μορφή πρέπει να έχει ο εκτιμητής (*estimand*) $P(Y_x = y)$. Χρησιμοποιώντας αυτή την εκτίμηση, είναι πιθανόν τότε από μια μελέτη παρατήρησης να εκτιμηθεί η μέση αιτιώδης επίδραση $E_u[Y_t(u) - Y_c(u)]$.

Από τη πλευρά του *Pearl* ένα σοβαρό μειονέκτημα του *RCM*, είναι ότι όλες οι υποθέσεις και οι βασικές γνώσεις που αναφέρονται σε ένα δεδομένο πρόβλημα πρέπει πρώτα να μεταφραστούν στη γλώσσα των *counterfactuals* (για παράδειγμα την αγνοησιμότητα), πριν ξεκινήσει – αρχίσει η ανάλυση. Στα *SEM* με σύγκριση ο *Pearl* [41] και ο *Heckman* [21] κρίνουν ότι οι βασικές γνώσεις εκφράζονται απευθείας σε ένα λεξιλόγιο, συνήθως επιστημονικού λόγου, επικαλούμενες την αιτία και την επίδραση μεταξύ υπαρκτών και όχι υποθετικών μεταβλητών. Το *RCM* επίσης μπορεί να συνδεθεί με τις βοηθητικές μεταβλητές [2] και άλλες τεχνικές για την αιτιώδη συμπερασματολογία.

Κεφάλαιο 2

Αιτιώδες μοντέλο Rubin

Το μοντέλο του *Neyman – Rubin* γίνεται ολοένα και περισσότερο δημοφιλές σε πολλά πεδία όπως η Στατιστική, η Ιατρική, οι Οικονομικές επιστήμες και η Κοινωνιολογία. Η δομή του στηρίζεται στο μη παραμετρικό μοντέλο του *Neyman* όπου κάθε μονάδα έχει δυο ενδεχόμενα αποτελέσματα, το ένα δίδεται από τη μονάδα που λαμβάνει θεραπεία και το άλλο από εκείνη που δεν λαμβάνει. Μια αιτιώδης επίδραση ορίζεται ως η διαφορά μεταξύ των δυο αποτελεσμάτων. Όμως μόνο ένα από τα δυο ενδεχόμενα αποτελέσματα μπορεί να παρατηρηθεί. Τα ενδεχόμενα αποτελέσματα προτάθηκαν για πρώτη φορά από τον *Jerzy Neyman* το [38] κατά τη διάρκεια απόκτησης του Μεταπτυχιακού του διπλώματος. Ο *Rubin* ανέπτυξε το μοντέλο μέσα σε ένα γενικότερο πλαίσιο για την αιτιώδη συμπερασματολογία, ενώ αργότερα ο *Holland* επινόησε την ονομασία ‘Αιτιώδες μοντέλο του *Rubin*’ και έγραψε άρθρο ισχυρής σημασίας τονίζοντας ορισμένες συνέπειες του μοντέλου.

2.1 Εισαγωγή

Στο κεφάλαιο αυτό παρουσιάζουμε ένα πλαίσιο για την αιτιώδη συμπερασματολογία, γνωστό ως ‘Αιτιώδες Μοντέλο του *Rubin*’ (*RCM, Holland, [23]*). Το πλαίσιο αποτελείται δύο βασικά μέρη. Το πρώτο μέρος του *RCM* ορίζει αιτιώδεις επιδράσεις μέσω των ενδεχόμενων αποτελεσμάτων. Το δεύτερο μέρος αυτού του πλαισίου για την αιτιώδη συμπερασματολογία αφορά τον μηχανισμό ανάθεσης. Περιγράφεται η κλασική χρήση του πλαισίου για τα τυχαιοποιημένα πειράματα, βάση τον *Neyman* [38] και τον *Fisher* [11], η οποία στη συνέχεια επεκτείνεται στις μη τυχαιοποιημένες μελέτες παρατήρησης. Επίσης εξετάζονται ορισμένα θέματα που περιλαμβάνουν, τη σπουδαιότητα της τυχαιοποίησης ως μη συγχυτική μέθοδο της ανάθεσης, μεθόδους για τη διαχείριση της μη συμμόρφωσης και των ελλεπών δεδομένων και μεθόδους για τον περιορισμό της μεροληψίας στην ανάλυση των παρατηρήσεων μελέτης.

2.1.1 Μονάδες, θεραπείες και ενδεχόμενα αποτελέσματα

Για την αιτιώδη συμπερασματολογία υπάρχουν διάφορες βασικές έννοιες στις οποίες πρέπει να στηριχθούμε. Μια μονάδα είναι ένα φυσικό αντικείμενο, για παράδειγμα ένα άτομο, μια συγκεκριμένη χρονική στιγμή. Μια θεραπεία είναι μια ενέργεια που μπορεί να εφαρμοστεί ή όχι σε αυτή τη μονάδα. Εστιάζουμε στην περίπτωση δυο θεραπειών, αν και η επέκταση σε περισσότερες από δύο θεραπείες είναι απλή αλλά όχι απαραίτητη όσον αφορά πραγματικά δεδομένα. Σε κάθε μονάδα αντιστοιχούν δύο ενδεχόμενα αποτελέσματα: η τιμή ενός αποτελέσματος μιας μεταβλητής Y σε μια χρονική στιγμή έπειτα από την εφαρμογή της ενεργούς θεραπείας και η τιμή του αποτελέσματος αυτής της μεταβλητής στην ίδια χρονική στιγμή όταν η ενεργή θεραπεία δεν εφαρμόζεται. **Στόχος μας είναι να μάθουμε για την αιτιώδη επίδραση της εφαρμογής της ενεργής θεραπείας σε σχέση με την θεραπεία ελέγχου (όταν η ενεργή θεραπεία δεν εφαρμόζεται) στην μεταβλητή Y .**

Για παράδειγμα, η μονάδα θα μπορούσε να είναι ένα άτομο αυτή τη στιγμή έχοντας πονοκέφαλο. Η ενεργή θεραπεία θα μπορούσε να είναι η λήψη της ασπιρίνης για τον πονοκέφαλο, και η θεραπεία ελέγχου θα μπορούσε να είναι η μη λήψη της ασπιρίνης. Το αποτέλεσμα Y θα μπορούσε να είναι η ένταση του πόνου του πονοκεφάλου του σε δύο ώρες, με δυο ενδεχόμενα αποτελέσματα να είναι η ένταση του πονοκεφάλου αν ληφθεί ασπιρίνη και αν δεν ληφθεί. Συμβολικά, έστω ότι η μεταβλητή W δηλώνει ποια θεραπεία έλαβε το άτομο: με $W = 1$ συμβολίζουμε την ενεργή θεραπεία, ενώ με $W = 0$ την θεραπεία ελέγχου. Επίσης έστω ότι $Y(1)$ είναι η τιμή του ενδεχόμενου αποτελέσματος αν η μονάδα έλαβε την ενεργή θεραπεία, και $Y(0)$ η τιμή αν η μονάδα έλαβε την θεραπεία ελέγχου. Παρουσιάζονται τρία παραδείγματα αιτιωδών επιδράσεων της ενεργής θεραπείας σε σχέση με την θεραπεία ελέγχου, που είναι η σύγκριση των τιμών $Y(0)$ και $Y(1)$. Στο πρώτο η απλή διαφορά $Y(1) - Y(0)$, ακολουθεί η μέθοδος *Gain Scores* όπου το αποτέλεσμα της μείωσης του πονοκεφάλου είναι $Y - X$, και τέλος παρουσιάζεται η ποσοστιαία μεταβολή, όπου το αποτέλεσμα της μείωσης του πονοκεφάλου είναι κλασματική, και δίνεται από το τύπο $Y^* = 1 - \frac{Y}{X}$ (στη περίπτωση αυτή ορίζουμε το κλάσμα $\frac{0}{0}$ να ισούται με 1).

Μπορούμε να παρατηρήσουμε μόνο μια τιμή της μεταβλητής $Y(W)$ όπως ορίστηκε από την W . Το κεντρικό πρόβλημα για την αιτιώδη συμπερασματολογία είναι ότι, για κάθε ανεξάρτητη μονάδα, είμαστε σε θέση να παρατηρήσουμε τη τιμή του ενδεχόμενου αποτελέσματος, βάση μόνο μιας εκ των δυο ενδεχόμενων θεραπειών. Δηλαδή τη θεραπεία την οποία πραγματικά ανατίθουμε, ενώ το ενδεχόμενο αποτέλεσμα βάση της άλλης θεραπείας

Μονάδα	Αρχικός Πονοκέφαλος	Ενδεχόμενα αποτελέσματα		Αιτιώδης επίδραση
	X	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
Άτομο	80	25	75	-50

Σχήμα 2.1: Ενδεχόμενα αποτελέσματα και αιτιώδης επιδράσεις: Απλή διαφορά

Μονάδα	Αρχικός Πονοκέφαλος	Ενδεχόμενα αποτελέσματα		Αιτιώδης επίδραση
	X	$Y(1) - X$	$Y(0) - X$	$Y(1) - X - [Y(0) - X]$
Άτομο	80	-55	-5	-50

Σχήμα 2.2: Ενδεχόμενα αποτελέσματα και αιτιώδης επιδράσεις: Gain Scores

Μονάδα	Αρχικός Πονοκέφαλος	Ενδεχόμενα αποτελέσματα		Αιτιώδης επίδραση
	X	$Y^*(1)$	$Y^*(0)$	$Y^*(1) - Y^*(0)$
Άτομο	80	$1 - \frac{25}{80} = 69\%$	$1 - \frac{75}{80} = 6\%$	$69\% - 6\% = 63\%$

Σχήμα 2.3: Ενδεχόμενα αποτελέσματα και αιτιώδης επιδράσεις: Ποσοστιαία μεταβολή

λείπει. Κατά συνέπεια, η συμπερασματολογία για τις αιτιώδεις επιδράσεις είναι ένα πρόβλημα ελλειπών δεδομένων. Για παράδειγμα, η μείωση της αρτηριακής πίεσης μια εβδομάδα μετά τη λήψη ενός φαρμάκου, μετράει μια αλλαγή στο χρόνο, ιδιαίτερα, μια αλλαγή πριν τη λήψη και μετά τη λήψη του φαρμάκου, και έτσι δεν αποτελεί μια αιτιώδη επίδραση χωρίς επιπρόσθετες υποθέσεις. Η σύγκριση της αρτηριακής πίεσης μετά την λήψη του φαρμάκου, με την αρτηριακή πίεση που θα μπορούσε να έχει το άτομο, στο ίδιο σημείο την ίδια χρονική στιγμή, με την μη λήψη φαρμάκου, αποτελεί μια αιτιώδη επίδραση.

2.1.2 Αιτιώδεις επιδράσεις: η αντιγραφή και η υπόθεση *SUTVA*

Με ποιον τρόπο μαθαίνουμε για τις αιτιώδεις επιδράσεις; Η απάντηση είναι η αντιγραφή περισσότερων μονάδων. Ο τρόπος με τον οποίο μαθαίνουμε από τη προσωπική μας εμπειρία είναι η αντιγραφή, η οποία εμπλέκει το ίδιο φυσικό αντικείμενο με περισσότερες μονάδες στο χρόνο. Δηλαδή, αν επιθυμούμε να μάθουμε για την επίδραση της λήψης της ασπιρίνης στον πονοκέφαλο μας, ο τρόπος είναι η επαναλήψεις στο χρόνο όταν λαμβάνουμε ή δεν λαμβάνουμε την ασπιρίνη, προκειμένου να ανακουφίσουμε τον πονοκέφαλο. Ως εκ τούτου λαμβάνουμε κάποιες από παρατηρήσεις $Y(1)$ και κάποιες $Y(0)$. Όταν θέλουμε να γενικεύσουμε σε μονάδες πέρα από τους ίδιους τους εαυτούς μας, τυπικά χρησιμοποιούμε περισσότερα αντικείμενα. Αυτό ακριβώς συμβαίνει στην επιδημιολογία και στα ιατρικά πειράματα.

Ας υποθέσουμε λοιπόν πως αντί για μια μονάδα έχουμε δύο. Τώρα, σε γενικές γραμμές, έχουμε το λιγότερο τέσσερα ενδεχόμενα αποτελέσματα για κάθε μονάδα: το αποτέλεσμα για τη μονάδα 1 αν και η μονάδα 1 και η μονάδα 2 λαμβάνουν τη θεραπεία ελέγχου, $Y_1(0, 0)$, το αποτέλεσμα για τη μονάδα 1 αν και οι δύο μονάδες λαμβάνουν την ενεργή θεραπεία, $Y_1(1, 1)$, το αποτέλεσμα για τη μονάδα 1 αν η μονάδα 1 λαμβάνει τη θεραπεία ελέγχου και η μονάδα 2 την ενεργή θεραπεία, $Y_1(0, 1)$, και το αποτέλεσμα για τη μονάδα 1 αν η μονάδα 1 λαμβάνει την ενεργή θεραπεία και η μονάδα 2 λαμβάνει την θεραπεία ελέγχου, $Y_1(1, 0)$. Ανάλογα αποτελέσματα λαμβάνουμε για τη μονάδα 2 με τις τιμές $Y_2(0, 0)$, $Y_2(1, 1)$, $Y_2(0, 1)$, $Y_2(1, 0)$. Επαναλαμβάνουμε το παράδειγμα με την ασπιρίνη, υποθέτοντας μόνο πως όλα τα δισκία είναι εξίσου αποτελεσματικά.

Θεραπεία που λαμβάνει το άτομο 1	1	0	1	0
Θεραπεία που λαμβάνει το άτομο 2	1	0	0	1
<u>Μονάδα</u>				
Άτομο 1	$Y_1(1, 1)=0$	$Y_1(0, 0)=100$	$Y_1(1, 0)=50$	$Y_1(0, 1)=75$
Άτομο 2	$Y_2(1, 1) = 0$	$Y_2(0, 0) = 100$	$Y_2(1, 0) = 100$	$Y_2(0, 1) = 0$

Σχήμα 2.4: Ενδεχόμενα αποτελέσματα με δυο μονάδες όπου επιτρέπεται η συσχέτιση μεταξύ τους (δεν ισχύει το (β) μέρος της υπόθεσης *SUTVA*)

Η αιτιώδης επίδραση του να λάβει το άτομο 2 την ασπιρίνη με το να μην την λάβει σε κάθε περίπτωση ανεξάρτητα από το άτομο 1, είναι 100. Θα μπορούσαμε να πούμε ότι η αιτιώδης επίδραση για το άτομο 2 είναι καλώς ορισμένη. Ο λόγος είναι ότι η διαφορά $Y_2(1, 1) - Y_2(0, 1)$,

η οποία είναι η αιτιώδης επίδραση του ατόμου 2 να λάβει την ασπιρίνη με το γεγονός να μη τη λάβει, όταν το άτομο 1 λαμβάνει την ασπιρίνη, είναι $0 - 100 = -100$. Και η διαφορά $Y_2(0, 1) - Y_2(0, 0)$, η οποία είναι η αιτιώδης επίδραση του ατόμου 2 να λάβει την ασπιρίνη με το γεγονός να μη τη λάβει, όταν το άτομο 1 δεν λαμβάνει την ασπιρίνη, είναι $0 - 100 = -100$. Οπότε το αποτέλεσμα για το άτομο 2 δεν εξαρτάται από το τι λαμβάνει το άτομο 1. Αντιθέτως, για το άτομο 1 η αιτιώδης επίδραση να λάβει την ασπιρίνη με το γεγονός να μη τη λάβει, εξαρτάται με το τι λαμβάνει το άτομο 2. Αν το άτομο 2 λάβει ασπιρίνη, η αιτιώδης επίδραση για το άτομο 1 είναι $Y_1(1, 1) - Y_1(0, 1) = 0 - 75 = -75$. Ενώ αν το άτομο 2 δεν λάβει ασπιρίνη, η αιτιώδης επίδραση για το άτομο 1 είναι $Y_1(1, 0) - Y_1(0, 0) = 50 - 100 = -50$, μια μικρότερη επίδραση. Το γεγονός ότι η αιτιώδης επίδραση του ατόμου 1 εξαρτάται από το τι θεραπεία λαμβάνει το άτομο 2, καθιστά την ανάλυση της κατάστασης περίπλοκη. Για το λόγο αυτό γίνεται η ‘υπόθεση σταθερής τιμής μονάδας θεραπείας’ (*SUTVA*), για την οποία προσπαθούμε σκληρά να κατασκευάσουμε καταστάσεις κάτω από τις οποίες ισχύει.

Η υπόθεση (*SUTVA*) αποτελείται από δυο μέρη, (α) υπάρχει μόνο μια μορφή ενεργής θεραπείας και μια μορφή θεραπείας ελέγχου, (β) δεν υπάρχει συσχέτιση μεταξύ των μονάδων. Επιστρέφοντας στο παράδειγμά μας, αξίζει να σημειωθεί πως δεν έχω θεωρήσει τη πιθανότητα να υπάρχουν περισσότερο και λιγότερο αποτελεσματικά δισκία ασπιρίνης. Αν υπήρχε μια τέτοια περίπτωση, στους προηγούμενους συμβολισμούς θα πρέπει να συμπεριλάβουμε και τους ‘1+’, για την περισσότερο αποτελεσματική ασπιρίνη, και ‘1-’ για τη λιγότερο αποτελεσματική. Ας σκεφτούμε τώρα τη περίπτωση που περιλαμβάνει ένα ολόκληρο κουτί με δισκία ασπιρίνης, καθένα από τα οποία ποικίλει σε αποτελεσματικότητα. Τότε η κατάσταση που θα έπρεπε να μελετήσουμε θα γινόταν εκθετικά πιο περίπλοκη, ακόμη και στη περίπτωση δυο μόνο μονάδων. Με περισσότερες από δυο μονάδες, η υπόθεση (*SUTVA*) είναι ακόμη πιο σημαντική και εξηγεί γιατί χρησιμοποιείται. Το σχήμα 2.5 με τιμές $X, Y(1)$ και $Y(0)$, αντιπροσωπεύει **επιστημονικά**, τι είναι ακριβώς αυτό που θέλουμε να μάθουμε.

Είναι προφανές, πως η *SUTVA* είναι μια πολύ βασική υπόθεση. Όμως δεν υπάρχει αιτιώδης συμπερασματολογία απαλλαγμένη από υποθέσεις. Είναι η ποιότητα των υποθέσεων που έχει σημασία, και όχι η ύπαρξή τους ή ακόμα και η απόλυτη ακρίβειά τους. Αρκετοί ερευνητές τείνουν να κάνουν κατάλληλες υποθέσεις από τον σχεδιασμό των μελετών τους. Για παράδειγμα, η *SUTVA*, γίνεται πιο κατάλληλη όταν οι μονάδες είναι απομονωμένες από τις υπόλοιπες για παράδειγμα σε ένα σχολείο. Κατά τη μελέτη μιας διαδικασίας όπως ένα πρόγραμμα πρόληψης καπνίσματος, προσδιορίζουμε τις μονάδες που είναι ανέπαφες με το κάπνισμα αντί των μεμονωμένων σπουδαστών ή των τάξεων στα σχολεία.

Η σταθερότητα της υπόθεσης (*SUTVA*) χρησιμοποιείται πολύ συχνά, ακόμα και αν δεν

Μονάδα	X	$Y(1)$	$Y(0)$	Αιτιώδης επίδραση σε μοναδιαίο επίπεδο
1	X_1	$Y_1(1)$	$Y_1(0)$	$Y_1(1) - Y_1(0)$
2	X_2	$Y_2(1)$	$Y_2(0)$	$Y_2(1) - Y_2(0)$
\vdots	\vdots	\vdots	\vdots	\vdots
i	X_i	$Y_i(1)$	$Y_i(0)$	$Y_i(1) - Y_i(0)$
\vdots	\vdots	\vdots	\vdots	\vdots
N	X_N	$Y_N(1)$	$Y_N(0)$	$Y_N(1) - Y_N(0)$

Σχήμα 2.5: Ενδεχόμενα αποτελέσματα για το παράδειγμα της ασπιρίνης, για N μονάδες υπό την υπόθεση $SUTVA$)

είναι πάντα κατάλληλη. Για παράδειγμα, θεωρήστε μια μελέτη της επίδρασης ενός εμβολιασμού για μια μεταδοτική ασθένεια. Όσο μεγαλύτερο είναι το ποσοστό του πληθυσμού που εμβολιάζεται, τόσο μικρότερη είναι η πιθανότητα κάθε μονάδας να προσβληφθεί από την ασθένεια, ακόμη και αν δεν έχει εμβολιαστεί-ένα παράδειγμα παρεμβολής. Σε ολόκληρο το κεφάλαιο, εφαρμόζουμε την υποθέση της $SUTVA$, αν και υπάρχουν και άλλες υποθέσεις που θα μπορούσαν να περιορίσουν τον αριθμό ενδεχομένων αποτελεσμάτων με αντιγραφή και χωρίς τις υποθέσεις.

Γενικότερα, μερικές από τις N μονάδες ενδέχεται να μην λαμβάνουν ούτε την ενεργή θεραπεία $W_i = 1$ αλλά ούτε και την θεραπεία ελέγχου $W_i = 0$. Λόγω του ότι ορισμένες από τις μονάδες μπορεί να υπάρξουν στο μέλλον, όπως όταν θέλουμε να γενικεύσουμε έναν πληθυσμό. Τότε επίσημα η W_i πρέπει να λάβει μια τρίτη τιμή, $W_i = *$ χωρίς να εκπροσωπεί ούτε το 1 αλλά ούτε το 0.

Παρουσιάζουμε ορισμένες αιτιώδεις επιδράσεις σε πληθυσμιακό επίπεδο

- Η σύγκριση των $Y_i(1)$ και $Y_i(0)$ σε ένα συνηθισμένο υποσύνολο μονάδων
- Ο μέσος όρος της αιτιώδους επίδραση της ‘1’ vs ‘0’ = $Ave[Y_i(1) - Y_i(0)] = \frac{1}{N} \sum_{i=1}^N [Y_i(1) - Y_i(0)] \equiv ACE$. Μια από τις πολλές χρήσιμες ιδιότητες της επίδρασης αυτής είναι η διάσπαση των αναμενόμενων τιμών που μας επιτρέπει να θεωρήσουμε τους μέσους όρους των ενδεχομένων αποτελεσμάτων ξεχωριστά, $ACE \equiv E[Y_i(1)] - E[Y_i(0)]$.
- Η μέση αιτιώδης επίδραση της ‘1’ vs ‘0’ = $Median\{Y_i(1) - Y_i(0)\}$
- Η διαφορά των μέσων των ενδεχομένων αποτελεσμάτων = $Median\{Y_i(1)\} - Median\{Y_i(0)\}$

- Αν η μεταβλητή X_i περιλαμβάνει άρρεν/θύλη για κάθε μονάδα τότε: Μέσος όρος της αιτιώδους επίδρασης της ‘1’ vs ‘0’ για τους άντρες= $Ave_{X_i=male}\{Y_i(1) - Y_i(0)\}$

Δυστυχώς οι μέσοι όροι $E[Y_i(1)]$, $E[Y_i(0)]$ δεν είναι πάντοτε παρατηρήσιμοι. Για το λόγο αυτό θα παρατηρήσουμε τους μέσους όρους των ενδεχόμενων αποτελεσμάτων στους υποπληθυσμούς που στη πραγματικότητα έλαβαν την ενεργή θεραπεία και τη θεραπεία ελέγχου. Ο Holland [23] καλεί αυτή την επίδραση *prima facie* μέσο όρο αιτιώδους επίδρασης:

$$ACE^{pf} \equiv E[Y_i(1)|X = 1] - E[Y_i(0)|X = 0]$$

2.1.3 Συμμεταβλητές

Οι συμμεταβλητές (όπως η ηλικία, η φυλή και το φύλο) κατέχουν σημαντικό ρόλο στις μελέτες παρατήρησης για τις αιτιώδεις επιδράσεις. Στην Επιδημιολογία και στη Στατιστική είναι γνωστές ως πιθανοί συγχυτικοί παράγοντες ή παράγοντες κινδύνου. Σε ορισμένες μελέτες, οι μονάδες που εκτίθενται στην ενεργή θεραπεία διαφέρουν ως προς την κατανομή των συμμεταβλητών, σε σημαντικό βαθμό από τις μονάδες που δεν εκτίθενται.

$$(X, Y(0), Y(1)) = \begin{bmatrix} X_1 & Y_1(0) & Y_1(1) \\ \vdots & \vdots & \vdots \\ X_i & Y_i(0) & Y_i(1) \\ \vdots & \vdots & \vdots \\ X_N & Y_N(0) & Y_N(1) \end{bmatrix}$$

Η πρώτη στήλη του παρακάτω πίνακα δίνει τις συμμεταβλητές $X = \{X_i\}$, που είναι, εξ'ορισμού, ανεπηρέαστες από την θεραπεία, όπως και η αρτηριακή πίεση πριν τη θεραπεία, η δεύτερη στήλη τα ενδεχόμενα αποτελέσματα, $Y(0) = \{Y_i(0)\}$, όταν δεν εκτίθενται οι μονάδες στην θεραπεία, και η τρίτη τα ενδεχόμενα αποτελέσματα, $Y(1) = \{Y_i(1)\}$, όταν εκτίθενται στην θεραπεία. Με $W_i = \{W_i\}$ συμβολίζουμε το δείκτη διανύσματος της θεραπείας για κάθε μονάδα στη μελέτη. Όλες οι αιτιώδεις εκτιμήσεις περιλαμβάνουν συγκρίσεις της $Y_i(0)$ και της $Y_i(1)$, είτε για όλες τις N μονάδες, ή σε ένα κοινό υποσύνολο μονάδων.

Οπότε, βάση της υπόθεσης *SUTVA*, όλες οι αιτιώδεις εκτιμήσεις μπορούν να υπολογιστούν από τον παραπάνω πίνακα με την i -οστή γραμμή : $(X_i, Y_i(0), Y_i(1))$. Εξ ορισμού, όλη οι σχετικές πληροφορίες κωδικοποιούνται στις $X_i, Y_i(0), Y_i(1)$, με αποτέλεσμα ο πίνακας να κατασκευάζεται κατα γραμμή με τυχαία μετάθεση των $1, \dots, N$. Με άλλα λόγια, ο πίνακας

N γραμμών είναι είναι ανταλλάξιμος κατά σειρές. Καλούμε αυτόν τον πίνακα ‘Επιστήμη’ διότι οι τιμές του δεν μπορούν να ελεγχθούν. Με την αλλαγή των θεραπειών, υπάρχει η πιθανότητα να διαπιστώσουμε ποιες τιμές παραγματικά παρατηρήθηκαν, αλλά όχι τις κάθε αυτού τιμές τους. Δηλαδή, οι παρατηρήσιμες τιμές της Y είναι οι $Y_{obs} = \{Y_{obs,i}\}$, όπου οι $Y_{obs,i} = Y_i(1)W_i + Y_i(0)(1 - W_i)$.

Προκειμένου να διαπιστώσουμε πως επιδρά το ζήτημα των συμμεταβλητών στο πλαίσιο του αιτιωδούς μοντέλου, πρέπει να ορίσουμε τον ‘μηχανισμό ανάθεσης’, έναν πιθανολογικό μηχανισμό που προσδιορίζει ποιες μονάδες λαμβάνουν την ενεργή θεραπεία και ποιες μονάδες λαμβάνουν τη θεραπεία ελέγχου.

2.1.4 Ιστορική αναφορά στα ενδεχόμενα αποτελέσματα

Η βασική ιδέα για τις αιτιώδεις επιδράσεις είναι οι συγκρίσεις των ενδεχόμενων αποτελεσμάτων, γεγονός που φαίνεται τόσο άμεσο, με αποτέλεσμα να δικαιολογείται η ύπαρξή του από παλιά. Στοιχεία αυτού του ορισμού των αιτιωδών επιδράσεων μπορούμε να βρούμε τόσο σε πειράματα όσο και σε κείμενα φιλοσόφων. Ας δούμε για παράδειγμα, το φιλόσοφο *John Stuart Mill*, ο οποίος συζητώντας και αναπτύσσοντας την οπτική του *Hume*, διατύπωσε [36]:

Αν ένα άτομο τρώει από ένα συγκεκριμένο πιάτο, με αποτέλεσμα να πεθάνει, δηλαδή, δεν θα είχε πεθάνει αν δεν είχε φάει το συγκεκριμένο φαγητό; κάποιιοι θα ήταν ικανοί να πουν ότι τρώγοντας από αυτό το πιάτο, ήταν η πηγή θανάτου του.

Ενώ ο *Fisher* [10] έγραψε:

Εάν πούμε, ‘Αυτό το αγόρι έγινε ψηλό γιατί έτρωγε σωστά’, δεν αποτυπώνουμε την αιτία και την επίδραση σε μεμονωμένη περίπτωση; Υποδηλώνουμε ότι το αγόρι στη πιθανότητα να μη τρεφόταν σωστά θα είχε υπάρξει κοντύτερος.

Παρά τα προφανή στοιχεία σε αυτές τις αναφορές, δεν υπήρξε κανένας επίσημος συμβολισμός για τα ενδεχόμενα αποτελέσματα μέχρι τον *Neyman* [38], που φαίνεται να υφίσταται η πρώτη μαθηματική ανάλυση που έχει γραφτεί για ένα τυχαίο πείραμα με αναφορά στα ενδεχόμενα αποτελέσματα. Η έννοια απέκτησε τυπικό χαρακτήρα όσον αφορά τα τυχαίοι πειράματα. Η προσήλωση του *Neyman* αποτέλεσε βασικό πλεονέκτημα γιατί

επέτρεψε τη πιο σαφή και συχνότερη πιθανολογική αιτιώδη συμπερασματολογία, να εξάγετε από δεδομένα που λαμβάνονται από τυχαιοποιημένα πειράματα.

Σχεδόν ταυτόχρονα, ο *Fisher* [11] δημιούργησε κατά κάποιο τρόπο διαφορετική μέθοδο συμπερασματολογίας για τα τυχαιοποιημένα πειράματα, επίσης με βάση στη τάξη των τυχαιοποιημένων μηχανισμών ανάθεσης. Η ιδέα της κεντρικής σημασίας των τυχαιοποιημένων πειραμάτων εκρημοσύσε μέχρι τη δεκαετία του 1920, αλλά ο *Fisher* ήταν ο πρώτος που πρότινε την απόδοση της τυχαιοποίησης των θεραπειών σε μονάδες και έπειτα τη χρήση της τυχαιοποίησης, για να δικαιολογήσει θεωρητικά την ανάλυση των αποτελεσμάτων; ένα σημείο στο οποίο δόθηκε έμφαση από τον *Neyman*

Παρά την άμεση αποδοχή της πρότασης του *Fisher*, στα τέλη του 1920, για τα τυχαιοποιημένα πειράματα και την έννοια του *Neyman* για τα ενδεχόμενα αποτελέσματα στα τυχαιοποιημένα πειράματα, το πλαίσιο αυτό δεν χρησιμοποιήθηκε πέραν των τυχαιοποιημένων πειραμάτων, για μισό αιώνα αργότερα, νωρίτερα από τον *Rubin* [44]. Η προσέγγιση που χρησιμοποιήθηκε σε μη τυχαιοποιημένα σύνολα, κατά τη διάρκεια του μισού αιώνα, και ακολουθεί τη καθιερωμένη εννοιολογία του *Neyman* για τα τυχαιοποιημένα πειράματα, βασίστηκε σε μαθηματικά μοντέλα, σχετίζοντας την παρατηρήσιμη τιμή της μεταβλητής αποτελέσματος $Y_{obs,i}$ με τις συμμεταβλητές και τους δείκτες θεραπείας που λαμβάνονται. Στη συνέχεια ορίζονται οι αιτιώδεις επιδράσεις ως παραμέτροι σε αυτά τα μοντέλα. Ένας στατιστικός θα μπορούσε φανερά να χρησιμοποιήσει τα ενδεχόμενα αποτελέσματα του *Neyman*, για να ορίσει τις αιτιώδεις επιδράσεις σε τυχαιοποιημένα πειράματα και το παρατηρήσιμο αποτέλεσμα στις μελέτες παρατήρησης.

Το πλαίσιο που περιγράφηκε, χρησιμοποιεί τα ενδεχόμενα αποτελέσματα για τον ορισμό των αιτιώδων επιδράσεων, το οποίο αποτελεί το πρώτο μέρος του *RCM*. Αυτή η σκοπιά απορρέει από όλα τα προβλήματα της στατιστικής συμπερασματολογίας για τις αιτιώδεις επιδράσεις, καθώς προβλήματα ελλειπών δεδομένων με ένα μηχανισμό δημιουργούν ελλειπή δεδομένα στα ενδεχόμενα αποτελέσματα. Το πλαίσιο των ενδεχόμενων αποτελεσμάτων έχει αποδεχτεί και υιοθετηθεί από τους περισσότερους ερευνητές κατά το τέλος του 20ου αιώνα. Ορισμένες φορές η αποδοχή αυτή και η υιοθέτηση γινόταν ρητά, αφαιρώντας το παρατηρούμενο αποτέλεσμα στο πλαίσιο των ενδεχόμενων αποτελεσμάτων και ορισμένες φορές γινόταν με λιγότερη σαφήνεια, προσπαθώντας ακόμα να δημιουργήσουν μια άλλη εκδοχή για την έννοια του παρατηρήσιμου αποτελέσματος, πριν την πλήρη αποδοχή του *RCM*.

Η παρότρυνση για τη χρήση των ενδεχομένων αποτελεσμάτων προκειμένου να προσδιοριστούν τα προβλήματα αιτιώδους συμπερασματολογίας, κυριαρχεί από την αρχή του 21ου

αίωνα, ειδικά στην Επιδημιολογία και στην Ιατρική στατιστική, καθώς επίσης και στις επιστήμες συμπεριφοράς.

2.2 Μηχανισμός ανάθεσης

Από τη προηγούμενη ενότητα διαπιστώσαμε πως οι αιτιώδεις επιδράσεις μοναδιαίου επιπέδου δε μπορούν να παρατηρηθούν. Συνεπάγεται ότι οι αιτιώδεις επιδράσεις δε μπορούν να παρατηρηθούν ούτε σε πληθυσμιακό επίπεδο, ακόμη και στη περίπτωση που ισχύει η υπόθεση *SUTVA*. Ένας μηχανισμός ανάθεσης απαιτείται, που θα ορίζει το τρόπο που θα επιλέγουμε ποια ενδεχόμενα αποτελέσματα παρατηρούνται για κάθε μονάδα. Τυπικά ο μηχανισμός ανάθεσης αποτελεί ένα πιθανολογικό μοντέλο που περιγράφει πως ορισμένες μονάδες επιλέγονται να λαμβάνουν την ενεργή θεραπεία και πως άλλες μονάδες επιλέγονται να λαμβάνουν τη θεραπεία ελέγχου. Εξηγεί τι πρέπει να κάνουμε (ή τι κάναμε) για να μάθουμε σχετικά με την ‘επιστήμη’ $X, Y(1), Y(0)$. Στην ενότητα αυτή αρχικά θα επεξηγήσουμε αυτό το μοντέλο μέσα από δύο τετριμμένα παραδείγματα, και στη συνέχεια θα παρουσιάσουμε μια επίσημη ερμηνεία για αυτό. Η διατύπωση του μηχανισμού ανάθεσης αποτελεί το δεύτερο μέρος του *RCM*.

2.2.1 Διευκρινίζοντας τη κρισιμότητα του μηχανισμού ανάθεσης

Θεωρούμε έναν ιατρό ο οποίος εξετάζει την εφαρμογή δύο εγχειρήσεων κάθε έναν από οκτώ ασθενείς, μια πρότυπη και μια νέα. Ο ιατρός αυτός είναι ένας ‘εξαιρετικός’: επιλέγει τη θεραπεία που είναι περισσότερο κατάλληλη για κάθε ασθενή! Στη περίπτωση που και οι δυο είναι εξίσου αποτελεσματικές, ρίχνει ένα αμερόληπτο νόμισμα. Ο Πίνακας 2.1 μας παρουσιάζει τα υποθετικά ενδεχόμενα αποτελέσματα, κατά τα έτη που έζησε μετά την εγχείρηση, σύμφωνα με κάθε θεραπεία, κάθε ένας από αυτούς τους οκτώ ασθενείς. Δίδει επίσης τις ανεξάρτητες αιτιώδεις επιδράσεις τους. Η στήλη W δείχνει ποια θεραπεία λαμβάνει κάθε ασθενής, $W_i = 0$ ή $W_i = 1$ για τον i -οστό ασθενή, ενώ οι δυο τελευταίες στήλες δείχνουν τα παρατηρήσιμα ενδεχόμενα αποτελέσματα. Το παράδειγμα αυτό επεξηγεί γιατί πρέπει να θεωρήσουμε το μηχανισμό ανάθεσης προκειμένου να εξάγουμε αξιόπιστα ενδεχόμενα αποτελέσματα.

Στο ιδανικό ιατρικό παράδειγμα που παρουσιάζουμε, η θεραπεία που λαμβάνει κάθε α-

Ιδανικό ιατρικό παράδειγμα

	Ενδεχόμενα αποτελέσματα		Παρατηρήσιμα δεδομένα		
	Y(0)	Y(1)	W	Y(0)	Y(1)
	13	14	1	?	14
	6	0	0	6	?
	4	1	0	4	?
	5	2	0	5	?
	6	3	0	6	?
	6	1	0	6	?
	8	10	1	?	10
	8	9	1	?	9
Πραγματικός μέσος όρος	7	5	Παρατηρήσιμος μέσος όρος	5.4	11

Πίνακας 2.1: Πραγματικός μέσος όρος αιτιώδους επίδρασης $\overline{Y(1)} - \overline{Y(0)} = 5 - 7 = -2$ και παρατηρήσιμος $\bar{y}_1 - \bar{y}_0 = 11 - 5.4 = 5.6 \neq -2$

σθενής εξαρτάται από τα ενδεχόμενα αποτελέσματα κάθε μονάδας. Σε αυτή τη περίπτωση είναι πολύ δύσκολο γίνει σωστή ανάλυση. Για παράδειγμα, αν θέλουμε να εξάγουμε ένα συμπέρασμα που βασίζεται στη παρατηρήσιμη διαφορά των μέσων (*means*), θα καταλήγαμε στο συμπέρασμα ότι η θεραπεία κατά μέσο όρο αφαιρεί δυο χρόνια ζωής. Επίσης από τους παρατηρήσιμους μέσους αντιλαμβανόμαστε ότι αν όλοι οι ασθενείς λάβουν τη νέα εγχείρηση, θα ζήσουν κατά μέσο όρο έντεκα ετών. Όμως από το πίνακα με τα ενδεχόμενα αποτελέσματα, γνωρίζουμε ότι αν όλοι λάβουν τη νέα εγχείρηση, θα ζήσουν κατά μέσο όρο πέντε ετών. Στο σημείο αυτό φαίνεται η λανθασμένη συμπερασματολογία. Ακόμη μελετώντας όλες τις παρατηρήσιμες τιμές, σημειώνεται ότι τα χρόνια ζωής μετά τη νέα εγχείρηση (9, 10, 14) είναι πολύ περισσότερα από τα χρόνια ζωής έπειτα από τη πρότυπη, (4,5,6,6,6). Μια παρατήρηση που μπορεί να οδηγεί σε επιπρόσθετα λανθασμένα αιτιώδη συμπεράσματα.

Ποιες είναι οι λανθασμένες υποθέσεις που έγιναν; Προκειμένου να έχουμε μια γενικότερη εικόνα για αυτό που συμβαίνει, δίδεται ο Πίνακας 2.2 που παρουσιάζει τι θα είχε παρατηρηθεί σε όλες τις πιθανές αναθέσεις αυτής της περίπτωσης. Υπάρχουν 56 πιθανές αναθέσεις σε κάθε μια από τις οποίες τρεις από τις οκτώ μονάδες λαμβάνουν κάποιου είδους θεραπεία. Υπολογίζουμε τη τιμή των παρατηρήσιμων μέσων, για τις 56 πιθανές αναθέσεις. Ο μέσος όρος όλων αυτών των πιθανών αναθέσεων ισούται με -2, το οποίο είναι ίδιο με τον πραγματικό

W	$\bar{y}_1 - \bar{y}_0$	$median(y_1) - median(y_0)$
11100000	-1.6	-5
11010000	-1.1	-4
11001000	-0.5	-3
11000100	-1.2	-5
11000010	2.2	4
11000001	1.9	3
10110000	-1.1	-4
10101000	-0.6	-3
10100100	-1.3	-5
10100010	2.1	4
10100001	1.8	3
10011000	-0.1	-3
10010100	-0.7	-4
10010010	2.7	4
10010001	2.3	3
10001100	-0.2	-3
10001010	3.2	4
10001001	2.9	3
10000110	2.5	4
10000101	2.2	3
10000011	5.6	4
01110000	-7.2	-7
01101000	-6.7	-7
01100100	-7.3	-7
01100010	-3.9	-5
01100001	-4.3	-5
01011000	-6.1	-6
01010100	-6.8	-7
01010010	-3.4	-4
01010001	-3.7	-4
01001100	-6.3	-7
01001010	-2.9	-3
01001001	-3.2	-3
01000110	-3.5	-5
01000101	-3.9	-5
01000011	-0.5	3
00111000	-6.2	-6
00110100	-6.9	-7
00110010	-3.5	-4
00110001	-3.8	-4
00101100	-6.3	-7
00101010	-2.9	-3
00101001	-3.3	-3
00100110	-3.6	-5
00100101	-3.9	-5
00100011	-0.5	3
00011100	-5.8	-6
00011010	-2.4	-3
00011001	-2.7	-3
00010110	-3.1	-4
00010101	-3.4	-4
00010011	0.0	3
00001110	-2.5	-3
00001101	-2.9	-3
00001011	0.5	3
00000111	-0.1	3
<i>Μεσος ορος</i>	-2	-2.3

Πίνακας 2.2: Όλες οι πιθανές αναθέσεις στο ιδανικό ιατρικό παράδειγμα

μέσο όρο. Αυτή η ισότητα δείχνει ότι όποια και αν ήταν η παρατηρηθείσα ανάθεση, που είχε επιλεχθεί τυχαία από τις 56 αναθέσεις, θα βρισκόταν σε συμφωνία με το μέσο όρο. Οι τυχαίες λοιπόν επιλογές δεν εξαρτώνται από τα ενδεχόμενα αποτελέσματα, οπότε ο συσχετιζόμενος

μηχανισμός ανάθεσης είναι μη συγχυτικός. Αντιθέτως, στο ιδανικό ιατρικό παράδειγμα, ο μηχανισμός ανάθεσης εξαρτάται από τα ενδεχόμενα αποτελέσματα οπότε είναι συγχυτικός. Παρατηρούμε με βεβαιότητα ότι ακόμη και στη πιο ακραία πιθανή ανάθεση, η διαφορά των παρατηρήσιμων μέσων απείχε από το πραγματικό έχοντας ακόμη και διαφορετικό πρόσημο (για την ανάθεση 11100000 ισούται με -1.6, και για την ανάθεση 00000111 με -0.1).

Καταλήγουμε στο συμπέρασμα ότι η παρατηρήσιμη διαφορά των μέσων είναι τελείως παραπλανητική σε αυτή τη περίπτωση. Το μεγαλύτερο πρόβλημα είναι ότι ουσιαστικά είχαμε προφασιστεί πως είχαμε ένα μη συγχυτικό παράγοντα. Το παράδειγμα αυτό δείχνει την σημαντικότητα του να βρούμε μια κατάλληλη στατιστική για το πραγματικό μηχανισμό ανάθεσης.

2.2.2 Το παράδοξο του *Lord*

Θεωρούμε τώρα ένα ‘παράδοξο’ της αιτιώδους συμπερασματολογίας, το οποίο εύκολα επιλύεται με τις απλές ιδέες που έχουμε ήδη παρουσιάσει. Ο *Lord* [34] παρουσίασε το παράδειγμα που ακολουθεί, προσπαθώντας να δείξει πόσο σημαντικό είναι να διατηρούμε αυτή την προοπτική ξεκάθαρη, όταν σκεφτόμαστε αιτιώδεις επιδράσεις παρεμβάσεων.

‘Ένα πανεπιστήμιο ενδιαφέρεται για την έρευνα των επιδράσεων του διαιτολογίου των φοιτητών, που είναι διαθέσιμη στην εστία του πανεπιστημίου, και την έρευνα των διαφορών μεταξύ των δυο φύλων στις επιδράσεις αυτές. Συγκεντρώνονται διάφοροι τύποι δεδομένων. Ιδίως, καταγράφεται το βάρος κάθε φοιτητή στην αρχή της σχολικής περιόδου, το Σεπτέμβρη και στο τέλος, τον Ιούλιο.’

Το αποτέλεσμα της μελέτης για το αρσενικό φύλο, είναι ότι το μέσο βάρος του είναι το ίδιο το τέλος της σχολικής χρονιάς σε σχέση με την αρχή. Στην πραγματικότητα, ολόκληρη η κατανομή των βαρών παραμένει αμετάβλητη, αν και κάποια άτομα αρσενικού φύλου έχασαν βάρος και μερικά πήραν, οι αυξήσεις και οι απώλειες ισορροπήστηκαν. Το ίδιο πράγμα συμβαίνει και για το θηλυκό φύλο. Με τη μόνη διαφορά ότι τα άτομα θηλυκού γένους ξεκίνησαν και τελείωσαν τη σχολική περίοδο ελαφρύτερες κατά μέσο όρο από τα άτομα αρσενικού γένους. Κατά μέσο όρο, δεν υπήρξε αύξηση βάρους ή απώλεια για κανένα από τα δυο φύλα. Από την περιγραφή του *Lord*, η ποσότητα που εκτιμάται είναι η διαφορά μεταξύ της αιτιώδους επίδρασης του διαιτολογίου που προσφέρει το πανεπιστήμιο στα άτομα αρσενικού γένους και η αιτιώδης επίδραση του ίδιου διαιτολογίου στα άτομα θηλυκού γένους.

Δηλαδή, ο αιτιώδης εκτίμηση είναι η διαφορά μεταξύ των αιτιωδών επιδράσεων για τα άτομα αρσενικού γένους και τα άτομα θηλυκού γένους.

Το παράδοξο παράγεται εξετάζοντας των αντιφατικών συμπερασμάτων των δύο στατιστικών αναλυτών, στους οποίους ζητήθηκε να σχολιάσουν τα δεδομένα. Ο στατιστικός 1 παρατήρησε ότι δεν υπάρχουν διαφορές μεταξύ των κατανομών βάρους του Σεπτεμβρίου και Ιουνίου, είτε για το αρσενικό είτε για το θηλυκό φύλο. Κατά συνέπεια, ο στατιστικός 1 συμπεραίνει ότι:

‘Όσο αυτά τα δεδομένα δεν μας απασχολούν, δεν υπάρχει καμία ένδειξη ενδιαφέρουσας επίδρασης του διαιτολογίου (ή οποιαδήποτε άλλη), για το βάρος των φοιτητών. Συγκεκριμένα, δεν υπάρχει ένδειξη διαφορικής επίδρασης στα δύο φύλα, εφόσον κανένα από τα δυο δεν δείχνει κάποια συστηματική αλλαγή.’

Ο στατιστικός 2 εξετάζει τα δεδομένα με ένα πιο σύνθετο τρόπο. Ουσιαστικά, εξετάζει τα άτομα αρσενικού και θηλυκού γένους, με με περίπου το ίδιο αρχικό βάρος Σεπτεμβρίου, μια υποομάδα που αποτελείται από υπέρβαρα άτομα θηλυκού γένους και μια υποομάδα υπέρβαρων ατόμων αρσενικού γένους. Παρατηρεί ότι αυτά τα άτομα αρσενικού γένους τείνουν να πάρουν βάρος κατά μέσο όρο, ενώ του θηλυκού γένους τείνουν να χάσουν βάρος κατά μέσο όρο. Παρατηρεί επίσης ότι αυτο το αποτέλεσμα είναι ισχύει, ανεξαρτήτως ποια ομάδα αρχικών βαρών εξετάζει (για την ακρίβεια, ο στατιστικός 2 του *Lord* χρησιμοποίησε προσαρμογή της συνδιακύμανσης, δηλαδή, προσαρμογή παλινδρόμησης). Το συμπέρασμά του, ωστόσο, είναι ότι μετά τον έλεγχο του αρχικού βάρους, το διαιτολόγιο έχει μια διαφορική θετική επίδραση στα άτομα αρσενικού φύλου σε σχέση με του θηλυκού, γιατί για τα άτομα αρσενικού και θηλυκού γένους με το ίδιο αρχικό βάρος, κατά μέσο όρο το αρσενικό φύλο παίρνει περισσότερο βάρος από το θηλυκό. Το αρχικό βάρος υπήρξε η τρίτη μεταβλητή, πέραν από το διαιτολόγιο (ανεξάρτητη μεταβλητή) και το βάρος (εξαρτημένη μεταβλητή). Ο έλεγχος του αρχικού βάρους (*controlling for*) αποτελεί μια μέθοδο μείωσης των συγχυτικών παραγόντων.

Ποιος από τους δυο στατιστικούς έχει δίκιο; Παρατηρούμε πως και οι δυο εστίασαν στην αποκτήση αποτελεσμάτων και να ανάκληση αυτών των αποτελεσμάτων, διότι δεν συγκρίναν τις αιτιώδεις επιδράσεις των ενδεχόμενων αποτελεσμάτων. Αν και οι δύο στατιστικοί περιόριζαν τις παρατηρήσεις τους στην περιγραφή των δεδομένων, και οι δύο θα μπορούσαν να είναι σωστοί, αλλά για την αιτιώδη συμπερασματολογία, και τα δύο είναι λάθος γιατί αυτά τα δεδομένα δεν μπορούν να υποστηρίξουν κάποια αιτιώδη συμπεράσματα για την επίδραση

του διαιτολογίου, χωρίς τον καθορισμό ισχυρών υποθέσεων.

Οι μονάδες που προφανώς είναι οι φοιτητές, και ο χρόνος της εφαρμογής της θεραπείας (το διαιτολόγιο του πανεπιστημίου) είναι ξεκάθαρα ο Σεπτέμβριος και ο χρόνος καταγραφής του αποτελέσματος Y είναι ξεκάθαρα ο Ιούνιος; Αποδεχόμαστε την υπόθεση σταθερότητας. Ποια είναι τα ενδεχόμενα αποτελέσματα και ποιος ο μηχανισμός ανάθεσης; Παρατηρούμε ότι η δήλωση του *Lord* στο πρόβλημα έχει επέλθει στην ήδη επικριμένη έννοια, το παρατηρήσιμο αποτέλεσμα, Y_{obs} , αντί του ενδεχόμενου αποτελέσματος.

Τα ενδεχόμενα αποτελέσματα είναι το βάρος τον Ιούνιο, σύμφωνα με το διαιτολόγιο του πανεπιστημίου, που περιγράφεται από την $Y_i(1)$ και σύμφωνα με το διαιτολόγιο ελέγχου περιγράφεται από την $Y_i(0)$. Οι συμμεταβλητές είναι το φύλο των φοιτητών, και το βάρος το Σεπτέμβρη. Αλλά ο μηχανισμός ανάθεσης έχει αναθέσει στον καθένα μια ενεργή θεραπεία! Δεν υπάρχει κάποιος, είτε αρσενικού είτε θηλυκού γένους, στο οποίο ανατέθηκε η θεραπεία ελέγχου. Ως εκ τούτου, δεν υπάρχει κάποια απόλυτα εμπειρική βάση στην οποία δύναται να συγκρίνουμε τις επιδράσεις, είτε απλές είτε διαφορικές, του διαιτολογίου του πανεπιστημίου με το διαιτολόγιο ελέγχου. Με την εισαγωγή των συμμεταβλητών, αρσενικού-θηλυκού γένους και αρχικού βάρους, περιπλέκουμε το πρόβλημα, ο *Lord* δημιούργησε μια μερική σύγχυση. Το ζήτημα εδώ είναι ότι το παράδοξο αμέσως επιλυθηκε μέσω χρήσης των ενδεχόμενων αποτελεσμάτων. Οποιαδήποτε απάντηση των δυο στατιστικών μπορούσε να είναι ορθή για την αιτιώδη συμπερασματολογία, ανάλογα με το τι επιθυμούμε να υποθέσουμε για τα ενδεχόμενα αποτελέσματα βάση του διαιτολογίου ελέγχου, που απουσιάζει εντελώς.

2.2.3 Είδη μηχανισμών ανάθεσης

Έχουμε καταλήξει στο συμπέρασμα πως για τη στατιστική συμπερασματολογία των αιτιωδών επιδράσεων είναι απαραίτητο ένα μοντέλο μηχανισμού ανάθεσης. Ο μηχανισμός ανάθεσης δίνει την δεσμευμένη πιθανότητα για κάθε διάλυσμα των αναθέσεων, δεδομένων των συμμεταβλητών και των ενδεχόμενων αποτελεσμάτων:

$$Pr(W|X, Y(0), Y(1)) \quad (2.1)$$

Εδώ το W είναι ένα $N \times 1$ διάλυσμα, όπως και προηγουμένως, και οι $X, Y(1)$ και $Y(0)$ είναι πίνακες N γραμμών. Ένα συγκεκριμένο παράδειγμα μηχανισμού ανάθεσης είναι ένα τελείως τυχαίο πείραμα με N μονάδες, όπου για τις $n < N$ μονάδες ανατίθεται η ενεργή θεραπεία, και στις υπόλοιπες $N - n$ μια θεραπεία ελέγχου:

$$Pr(W|X, Y(0), Y(1)) = \begin{cases} 1/\binom{N}{n} & \text{εάν } \sum W_i = n \\ 0 & \text{διαφορετικά} \end{cases} \quad (2.2)$$

Ένας μη συγχυτικός μηχανισμός ανάθεσης είναι ελεύθερος από την εξάρτηση είτε στην $Y(0)$ είτε στην $Y(1)$:

$$Pr(W|X, Y(0), Y(1)) = Pr(W|X) \quad (2.3)$$

Ο μηχανισμός ανάθεσης είναι πιθανολογικός αν κάθε μονάδα έχει μια θετική πιθανότητα στην λήψη οποιαδήποτε θεραπείας (Πιθανολογικό είναι ένα μοντέλο το οποίο έχει πολλαπλά ενδεχόμενα αποτελέσματα, όπου κάθε ένα έχει διαφορετικό βαθμό βεβαιότητας ή αβεβαιότητας, εμφάνισης. Με τον όρο αβεβαιότητα, εννούμε ότι μια παρατηρήσιμη τιμή μπορεί να διαφέρει από τη πραγματική):

$$0 < Pr(W_i|X, Y(0), Y(1)) < 1 \quad (2.4)$$

Αν ο μηχανισμός ανάθεσης είναι μη συγχυτικός (όταν η ανάθεση της θεραπείας ελέγχου ή της ενεργής θεραπείας, δεν εξαρτάται από όλα τα ενδεχόμενα αποτελέσματα, παρατηρήσιμα και μη) και πιθανολογικός, ονομάζεται ισχυρά αγνοήσιμος και είναι θεμελιώδης στην αιτιώδη συμπερασματολογία. Επειδή η αιτιώδης συμπερασματολογία είναι βασικά ένα πρόβλημα με ελλιπή δεδομένα, με τουλάχιστον τα μισά των ενδεχόμενων αποτελεσμάτων να απουσιάζουν, όταν δεν έχουμε αντίληψη της διαδικασίας που δημιουργεί τα ελλιπή δεδομένα, δεν υπάρχει πιθανότητα να εξάγουμε κανένα συμπέρασμα σχετικά με τις ελλειπείς τιμές. Δηλαδή, χωρίς ένα στοχαστικό μοντέλο που περιγράφει τον τρόπο με τον οποίο οι θεραπείες ανατίθενται στα άτομα, η τυπική αιτιώδης συμπερασματολογία είναι αδύνατη. Αυτό δεν σημαίνει ότι χρειάζεται να γνωρίζουμε τον μηχανισμό ανάθεσης, αλλά χωρίς τη θέσπιση ενός, δεν μπορούμε να κάνουμε κάποιο στατιστικό ισχυρισμό για τις αιτιώδεις επιδράσεις, όπως μια αμερόληπτη εκτίμηση ή την κάλυψη των διαστημάτων εμπιστοσύνης ή η κάλυψη των *posterior Bayesian* διαστημάτων.

Οι ισχυρά αγνοήσιμοι μηχανισμοί ανάθεσης, συχνά επιτρέπουν την άμεση εκτίμηση των αιτιωδών επιδράσεων. Επομένως, αυτοί οι μηχανισμοί ανάθεσης αποτελούν τη βάση για συμπερασματολογία για τις αιτιώδεις επιδράσεις σε πιο περίπλοκες καταστάσεις, όταν οι πιθανότητες ανάθεσης που εξαρτώνται με τρόπο που δεν γνωρίζουμε από τις συμμεταβλητές, ή όταν υπάρχει μη συμμόρφωση με τη θεραπεία που έχει ανατεθεί. Οι ισχυρά μη αγνοήσιμοι μηχανισμοί ανάθεσης, που βασικά είναι συλλογές απολύτως ξεχωριστών τυχαιοποιημένων

πειραμάτων κάθε τιμής της X_i , με μια διακριτή πιθανότητα ανάθεσης θεραπείας, αποτελεί τη βάση για την ανάλυση των μη τυχαιοποιημένων μελετών παρατήρησης.

Γνωρίζουμε ότι ο συγχυτικός μηχανισμός ανάθεσης, εξαρτάται από τα ενδεχόμενα αποτελέσματα και ισχύει:

$$Pr(W|X, Y(0), Y(1)) \neq Pr(W|X) \quad (2.5)$$

Μία ειδική κατηγορία πιθανών συγχυτικών μηχανισμών ανάθεσης που είναι ιδιαίτερα σημαντικοί στην στην *Bayesian* συμπερασματολογία είναι οι αγνοήσιμοι μηχανισμοί ανάθεσης. Οι αγνοήσιμοι μηχανισμοί ανάθεσης ορίζονται από την ανεξαρτησία τους από τα ελλειπή ενδεχόμενα αποτελέσματα:

$$Pr(W|X, Y(0), Y(1)) = Pr(W|X, Y_{obs}) \quad (2.6)$$

Αγνοήσιμοι συγχυτικοί μηχανισμοί ανάθεσης προκύπτουν κατά την εφαρμογή τους, ειδικά σε διαδοχικά πειράματα. Στη περίπτωση αυτή, η πιθανότητα της επόμενης μονάδας να εκτεθεί στην ενεργή θεραπεία, εξαρτάται από τα παρατηρήσιμα αποτελέσματα όσων προηγουμένως ήταν εκτεθειμένα στην ενεργή θεραπεία έναντι όσων τα παρατηρήσιμα αποτελεσμάτων ήταν εκτεθειμένα στην θεραπεία ελέγχου.

Όλοι οι μη συγχυτικοί μηχανισμοί ανάθεσης είναι αγνοήσιμοι, αλλά όλοι οι αγνοήσιμοι μηχανισμοί δεν είναι μη συγχυτικοί. Είναι εύκολο να παρατηρήσει κανείς γιατί οι αγνοήσιμοι και οι ισχυρά αγνοήσιμοι μηχανισμοί κατέχουν σημαντικό ρόλο στην αιτιώδη συμπερασματολογία, μέσα από τα δύο τετριμμένα παραδείγματα των Ενοτήτων 2.2.1 και 2.2.2: Το πρώτο, με τον 'εξαιρετικό' ιατρό, που αφορούσε έναν μη αγνοήσιμο μηχανισμό ανάθεσης θεραπείας, και το δεύτερο, το παράδοξο του *Lord* που αφορούσε ένα μη συγχυτικό παράγοντα και ένα μη πιθανολογικό μηχανισμό ανάθεσης, και δεν ήταν ισχυρά αγνοήσιμος.

2.3 Μορφές αιτιώδους συμπερασματολογίας που βασίζονται στην ανάθεση

Ουσιαστικά, υπάρχουν τρεις επίσημες στατιστικές μορφές της αιτιώδους συμπερασματολογίας, η μια είναι η *Bayesian*, που αντιμετωπίζει τα ενδεχόμενα αποτελέσματα σαν τυχαίες μεταβλητές, και οι άλλες δυο που βασίζονται μόνο στο μηχανισμό ανάθεσης, που αντιμετωπίζουν τα ενδεχόμενα αποτελέσματα σαν σταθερές αλλά άγνωστες ποσότητες. Ο *Rubin*

περιγράφει αυτές τις τρεις μορφές καθώς και τον συνδυασμό τους, τον οποίο όμως δεν θα αναπτύξουμε στο σημείο αυτό. Από τις δυο διακριτές μορφές της συμπερασματολογίας που βασίζονται στην ανάθεση, η μια οφείλεται στον *Neyman* [38] και η άλλη οφείλεται στον *Fisher* [11]. Θα περιγράψουμε μόνο τις δυο τελευταίες, αρχικά χωρίς την παρουσία των συμμεταβλητών X . Οι μορφές συμπερασματολογίας που βασίζονται στην ανάθεση, όπως αναπτύχθηκαν από τον *Fisher* και τον *Neyman* ήταν μορφές που βασίζονταν στην τυχαιοποίηση της συμπερασματολογίας, γιατί και οι δύο θεωρούνται τυχαιοποιημένα πειράματα. Στις πιο κάτω υποενότητες θα παρουσιάσουμε μια γενίκευση αυτών των μορφών που βασίζονται στην τυχαιοποίηση.

2.3.1 Συμπερασματολογία βασισμένη στη τυχαιοποίηση κατά *Fisher*

Η προσέγγιση του *Fisher* είναι πιο άμεση εννοιολογικά και για το λόγο αυτό παρουσιάζεται πρώτη. Είναι στενά συνδεδεμένη με τη μαθηματική ιδέα της απόδειξης με αντίφαση. Είναι κατά βάση μια στοχαστική απόδειξη με αντίφαση, δίνοντας το επίπεδο σημαντικότητας (ή την p -τιμή) - στη πραγματικότητα την εγκυρότητα - της μηδενικής υπόθεσης, που συχνά σημαίνει ότι δεν υπάρχει μια απολύτως επίδραση θεραπείας. Η μέθοδος του *Fisher* ισχύει μόνο για το σύνολο των μονάδων με $W_i = 1$ ή 0 , και όχι για τις μονάδες με $W_i = *$. Οπότε στην υποενότητα αυτή, υποθέτουμε ότι όλες οι μονάδες εκτίθενται είτε στην ενεργή θεραπεία είτε στην θεραπεία ελέγχου.

Το πρώτο στοιχείο της μορφής συμπερασματολογίας του *Fisher* είναι η μηδενική υπόθεση, που συχνά είναι $Y_i(1) \equiv Y_i(0)$ για όλες τις μονάδες: η θεραπεία δεν έχει απολύτως καμία επίδραση στα ενδεχόμενα αποτελέσματα. Βάσει της μηδενικής υπόθεσης, όλα τα ενδεχόμενα αποτελέσματα δίδονται από τις παρατηρήσιμες τιμές των ενδεχόμενων αποτελεσμάτων, Y_{obs} , γιατί $Y(1) \equiv Y(0) \equiv Y_{obs}$. Συνεπάγεται ότι, βάσει της μηδενικής υπόθεσης, η τιμή κάθε στατιστικής διαφοράς, όπως των παρατηρήσιμων μέσων όρων για τις μονάδες που εκτίθενται στην αγωγή 1 και στις μονάδες που εκτίθενται στην αγωγή 0, δηλαδή η διαφορά $\bar{y}_1 - \bar{y}_0$, είναι γνωστές, όχι μόνο για την παρατηρήσιμη ανάθεση, αλλά για όλες τις πιθανές αναθέσεις της W .

Ας υποθέσουμε ότι επιλέγουμε μια αριθμητική τιμή S όπως η διαφορά $\bar{y}_1 - \bar{y}_0$, και υπολογίζουμε την τιμή της σύμφωνα με κάθε πιθανή ανάθεση (υποθέτοντας την μηδενική υπόθεση) και επίσης υπολογίζουμε την πιθανότητα κάθε ανάθεσης σύμφωνα με τον τυχαιοποιημένο μηχανισμό ανάθεσης. Σε ορισμένα κλασσικά πειράματα, οι πιθανότητες αυτές είναι

είτε μηδέν είτε έχουν μια κοινή τιμή για όλες τις πιθανές αναθέσεις. Για παράδειγμα, σε ένα πλήρως τυχαιοποιημένο πείραμα με $N = 2n$ μονάδες, οι n επιλέγονται τυχαία ώστε να λαμβάνουν την θεραπεία 1, και οι υπόλοιπες n ώστε να λαμβάνουν την θεραπεία 0. Τότε οποιαδήποτε ανάθεση W που έχει την τιμή 1 ή 0, έχει πιθανότητα $1/\binom{N}{n}$, και οποιαδήποτε άλλη ανάθεση της W έχουν μηδενική πιθανότητα. Γνωρίζοντας τη τιμή της διαφοράς S για κάθε ανάθεση W , και την πιθανότητα του, μπορούμε να υπολογίσουμε τη πιθανότητα (βάση το μηχανισμό ανάθεσης και της μηδενικής υπόθεσης) που θα μπορούσε να παρατηρήσει μια τιμή της S τόσο ‘ασυνήθιστη’, ή περισσότερο ασυνήθιστη από, την παρατηρήσιμη τιμή της S, S_{obs} . ‘Ασυνήθιστη’ ορίζεται μια *a priori*, κατά βάση από το πόσο αποκλείει η S_{obs} από το 0. Αυτή η πιθανότητα είναι η εγκυρότητα (η p -τιμή ή το επίπεδο σημαντικότητας) της παρατηρήσιμης τιμής της στατιστικής διαφοράς S , βάσει της μηδενικής υπόθεσης: η πιθανότητα ενός αποτελέσματος (που αντιπροσωπεύεται από τη τιμή S_{obs} της στατιστικής διαφοράς, S) όσο σπάνιο, ή περισσότερο σπάνιο, από το πραγματικό παρατηρήσιμο αποτέλεσμα, αν η μηδενική υπόθεση ήταν αληθής, όπου η πιθανότητα είναι πάνω από την κατανομή που προκαλείται από το μηχανισμό ανάθεσης.

Η μορφή αυτή της συμπεραματολογίας είναι εκλεκτή: Αν τα δεδομένα δεν υποδηλώνουν ότι η μηδενική υπόθεση απολύτως καμίας επίδρασης θεραπείας είναι λανθασμένη, δεν είναι εύκολο να ισχυριστούμε απόδειξη διαφορετικών αποτελεσμάτων των ενεργών θεραπειών και των θεραπειών ελέγχου.

Η προσέγγιση του *Fisher* μπορεί να επεκταθεί και σε άλλες πιο αιχμηρές μηδενικές υποθέσεις, δηλαδή, μια μηδενική υπόθεση όπως από την γνώση της Y_{obs} , οι τιμές της $Y(1)$ και $Y(0)$ είναι γνωστές. Για παράδειγμα, από μια πρόσθετη μηδενική υπόθεση, που βεβαιώνει πως για κάθε μονάδα, η διαφορά $Y_i(1) - Y_i(0)$ είναι μια ειδική σταθερά, π.χ. ίση με τη τιμή 3. Η συλλογή τέτοιων μηδενικών υποθέσεων που δεν οδηγούν σε ακραίες p τιμές, μπορεί να χρησιμοποιηθεί για να δημιουργήσει ένα διάστημα εκτιμήσεων της αιτιώδους επίδρασης υποθέτοντας την προσθετικότητα. Οι επεκτάσεις σε άλλες στατιστικές και άλλους ειδικούς μηχανισμούς ανάθεσης, συμπεριλαμβάνοντας τους αντίστοιχους μη συγχυτικούς, ακόμα και τους μη αγνοήσιμους, είναι άμεσες γιατί όλα τα ενδεχόμενα αποτελέσματα είναι γνωστά από την Y_{obs} . Κατά συνέπεια οι πιθανότητες κάθε ανάθεσης είναι γνωστές. Ο *Fisher*, ωστόσο, ποτέ δεν αναφέρθηκε σε τέτοιου είδους επεκτάσεις. Παρατηρούμε ότι η άποψη του *Fisher*, δεν μας παρέχει την ικανότητα να γενικεύσουμε εκτός των μονάδων σε ένα πείραμα, είτε να θεωρήσουμε ‘ενοχλητικές’ τις μηδενικές υποθέσεις, όταν υπάρχουν πολλαπλές θεραπείες. Αυτοί οι περιορισμοί δεν παρουσιάζονται στην προσέγγιση του *Neymann*.

Ο έλεγχος *Fisher* σε ένα πλήρως τυχαιοποιημένο πείραμα

Πριν από την εξέταση των παρατηρήσιμων δεδομένων Y

1. Καθορισμός μιας αιχμηρής μηδενικής υπόθεσης (μια υπόθεση σχετικά με το μέγεθος της επίδρασης της θεραπείας σε κάθε μονάδα). Συνήθως χρησιμοποιείται η μηδενική υπόθεση, η οποία δεν έχει απολύτως καμία επίδραση θεραπείας ($Y_i(1) \equiv Y_i(0)$ για όλες τις μονάδες)
2. Προσδιορισμός ενός στατιστικού ελέγχου για την εκτίμηση της επίδρασης της θεραπείας και την αξιολόγηση της μηδενικής υπόθεσης. Συχνά χρησιμοποιείται η διαφορά των παρατηρήσιμων μέσων ομάδων θεραπείας και ελέγχου ($\bar{y}_1 - \bar{y}_0$).

Χρησιμοποιώντας τα δεδομένα Y_{obs}

3. Υπολογίζεται η τιμή του στατιστικού ελέγχου, και καθορίζονται οι ακραίες τιμές (δηλαδή οι ασυνήθιστες)
4. Συμπλήρωση των ελλειπών ενδεχόμενων αποτελεσμάτων με τη χρήση της αιχμηρής μηδενικής υπόθεσης και των παρατηρήσιμων Y τιμών.

Λήψη της p -τιμής

5. Για κάθε πιθανή ανάθεση, υπολογίζεται η τιμή ενδιαφέροντος του στατιστικού ελέγχου που θα είχε παρατηρηθεί υπό αυτή την ανάθεση (ίδιος υπολογισμός με το 3ο βήμα με τη μόνη διαφορά των παρατηρήσιμων τιμών).
6. Καθορισμός του πόσο ακραίες είναι οι τιμές που παρατηρήθηκαν στο 3ο βήμα. Αυτό καλείται επίπεδο σημαντικότητας ή p -τιμή.

2.3.2 Συμπερασματολογία βασισμένη στη τυχαιοποίηση κατά *Neyman*

Η μορφή του *Neymann* για τη συμπερασματολογία βάσει τη τυχαιοποίηση, μπορεί να θεωρηθεί ως την εξαγωγή συμπερασμάτων, αξιολογώντας τις αναμενόμενες τιμές σύμφωνα με τη κατανομή που επάγεται από το μηχανισμό ανάθεσης, με σκοπό τον υπολογισμό διαστήματος εμπιστοσύνης για την τυπική αιτιώδη επίδραση. Ουσιαστικά, μια αμερόληπτη εκτιμήτρια της επίδρασης της θεραπείας δημιουργείται (η τυπική αιτιώδης επίδραση, για παράδειγμα, ο μέσος όρος, ή ο μέσος). Επίσης διαπιστώνουμε την ύπαρξη μια αμερόληπτης εκτιμήτριας της διακύμανσης, της αντίστοιχης αμερόληπτης εκτιμήτριας. Κατόπιν χρησιμοποιούμε το Κεντρικό Οριακό Θεώρημα για την κανονικότητα της εκτιμήτριας, σύμφωνα με την κατανομή τυχαιοποίησής της, από όπου λαμβάνεται ένα διάστημα εμπιστοσύνης για την αιτιώδη εκτιμήτρια.

Πιο συγκεκριμένα, η αιτιώδης ποσότητα που εκτιμούμε είναι συνήθως ο μέσος όρος της αιτιώδους επίδρασης, $\overline{Y(1)} - \overline{Y(0)}$, όπου οι μέσοι όροι προέρχονται από όλες τις μονάδες του πληθυσμού που μελετάται και η κλασική στατιστική για την εκτίμηση αυτής της επίδρασης, είναι η διαφορά των παρατηρήσιμων μέσων για τις δύο ομάδες, $\bar{y}_1 - \bar{y}_0$. Μπορεί να αποδειχτεί ότι η τελευταία είναι αμερόληπτη για την $\overline{Y(1)} - \overline{Y(0)}$ σε ένα πλήρως τυχαιοποιημένο σχεδιασμό. Μια κοινή επιλογή για την εκτίμηση της διακύμανσης της $\bar{y}_1 - \bar{y}_0$ σύμφωνα με την τυχαιοποιημένη κατανομή της, σε πλήρως τυχαιοποιημένα πειράματα με $N = n_1 + n_0$ μονάδες, είναι η $se^2 = s_1^2/n_1 + s_0^2/n_0$. Όπου οι s_1^2, s_0^2, n_1 και n_0 είναι οι διακυμάνσεις των παρατηρήσιμων δειγμάτων και τα μεγέθη δείγματος στις δύο ομάδες θεραπείας. Ο *Neyman* έδειξε ότι η ποσότητα se^2 υπερεκτιμά την πραγματική διακύμανση της $\bar{y}_1 - \bar{y}_0$, εκτός κι αν η ισχύει η προσθετικότητα (δηλαδή εκτός και αν όλες οι αιτιώδεις επιδράσεις είναι σταθερές). Σε αυτή τη περίπτωση η ποσότητα se^2 είναι αμερόληπτη για την διακύμανση $\bar{y}_1 - \bar{y}_0$. Το 95% διάστημα εμπιστοσύνης για την $\overline{Y(1)} - \overline{Y(0)}$ είναι το $\bar{y}_1 - \bar{y}_0 \pm 1.96se$, το οποίο για μεγάλα δείγματα, περιλαμβάνει την $\overline{Y(1)} - \overline{Y(0)}$ τουλάχιστον στο 95% των πιθανών τυχαίων αναθέσεων.

Η μορφή του *Neyman* για τη συμπερασματολογία είναι λιγότερο άμεση από του *Fisher*. Αποσκοπεί στην εκτίμηση των διαδικασιών: Στις επαναλαμβανόμενες εφαρμογές θεραπειών, πόσο συχνά το διάστημα $\bar{y}_1 - \bar{y}_0 \pm 1.96se$ περιλαμβάνει την $\overline{Y(1)} - \overline{Y(0)}$; Ωστόσο, αποτελεί τη θεωρητική βάση για το μεγαλύτερο σύνολο των εφαρμογών, συμπεριλαμβανομένων των ιατρικών πειραμάτων. Εν τούτοις η προσέγγιση του *Neymann*, δεν γίνεται με σκοπό να μας εξηγήσει τον τρόπο να δημιουργήσουμε μια επαγωγική διαδικασία, αλλά μας δείχνει πως να αξιολογήσουμε μια προτεινόμενη διαδικασία για την εξαγωγή αιτιωδών συμπερασμάτων. Κατά συνέπεια, δεν είναι η κατάλληλη για την αντιμετώπιση σύνθετων προβλημάτων, με εξαίρεση τον τρόπο αξιολόγησης των προτεινόμενων απαντήσεων που λαμβάνονται βάσει γνώσης ή άλλης μεθόδου. Η προσέγγιση του *Fisher* επίσης μειονεκτεί της έλλειψης οδηγιών, στην πραγματικότητα, περισσότερο γιατί υπάρχει μικρή καθοδήγηση στην προσέγγισή του για το ποιους στατιστικούς έλεγχους να χρησιμοποιήσουμε ή για τον τρόπο ορισμού του 'πιο ασυνήθιστου'.

2.3.3 Συμπερασματολογία βασισμένη σε ένα μοντέλο

Η τρίτη μορφή στατιστικής συμπερασματολογίας για τις αιτιώδεις επιδράσεις είναι η συμπερασματολογία που βασίζεται σε ένα μοντέλο, στην οποία μορφή το μοντέλο για τον μηχανισμό ανάθεσης, $Pr(W|X, Y(1), Y(0))$, συμπληρώνεται με ένα μοντέλο για τα δεδομένα,

$Pr(Y(1), Y(0)|X)$, που σχεδόν πάντα συντάσσεται από άγνωστες παραμέτρους θ . Κατά τη γνώμη μας, η *Bayesian* μορφή της συμπερασματολογίας που βασίζεται σε ένα μοντέλο, είναι πιο γενική και εννοιολογικά πιο ικανοποιητική. (Στη προσέγγιση αυτή, η παράμετρος θ δίνει μια *a priori* κατανομή και η αιτιώδης συμπερασματολογία λαμβάνεται από μια δεσμευμένη κατανομή του αιτιώδους εκτιμητή δεδομένων των παρατηρήσιμων δεδομένων). Αναλυτικότερα, ας υποθέσουμε ότι ο αιτιώδης εκτιμητής είναι ο $\overline{Y(1)} - \overline{Y(0)}$, όπως προηγουμένως. Οπότε η *posterior* κατανομή (δηλαδή η δεδομένη κατανομή των δεδομένων των προδιαγραφών του μοντέλου και των παρατηρήσιμων τιμών των W, X και Y), προκύπτει από την *posterior* κατανομή πρόβλεψης των ελλειπών τιμών Y_{mis} , αξιολογημένες στις παρατηρήσιμες τιμές των W, X και Y_{obs} . Για μεγάλα δείγματα, η ανάλυση κατά *Bayesian* για ένα ανεξάρτητο κανονικό μοντέλο είναι σύμφωνο με τη προσέγγιση του *Neyman* για τα διαστήματα εμπιστοσύνης. Η *posterior* κατανομή της $\overline{Y(1)} - \overline{Y(0)}$ είναι κανονική, με μέσο $\bar{y}_1 - \bar{y}_0$ και η διασπορά $s_1^2/n_1 + s_0^2/n_0 - c$, όπου η c είναι μηδέν για την συμπερασματολογία των μέσων σε έναν άπειρο πληθυσμό. Σε ένα πεπερασμένο πληθυσμό η c είναι ανάλογη με την διασπορά των ατομικών αιτιωδών επιδράσεων και ως εκ τούτου είναι μηδενική όταν η αιτιώδης επίδραση είναι προσθετική. Θα πρέπει να επισημάνουμε σε αυτό το σημείο ότι η παράμετρος θ κάθε μοντέλου (για παράδειγμα, οι συντελεστές παλινδρόμησης σε κανονικά γραμμικά μοντέλα παλινδρόμησης) δεν είναι αιτιώδεις επιδράσεις. Για παράδειγμα, ας υποθέσουμε ότι η Y είναι μια δυαδική μεταβλητή που υποδεικνύει 5-χρόνια επιβίωση. Μπορεί να είναι λογικό να διαμορφώσουμε τη *log - odds* της πιθανότητας επιβίωσης, σαν μια επιπρόσθετη συνάρτηση της ομάδας θεραπείας και των συμμεταβλητών. Από την άλλη πλευρά, οι αιτιώδεις επιδράσεις που μας ενδιαφέρουν, μπορεί να είναι άμεσες συγκρίσεις των ποσοστών επιβίωσης σύμφωνα με την ενεργή θεραπεία και του ελέγχου. Ο ερευνητής ορίζει τον αιτιώδη εκτιμητή, αλλά όχι τη μορφή της κατανομής για τις μεταβλητές.

Η *Bayesian* προσέγγιση που είναι βασισμένη σε ένα μοντέλο, είναι από τις πιο άμεσες και ευέλικτες μορφές συμπερασματολογίας για τις αιτιώδεις επιδράσεις. Ωστόσο, οι προσεγγίσεις βασιζόμενες σε μοντέλο πετυχαίνουν αυτά τα πλεονεκτήματα με το να αξιώνουν μια κατανομή για τα δεδομένα $Pr(Y(1), Y(0)|X)$, που αποφεύγουν οι προσεγγίσεις που βασίζονται στην τυχαιοποίηση. Μια τέτοια κατανομή μπορεί να είναι πολύ χρήσιμη, αλλά μπορεί να είναι και σαν ένα όπλο στα χέρια ενός παιδιού, δηλαδή γεμάτο κίνδυνο για τον αφελή αναλυτή δεδομένων. Τα μοντέλα χρειάζεται να είναι πολύ προσεκτικά διαλεγμένα για να ανταλλάκουν τα βασικά χαρακτηριστικά του προβλήματος και θα πρέπει να ελέγχονται για σημαντικές παραβιάσεις και να τροποποιούνται όταν είναι απαραίτητο.

Εν περιλήψει, κάποιος πρέπει να είναι πρόθυμος για να χρησιμοποιήσει τα καλύτερα

χαρακτηριστικά όλων αυτών των προσεγγίσεων συμπερασματολογίας. Σε απλά κλασσικά τυχαιοποιημένα πειράματα με κανονικοποιημένα δεδομένα, οι τρεις προσεγγίσεις δίνουν παρόμοιες πρακτικές απαντήσεις, αλλά δεν ανταποκρίνονται σε πιο δύσκολες υποθέσεις στις οποίες κάθε προοπτική παρέχει διαφορετικές δυνατότητες.

2.3.4 Ο ρόλος των συμμεταβλητών στα τυχαιοποιημένα πειράματα

Όπως αναφέραμε και νωρίτερα, οι συμμεταβλητές είναι μεταβλητές των οποίων οι τιμές δεν επηρεάζονται από την θεραπεία ανάθεσης. Αποτελούν μεταβλητές των οποίων οι τιμές καθορίζονται πριν από την τυχαιοποίηση στις ομάδες θεραπείας (για παράδειγμα, το έτος γέννησης, η αρχική αρτηριακή πίεση ή η χοληστερίνη). Σε κλασσικά τυχαιοποιημένα πειράματα, αν μια συμμεταβλητή χρησιμοποιείται στο μηχανισμό ανάθεσης, για παράδειγμα ως μεταβλητή *block* σε ένα τυχαιοποιημένο *block* σχεδιασμό, πρέπει να αντικατοπτρίζεται στην ανάλυση, διότι επηρεάζει την κατανομή τυχαιοποίησης που επάγεται από το μηχανισμό ανάθεσης. Επίσης, οι συμμεταβλητές μπορούν να χρησιμοποιηθούν για την αύξηση αποδοτικότητας εκτίμησης, ακόμα κι όταν δεν χρησιμοποιούνται στο μηχανισμό ανάθεσης.

Το νόημα για την αύξηση αποδοτικότητας μπορεί να φανεί στο πλαίσιο ενός πλήρως τυχαιοποιημένου πειράματος στην Ιατρική με την μεταβλητή X να αποτελεί την αρχική χοληστερίνη και την Y τη μετά τη θεραπεία χοληστερίνη. Είτε από τη προοπτική του *Fisher* είτε του *Neyman*, μπορούμε να χρησιμοποιήσουμε τις συμμεταβλητές για να ορίσουμε μια νέα στατιστική, για την εκτίμηση των αιτιωδών ποσοτήτων. Για παράδειγμα, κάποιος μπορεί να χρησιμοποιήσει την διαφορά του μέσου όρου της παρατηρήσιμης μείωσης χοληστερίνης, $(\bar{y}_1 - \bar{x}_1) - (\bar{y}_0 - \bar{x}_0)$, όπου η \bar{x}_1 και η \bar{x}_0 είναι οι μέσες παρατηρήσιμες X τιμές για όσους εκτέθηκαν στην $W = 1$ και στην $W = 0$ θεραπεία, αντίστοιχα, παρά την διαφορά των μέσων Y τιμών, η $\bar{y}_1 - \bar{y}_0$, για την εκτίμηση της $\overline{Y(1)} - \overline{Y(0)}$. Υποθέτουμε ότι οι μεταβλητές X και Y συσχετίζονται, το οποίο είναι αναμενόμενο για την αρχική και τη μετά τη θεραπεία χοληστερίνη. Βάσει άποψης του *Neyman*, η διακύμανση της διαφοράς της μέσης $Y - X$ αλλαγής, θα πρέπει να είναι μικρότερη από τη διακύμανση της διαφοράς των μέσων Y τιμών, το οποίο εκφράζεται σε μικρότερες εκτιμήσεις διακυμάνσεων και κατά συνέπεια σε μικρότερα διαστήματα εμπιστοσύνης. Βάσει άποψης του *Fisher*, η μειωμένη αυτή διακύμανση εκφράζεται σε πιο σημαντικές p -τιμές, όταν η μηδενική υπόθεση είναι λανθασμένη.

Το ζήτημα αυτό απεικονίζεται εύκολα μέσα από παραδείγματα. Υποθέτουμε μια ακραία περίπτωση, όπου η νέα θεραπεία αφαιρεί 10 μονάδες της αρχικής χοληστερίνης σε όλες τις

μονάδες μελέτης, όπου η παλιά θεραπεία δεν προκαλεί καμία αφαίρεση. Οι παρατηρήσιμες $Y - X$ μεταβολές έχουν ουσιαστικά μηδενική διακύμανση σε κάθε ομάδα θεραπείας, όπου οι Y τιμές έχουν ίδιες διακυμάνσεις για κάθε ομάδα θεραπείας, όπως και οι X τιμές. Το αποτέλεσμα σημαίνει ότι το διάστημα εμπιστοσύνης βάσει τον *Neyman* για την επίδραση θεραπείας, που στηρίζεται στη διαφορά του μέσου όρου των ατόμων που λαμβάνουν θεραπεία μείον όσων λαμβάνουν τη θεραπεία ελέγχου, στις $Y - X$ μεταβολές, είναι πολύ μικρότερες από τις αντίστοιχο διάστημα που βασίζεται στις Y τιμές. Επίσης, η παρατηρήσιμη τιμή της διαφοράς μεταβολών $Y - X$, είναι η πιο ακραία τιμή που μπορεί να παρατηρηθεί από τη μηδενική υπόθεση του *Fisher*, οπότε το παρατηρήσιμο αποτέλεσμα με τις $Y - X$ μεταβολές είναι τόσο σημαντικό, σε αντίθεση με τη διαφορά των Y τιμών.

2.3.5 Πιο σύνθετα τυχαιοποιημένα πειράματα

Η διατύπωση των ενδεχόμενων αποτελεσμάτων των αιτιωδών επιδράσεων επεκτείνεται με έναν φυσικό τρόπο σε πιο σύνθετους τυχαιοποιημένους σχεδιασμούς. Για παράδειγμα, αν οι $K > 2$ θεραπείες είναι τυχαιοποιημένες, τότε κάθε μονάδα έχει K ενδεχόμενα αποτελέσματα, ένα για κάθε θεραπεία. Αν οι θεραπείες είναι συνδυασμοί σε έναν παραγοντικό σχεδιασμό με J παράγοντες, με τον παράγοντα j να έχει k_j επίπεδα, τότε υπάρχουν $K = k_1 \times k_2 \times \dots \times k_j$ συνδυασμοί θεραπείας, και οι μέθοδοι ανάλυσης του κλασικού πειραματικού σχεδιασμού μπορεί να εφαρμοστεί για τη μελέτη των αιτιωδών επιδράσεων των παραγόντων και των αλληλεπιδράσεών τους.

Σε μια μελέτη με διαδοχικές υποθέσεις θεραπειών και ικανότητας αγνόησης, υποθέτουμε ότι υπάρχουν J χρονικά σημεία στα οποία οι θεραπείες αναθέτονται, και στο χρονικό σημείο j υπάρχουν k_j πιθανές επιλογές θεραπείας. Όπως προηγουμένως, υπάρχουν $K = k_1 \times k_2 \times \dots \times k_j$ πιθανοί συνδυασμοί θεραπείας και K ενδεχόμενα αποτελέσματα για κάθε άτομο στην μελέτη. Η ανάθεση μιας επιλογής θεραπείας σε ένα συγκεκριμένο χρονικό σημείο, ενδεχομένως να εξαρτάται από την Y_{obs} μέσω των αποτελεσμάτων για τις θεραπείες που έχουν καταγραφεί πριν από το χρονικό σημείο.

Όταν η πιθανότητα ανάθεσης θεραπείας επιτρέπεται να εξαρτάται από τις συμμεταβλητές ή τα αποτελέσματα που έχουν καταγραφεί, έπεται ότι η πιθανότητα αυτή ποικίλει από μονάδα σε μονάδα, και οι ατομικές πιθανότητες ανάθεσης $p_i \equiv Pr(W_i | X_i, Y_{obs})$ ονομάζονται βαθμολογίες τάσης. Για τη *Neyman* συμπερασματολογία, απαιτούνται να είναι αυστηρά μεταξύ 0 και 1, και είναι βασικές ποσότητες την ανάλυση. Όταν οι βαθμολογίες τάσης, είναι γνωστές, ο μηχανισμός ανάθεσης είναι βασικά γνωστός, και απλές γενικεύσεις των μορφών

Fisher και *Neyman*, μπορούν να εφαρμοστούν.

2.4 Κίνδυνοι μη συγχυτικών μηχανισμών ανάθεσης

Στην Ιατρική η συμμόρφωση (*compliance*) περιγράφει το βαθμό στον οποίο ένας ασθενής ακολουθεί σωστά τις ιατρικές συμβουλές. Με τον όρο μη-συμμόρφωση (*noncompliance*) αναφερόμαστε στο ακριβώς αντίθετο γεγονός. Η μη-συμμόρφωση αποτελεί ένα σημαντικό εμπόδιο για την αποτελεσματική παροχή της υγειονομικής περίθαλψης. Οι τυχαιοποιημένες μελέτες πλήρους συμμόρφωσης και χωρίς ελλειπή δεδομένα, αναλύονται σχετικά εύκολα από μεθόδους που περιγράφηκαν προηγουμένως. Ωστόσο πρακτικά η συμπερασματολογία για τις αιτιώδεις επιδράσεις, είναι συχνά περίπλοκη λόγω έλλειψης συμμόρφωσης και ελλειπών δεδομένων. Η θεώρηση των ενδεχόμενων αποτελεσμάτων, παρέχει εννοιολογική σαφήνεια και ένα σύνολο μέσων για την χειρισμό αυτών των περιπλοκών. Στην ενότητα αυτή εξετάζουμε το ζήτημα της μη-συμμόρφωσης, και στη συνέχεια το ζήτημα των ελλειπών δεδομένων.

2.4.1 Μη-συμμόρφωση

Η ανάλυση μιας μη τυχαιοποιημένης μελέτης είναι περίπλοκη όταν τα αντικείμενα μελέτης της δεν συμμορφώνονται με την θεραπεία που ανατίθεται. Μια ανάλυση του σκοπού της θεραπείας (*intention – to – treat, ITT*), συγκρίνει τα αποτελέσματα των αντικειμένων που προέρχονται από τις τυχαιοποιημένες ομάδες (όπως μια τυχαιοποιημένη ανάλυση), αγνοώντας τις πληροφορίες συμμόρφωσης. Η επίδραση *ITT* μετρά την επίδραση της θεραπείας τυχαιοποίησης αντί να μετρά την επίδραση της θεραπείας όσων πραγματικά την λαμβάνουν. Ο *ITT* εκτιμητής προστατεύεται από τη μεροληψία επιλογής από την τυχαιοποιημένη θεραπεία ανάθεσης, αλλά σε γενικές γραμμές είναι ένα ψευδές μέτρο της επίδρασης αυτής της θεραπείας. Εναλλακτικά, τα αντικείμενα μελέτης μπορούν να ταξινομηθούν από τις θεραπείας που πραγματικά λαμβάνονται (ανάλυση *as – treated*). Το πρόβλημα τότε είναι ότι παραβιάζεται η τυχαιοποίηση, και οι συγχυτικοί παράγοντες συσχετιζόμενοι με την εναλλαγή των θεραπειών ανάθεσης, ενδεχομένως για να αλλοιώσουν την αιτιώδη ερμηνεία των επιδράσεων θεραπείας.

Ο ορισμός του ενδεχόμενου αποτελέσματος των αιτιωδών επιδράσεων, παρέχει μια χρησιμική βάση για την κατανόηση προβλημάτων μη-συμμόρφωσης και των υποθέσεων που συνεπάγονται από διάφορες στρατηγικές εκτίμησης. Οι συμμετέχοντες ταξινομούνται σε έναν από τους τέσσερις τύπους, στους *compliers*, στους *defiers*, στους *never – takers* και

στους *always – takers*. Οι *compliers* είναι ασθενείς που μπορούν να ενστερνιστούν οποιαδήποτε θεραπεία τους ανατεθεί, οι *never – takers* είναι ασθενείς που θα λάβουν τη θεραπεία ελέγχου ανεξάρτητα με όποια θεραπεία τους ανατεθεί, οι *always – takers* θα λάβουν την ενεργή θεραπεία ανεξάρτητα με όποια θεραπεία τους έχει ανατεθεί, και οι *defiers* είναι ασθενείς που εστερνίζονται την αντίθετη θεραπεία που τους ανατίθεται. Ας συμβολίσουμε με τη μεταβλητή C το δείκτη συμμόρφωσης του συμμετέχοντα, με $C = 1$ για τους *compliers* και $C = 0$ για όλους τους *noncompliers*: τους *defiers*, τους *never – takers* και τους *always – takers*.

Παρατηρούμε τη κατάσταση συμμόρφωσης μόνο για την θεραπεία που έχει ανατεθεί, έτσι η κατάσταση πλήρους συμμόρφωσης των αντικειμένων μελέτης να μην παρατηρείται πλήρως. Ειδικότερα, αν στο αντικείμενο ανατίθεται η νέα θεραπεία και συμμορφώνεται, τότε το συγκεκριμένο αντικείμενο μπορεί να είναι ένας *complier* ή ένας *always – taker*. Αν σε ένα αντικείμενο μελέτης ανατίθεται μια νέα θεραπεία και αποτυγχάνει να συμμορφωθεί, τότε το αντικείμενο μπορεί να είναι ένας *never – taker* ή ένας *defier*. Αν σε έναν ασθενή ανατίθεται η θεραπεία ελέγχου και συμμορφώνεται, τότε το αντικείμενο μπορεί να είναι ένας *complier* ή ένας *never – taker*. Αν σε ένα αντικείμενο μελέτης ανατίθεται η θεραπεία ελέγχου και λάβει τη νέα θεραπεία, τότε το αντικείμενο μπορεί να είναι ένας *always – taker* ή ένας *defier*. Ο *Imbens & Rubin* αντιμετωπίζουν αυτή την επιμέρους πληροφόρηση για την συμμόρφωση σαν ένα πρόβλημα με ελλιπή δεδομένα.

Ας κάνουμε την *SUTVA* υπόθεση, έτσι ώστε οι αιτιώδεις επιδράσεις μοναδιαίου επιπέδου να μπορούν να οριστούν χωρίς αναφορά σε άλλα άτομα της μελέτης. Για έναν ασθενή που ανταποκρίνεται σε μια θεραπεία και καταχωρείται στην ομάδα W , υποθέτουμε ότι $A(W)$ είναι η θεραπεία που ενστερνίζεται και έστω ότι $Y(W)$ το ενδεχόμενο αποτέλεσμα. Η αιτιώδης επίδραση μοναδιαίου επιπέδου (W) της θεραπείας που ανατίθεται, στη θεραπεία (A) που ενστερνίζεται, είναι η $A(1) - A(0)$. Και η αιτιώδης επίδραση μοναδιαίου επιπέδου (W) της θεραπείας που ανατίθεται στο αποτέλεσμα (Y), είναι η $Y(1) - Y(0)$. Αυτές οι επιδράσεις μοναδιαίου επιπέδου της W στις θεραπείες A και Y , είναι γενικά μη παρατηρήσιμες, διότι στις μονάδες δεν μπορούν να ανατεθούν ταυτόχρονα στη θεραπεία ελέγχου και στη πειραματική θεραπεία. Ωστόσο, οι επίδρασεις κατά μέσο όρου *ITT*, ταυτόχρονα στην A και στην Y μπορούν να εκτιμηθούν εύκολα από τους μέσους των ομάδων που αποκρίνονται σε μια τυχαιοποιημένη μελέτη.

Μια άλλη αιτιώδη επίδραση που χρησιμοποιείται σε πολλές μελέτες θεραπειών, είναι η *CACE* (*Complier Average Causal Effect*), που είναι η κατά μέσο όρο αιτιώδη επίδραση για τον υποπληθυσμό των *compliers*, και συμβολίζεται με $C = 1$. Για τους *compliers*, για

τους οποίους $A(0) = 0$ και $A(1) = 1$, ισχύει

$$CACE = E(Y(1) - Y(0)|C = 1) \quad (2.7)$$

όπου η συνθήκη $C = 1$ είναι ταυτόσημη με τις συνθήκες $A(0) = 0$ και $A(1) = 1$. Η $CACE$ αποτελεί μια έγκυρη αιτιώδη επίδραση γιατί είναι ένα συνοπτικό μέτρο των επιδράσεων σε μοναδιαίο επίπεδο σε έναν υποπληθυσμό ενδιαφέροντος, που ονομάζονται *compliers*. Η $CACE$ και η επίδραση ITT συλλέγουν δυο χαρακτηριστικά θεραπείας του ενδεχόμενου αποτελέσματος ενδιαφέροντος - την επίδραση της θεραπείας σε εκείνους που μπορεί να οδηγηθούν να λάβουν μέρος σε αυτό το πείραμα και τις γενικές επιπτώσεις της θεραπείας σε όλο τον πληθυσμό, συμπεριλαμβανομένων των μη *noncompliers*.

Υπάρχουν δυο κοινές μέθοδοι ανάλυσης, η ανάλυση *as - treated* και η *per - protocol*, που είναι γενικότερα ελαττωματικές γιατί δεν εκτιμούν την επίδραση $CACE$ ή οποιαδήποτε άλλη αιτιώδη επίδραση μοναδιαίου επιπέδου [54]. Η *as - treated* ανάλυση ταξινομεί τα αντικείμενα μελέτης σύμφωνα με τη θεραπεία που ενστερνίζονται $A(W)$ και εκτιμά

$$E(Y(W)|A(W) = 1) - E(Y(W)|A(W) = 0)$$

Η τελευταία σχέση δεν αποτελεί το μέσο όρο των αιτιωδών επιδράσεων μοναδιαίου επιπέδου, διότι συγκρίνει τους μέσους όρους της Y για ομάδες με διαφορετικά χαρακτηριστικά. Συγκεκριμένα, συγκρίνει το μέσο όρο των αποτελεσμάτων όσων ενστερνίζονται τη θεραπεία 1 (δηλαδή, στους *compliers* ανατίθεται η θεραπεία 1, στους *defiers* ανατίθεται η θεραπεία 0, και στους *always - takers* ανατίθεται μια από τις δυο θεραπείες), με το μέσο όρο των αποτελεσμάτων όσων ενστερνίζονται τη θεραπεία 0 (δηλαδή, στους *compliers* ανατίθεται η θεραπεία 0, στους *defiers* ανατίθεται η θεραπεία 1, και στους *never - takers* ανατίθεται μια από τις δυο θεραπείες). Μια *per - protocol* ανάλυση συγκρίνει τα αντικείμενα μελέτης που πραγματικά ενστερνίζονται τις θεραπείες που τους ανατίθενται και εκτιμά

$$E(Y(1)|A(1) = 1) - E(Y(0)|A(0) = 0)$$

η οποία επίσης δεν αποτελεί το μέσο όρο των αιτιωδών επιδράσεων μοναδιαίου επιπέδου. Συγκρίνει τα αποτελέσματα όσων ανατίθεται και ενστερνίζονται τη θεραπεία 1 (δηλαδή στους *compliers* και στους *always - takers* ανατίθεται η θεραπεία 1) με τα αποτελέσματα όσων ανατίθεται και ενστερνίζονται τη θεραπεία 0 (δηλαδή, στους *compliers* και στους *never - takers* ανατίθεται η θεραπεία 0). Η διατύπωση του ενδεχόμενου αποτελέσματος των αιτιωδών επιδράσεων συμβάλλει να διευκρινιστούν οι ανεπάρκειες αυτών των κοινών

μεθόδων ανάλυσης. Το πρόβλημα της επίδρασης *CACE* είναι η διάκριση μεταξύ της ερμηνείας ενός πραγματικού *complier* (η συμμόρφωση σύμφωνα και με τις δύο θεραπείες) και ενός παρατηρήσιμου *complier* (η συμμόρφωση σύμφωνα με την θεραπεία που πραγματικά ανατίθεται).

Η επίδραση *CACE* στην εξίσωση (2.7) είναι ένας έγκυρος αιτιώδης εκτιμητής, αν και δεν είναι άμεσα εμφανές ο τρόπος με τον οποίο εκτιμάται γιατί η κατάσταση συμμόρφωσης C των ατόμων είναι γενικά άγνωστη. Ωστόσο, η επίδραση *CACE* μπορεί να εκτιμηθεί από τα δεδομένα σύμφωνα με την υπόθεση *SUTVA* και τη τυχαία ανάθεση της W , αν κάνουμε τις ακόλουθες επιπρόσθετες υποθέσεις [2]:

1. Περιορίζεται ο αποκλεισμός της θεραπείας ανάθεσης δεδομένης της θεραπείας που λαμβάνεται: για τους *never – takers* και για τους *always – takers*, των οποίων η θεραπεία που ενστερνίζονται είναι η ίδια ανεξάρτητα της θεραπείας που ανατίθεται, το αποτέλεσμα Y είναι το ίδιο ανεξάρτητα από τη θεραπεία που ανατίθεται, δηλαδή $Y(0) = Y(1)$ αν $A(0) = A(1)$.
2. Η θεραπεία ανάθεσης είναι προτιμότερη από τη θεραπεία που ενστερνίζεται κάποιος ασθενής (δεν υπάρχουν *defiers*).
3. Μη μηδενικός παρονομαστής (ο πληθυσμός περιλαμβάνει ορισμένους *compliers*).

Προκειμένου να παράγουμε μια εκτίμηση της *CACE* σύμφωνα με αυτές τις υποθέσεις, σημειώνουμε ότι:

$$E(Y(1) - Y(0)) = E(Y(1) - Y(0)|C = 1)Pr(C = 1) + E(Y(1) - Y(0)|C \neq 1)Pr(C \neq 1) \quad (2.8)$$

Σύμφωνα με τη 2η υπόθεση, η επίδραση της θεραπείας για τους *never – takers* και για τους *always – takers* είναι μηδέν. Αν, επιπροσθέτως, ισχύει η 3η υπόθεση (δεν υπάρχουν *defiers*), τότε ο δεύτερος όρος του δεξιού μέλους της εξίσωσης (2.8) είναι μηδέν, και

$$E(Y(1) - Y(0)|C = 1) = E(Y(1) - Y(0))/Pr(C = 1) \quad (2.9)$$

δηλαδή, η επίδραση *CACE* είναι η *ITT* επίδραση διαιρούμενη από το ποσοστό των συμμορφωμένων *compliers*. Τώρα σύμφωνα με την τυχαιοποιημένη κατανομή θεραπείας, ο *ITT* εκτιμητής για την επίδραση της W στην Y , η διαφορά των μέσων δειγμάτων $\bar{y}_1 - \bar{y}_0$ είναι

μια αμερόληπτος εκτίμηση του αριθμητή της εξίσωσης (2.9). Επίσης, έστω ότι p_1 είναι το ποσοστό των αντικείμενων της ομάδας θεραπείας που ενστερνίζονται τη νέα θεραπεία και ως υποθέσουμε ότι η p_0 είναι το ποσοστό των αντικειμένων της ομάδας ελέγχου που ενστερνίζονται τη νέα θεραπεία. Τότε η p_1 αποτελεί μια αμερόληπτη εκτιμήτρια του ποσοστού των *compliers* ή των *always – takers* (εκείνων που ενστερνίζονται τη θεραπεία που τους ανατίθεται), και η p_0 είναι μια αμερόληπτη εκτιμήτρια της ροπής των ληπτών (εκείνων που ενστερνίζονται τη θεραπεία όταν τους ανατίθεται η θεραπεία ελέγχου). Ως εκ τούτου η διαφορά $p_1 - p_0$ είναι μια αμερόληπτη εκτιμήτρια του ποσοστού των *compliers*, $Pr(C = 1)$. Κατά συνέπεια ένας προσεγγιστικά αμερόληπτος εκτιμητής της επίδρασης *CACE* είναι ο:

$$IVE = \frac{\bar{y}_1 - \bar{y}_0}{p_1 - p_0} \quad (2.10)$$

που είναι μια εκτιμώμενη *ITT* επίδραση διαιρούμενη με τη διαφορά των ποσοστών που ενστερνίζονται τη νέα θεραπεία, στις ομάδες ελέγχου και στις ομάδες θεραπείας. Η 3η υπόθεση βεβαιώνει ότι ο παρονομαστής στην εξίσωση (2.10) είναι μη μηδενικός. Η εξίσωση (2.10) ορισμένες φορές ονομάζεται ο εκτιμητής της βοηθητικής μεταβλητής (*IV*) (Για περισσότερες εφαρμογές αυτής της μεθόδου δείτε τον *Sommer & Zeger* [59] και *Etter* [9]). Ειδικότερα, η 1η υπόθεση, έχει βασικό ρόλο, και δεν αποτελεί συνέπεια της τυχαιοποίησης των θεραπειών.

Ένα ακόμη πλεονέκτημα της διατύπωσης του ενδεχόμενου αποτελέσματος είναι ότι οδηγεί σε περισσότερο αποδοτικούς εκτιμητές της *CACE* από τον εκτιμητή *IV* (εξίσωση (2.10)). Οι *Bayesian* μορφές για την συμπερασματολογία όσον αναφορά την επίδραση *CACE*, μεταχειρίζεται τον δείκτη συμμόρφωσης *C* σαν ελλείπων δεδομένα [28] και χρησιμοποιούν επαναληπτικές τεχνικές προσομοίωσης όπως την αύξηση δεδομένων [62].

2.4.2 Ελλειπή δεδομένα

Τα ελλειπή δεδομένα [32] είναι ένα διαδεδομένο πρόβλημα στις επιδημιολογικές και κλινικές μελέτες, ιδιαίτερα όταν περιλαμβάνουν επαναλαμβανόμενες μετρήσεις. Η στατιστική ανάλυση όταν υπάρχουν ελλειπείς τιμές, συχνά περιορίζεται σε απλές και σχετικά ειδικές αποτυπώσεις, όπως η απόρριψη των ελλειπών περιπτώσεων ή την απόδοση των μη υπό-συνθήκη μέσων ή την τελευταία καταγεγραμμένη τιμή σε μια *longitudinal* μελέτη. Αυτές οι αναλύσεις κάνουν ισχυρές υποθέσεις σχετικά με τη φύση του μηχανισμού των ελλειπών δεδομένων και συχνά αποδίδουν μεροληπτικές συμπερασματολογίες.

Οι μέθοδοι οι οποίες αποδίδουν τις ελλειπείς τιμές, έχουν το πλεονέκτημα ότι οι παρατηρήσιμες τιμές διατηρούνται στις ελλειπείς περιπτώσεις. Κατά συνέπεια, στις τυχαιοποιημένες μελέτες διατηρείται η εξισορρόπηση για τους παράγοντες σύγχυσης. Καλύτερες μέθοδοι απόδοσης, εξάγουν μια απόδοση από μια κατανομή πρόβλεψης των ελλειπών τιμών, δεδομένων των παρατηρήσιμων δεδομένων. Για παράδειγμα, σε μια απόδοση στοχαστικής παλινδρόμησης, κάθε ελλειπή τιμή αντικαθίσταται από την αντίστοιχη παλινδρόμηση πρόβλεψης συν ένα τυχαίο σφάλμα, με διακύμανση ίση με την εκτιμώμενη διακύμανση των υπολοίπων. Ένα σοβαρό μειονέκτημα όλων των μεθόδων απόδοσης, είναι πως δημιουργούν δεδομένα που δεν υπάρχουν. Συγκεκριμένα, μια απλή τιμή απόδοσης, δεν μπορεί να αντιπροσωπεύσει όλη την αβεβαιότητα για το ποια τιμή θα αποδοθεί, οπότε αναλύει εκείνες που παρουσιάζονται ως τιμές απόδοσης όπως στις παρατηρήσιμες τιμές που υποτιμούν την αβεβαιότητα.

Η πολλαπλή απόδοση [51][52] είναι μια τροποποιημένη μορφή απόδοσης η οποία διορθώνει αρκετές ατέλειες αυτής της μεθόδου. Αντί της απόδοσης ενός απλού συνόλου για την εξαγωγή των ελλειπών τιμών, δημιουργείται ένα σύνολο M συνόλου δεδομένων, κάθε ένα από τα οποία περιέχει διαφορετικά σύνολα για την εξαγωγή ελλειπών τιμών βάσει της κατανομής πρόβλεψης. Αναλύουμε τότε κάθε σύνολο M και συνδυάζουμε τα αποτελέσματα με έναν απλό τρόπο. Συνήθως η πολλαπλή μέθοδος απόδοσης δεν είναι δυσκολότερη από την απλή. Ο επιπρόσθετος υπολογισμός από την επαναλαμβανόμενη ανάλυση, M φορές δεν είναι βάρος, και οι μέθοδοι για τη συμπερασματολογία είναι άμεσες. Το σημαντικότερο είναι η δημιουργία 'καλών' κατανομών πρόβλεψης για τις ελλειπείς τιμές. Η βασική θεωρία της πολλαπλής απόδοσης είναι η *Bayesian*, και αυτή η μέθοδος είναι στενά συνδεδεμένη με τις μεθόδους προσομοίωσης του *Bayes*[62]. Ωστόσο, η μέθοδος έχει συνήθως καλές ιδιότητες αν το μοντέλο απόδοσης είναι προσεκτικά διευκρινισμένο[51].

Ακόμη μια ισχυρή και βασική προσέγγιση για το χειρισμό των ελλειπών δεδομένων, ειδικότερα σε μεγάλα δείγματα, είναι η μέγιστη πιθανόφάνεια με βάση ένα στατιστικό μοντέλο για τα δεδομένα και τον μηχανισμό ελλειπών δεδομένων[32]. Αυτή η μέθοδος σχετίζεται με τη πολλαπλή απόδοση, χρησιμοποιώντας όλες τις πληροφορίες σε πλήρεις και ατελείς περιπτώσεις, και παράγει συμπερασματολογίες που λαμβάνονται υπόψη για τις πληροφορίες που λείπουν, από την ύπαρξη των ελλειπών τιμών.

2.5 Η αιτιώδης συμπερασματολογία στις μελέτες παρατήρησης

Σε μελέτες παρατήρησης, στις οποίες η ανάθεση θεραπειών δεν είναι υπό τον έλεγχο του ερευνητή, πολλές επιδημιολογικές μελέτες υποθέτουν ότι η ανάθεση θεραπείας είναι αγνοήσιμη, αφού ενδεχομένως ελέγχεται για καταγεγραμμένους συγχυτικούς παράγοντες. Δηλαδή, υπάρχει μια υπόθεση, που μπορεί ή μπορεί να μην καθιστά σαφές ότι η ανάθεση θεραπείας ανταποκρίνεται στο τυχαίοποιημένο μη συγχυτικό πείραμα. Ακόμη, είναι μια βασική υπόθεση η οποία δεν έχει επαληθευτεί, που εξαρτάται από το μέγεθος και τη φύση των καταγεγραμμένων συμμεταβλητών. Επίσης έχει ακόμη πιο βασικό ρόλο στα τυχαίοποιημένα πειράματα, διότι η προσαρμογή σε αυτά είναι απαραίτητη για να μειώσει ή να αποφύγει την μεροληψία.

Οι ‘καλές’ μελέτες παρατήρησης, όπως και τα ‘καλά’ πειράματα, σχεδιάζονται, δεν ανακαλύπτονται. Το έργο του *William Cochran* σε αυτό το θέμα πριν τα μέσα του 1970 είναι ιδιαίτερα σημαντικό. Όταν σχεδιάζεται ένα πείραμα, δεν έχουμε καμία ιδέα για τα δεδομένα αποτελέσματος, αλλά σχεδιάζουμε τη συλλογή, την οργάνωση και την ανάλυση των δεδομένων για να βελτιώσουμε τις πιθανότητές μας με την απόκτηση έγκυρων, αξιόπιστων και σε ακρίβεια αιτιωδών απαντήσεων. Η ίδια λογική, πρέπει να εφαρμοστεί σε μια μελέτη παρατήρησης. Σε μια μελέτη παρατήρησης συγκρίνοντας δύο θεραπείες, το βασικό βήμα είναι να συναρμολογήσουμε τα δεδομένα έτσι ώστε, μέσα από ένα *block*, η κατανομή των συμμεταβλητών για όσους ανήκουν στην ομάδα ελέγχου να είναι προσεγγιστικά ίδια με την ομάδα θεραπείας. Όταν τα *block* έχουν ακριβώς μια μονάδα ελέγχου και μια μονάδα θεραπείας, αυτό ανταποκρίνεται σε ένα προσαρμοσμένο ζευγάρι σχεδιασμού (*matched – pairs design*). Η χρήση των εκτιμώμενων βαθμολογιών τάσης ενδέχεται να είναι ιδιαίτερα χρήσιμη σε αυτή τη περίπτωση.

2.5.1 Μέθοδοι προσαρμογής για συγχυτικούς παράγοντες

Όταν ο μηχανισμός ανάθεσης μπορεί να θεωρηθεί συγχυτικός, μετά τον έλεγχο (*conditioning on*) των παρατηρήσιμων συγχυτικών παραγόντων, μπορούν να αποκτηθούν έγκυρες αιτιώδεις συμπερασματολογίες με τη προσαρμογή των εκτιμώμενων επιδράσεων θεραπείας για τις συμμεταβλητές X . Οι μέθοδοι για την προσπάθεια αυτής της επίτευξης περιλαμβάνουν τα ακόλουθα:

1. Παλινδρόμηση των αποτελεσμάτων των ψευδομεταβλητών για τις θεραπείες και για τις συμμεταβλητές, συμπεριλαμβανομένων των αλληλεπιδράσεων, αν αυτές θεωρούνται

σημαντικές. Αυτή η μέθοδος είναι πολύ κοινή στις επιδημιολογικές μελέτες. Μπορεί να είναι αρκετά μη αξιόπιστη, όταν οι ενεργές ομάδες θεραπείας και οι ομάδες ελέγχου διαφέρουν σημαντικά στις κατανομές των συμμεταβλητών, γιατί η παρεκβολή που συνεπάγεται από το μοντέλο παλινδρόμησης, στηρίζεται σε υποθέσεις όπως η γραμμικότητα.

2. Οι μέθοδοι που προσαρμόζονται για μια εκτίμηση της βαθμολογίας τάσης $p(W = 1|X)$, συνήθως λαμβάνονται από τη λογιστική παλινδρόμηση των δεικτών θεραπείας, στις συμμεταβλητές. Αυτές οι μέθοδοι περιλαμβάνουν τη προσθήκη της εκτιμώμενης βαθμολογίας τάσης σαν μια συμμεταβλητή σε ένα μοντέλο παλινδρόμησης, ή τη καλύτερη αντιστοίχιση, ή την υποδιαίρεση των μονάδων που λαμβάνουν την ενεργή θεραπεία και εκείνων που λαμβάνουν τη θεραπεία ελέγχου και βασίζονται στην εκτιμώμενη βαθμολογία τάσης. Ένα σημαντικό πλεονέκτημα αυτών των μεθόδων πέρα από την προσαρμογή παλινδρόμησης, είναι ότι ο ερευνητής μπορεί να διαπιστώσει ότι δεν υπάρχει αλληλοεπικάλυψη στις κατανομές των συμμεταβλητών, στις δυο ομάδες. Σε αυτή την περίπτωση, δεν υπάρχει ελπίδα εξαγωγής έγκυρων αιτιωδών συμπερασματολογιών από αυτά τα δεδομένα, χωρίς να κάνουμε επιπρόσθετες ισχυρές υποθέσεις. Σε τέτοιες περιπτώσεις, είναι σημαντικό να παραδεχτούμε ότι αυτό το σύνολο δεδομένων, δε μπορεί να στηρίξει μια ισχυρή αιτιώδη συμπερασματολογία.
3. Συνδυασμοί των μεθόδων των βαθμολογιών τάσης και της μοντελοποίησης παλινδρόμησης. Παραδείγματα αποτελούν η προσαρμογή ενός μοντέλου παλινδρόμησης σε κάθε υποσύνολο βαθμολογίας τάσης, και ο συνδυασμός αυτών των εκτιμήσεων, την προσαρμογή σε ένα μοντέλο παλινδρόμησης με προσαρμοσμένα ζευγάρια διαφορών.
4. Οι μέθοδοι βασίζονται στην υπόθεση της διαδοχικής ικανότητας αγνόησης που μπορεί να εφαρμοστεί σε *longitudinal* μελέτες παρατήρησης όταν αυτό κρίνεται απαραίτητο.

Κεφάλαιο 3

Δομικά μοντέλα εξισώσεων

Στο κεφάλαιο αυτό θα ασχοληθούμε με τα δομικά μοντέλα εξισώσεων (*Structural Equation Modeling, SEM*). Αποτελούν μία στατιστική τεχνική για τον έλεγχο και την εκτίμηση των αιτιωδών σχέσεων χρησιμοποιώντας ένα συνδυασμό στατιστικών δεδομένων και ποιοτικών αιτιωδών υποθέσεων. Ο ορισμός αυτός των *SEM* διατυπώθηκε αρχικά από τον γενετιστή *Sewall Wright* [64], και τους οικονομολόγους *Trygve Haavelmo* [20] και *Herbert A. Simon* [55]. Επίσημα ορίστηκε όμως από τον *Judea Pearl* [41] ο οποίος χρησιμοποίησε την λογική των *counterfactuals*.

3.1 Εισαγωγή

Οι ιδρυτές των δομικών μοντέλων εξισώσεων θεωρούσαν τα *SEM* ως ένα μαθηματικό εργαλείο για την εξαγωγή συμπερασμάτων από τον συνδυασμό παρατηρούμενων δεδομένων και θεωρητικών υποθέσεων. Ήταν κατηγορηματικοί για την αξία τους και ανένδοτοι για την ξεκάθαρη αιτιώδη ερμηνεία τους. Ωστόσο στη πορεία τα δομικά μοντέλα εξισώσεων, καθώς και το θεωρητικό υπόβαθρο στο οποίο στηρίζονται, αμφισβητήθηκαν από αρκετούς στατιστικολόγους. Αυτό συνέβη εν μέρη εξαιτίας της ραγδαίας ανάπτυξης των υπολογιστών, που έκανε τους ερευνητές των *SEM* να χάσουν την ικανότητά τους να διακρίνουν τη σχέση μεταξύ της θεωρίας και της απόδειξης. Όπως επίσης και της βαθμιαίας μείωσης της κατανόησης της λογικής των *SEM*.

Στη δημοσίευσή του για τα *SEM* ο *Freedman* [16] αμφισβητεί την αιτιώδη ερμηνεία τους και τα θεωρεί αντιφατικά. Είναι αρκετοί οι στατιστικολόγοι που παρεμέλησαν τη μελέτη των *SEM* και προχάλεσαν άλλους ερευνητές να τα θεωρήσουν ως μία κωδικοποίηση συνολικών παραμέτρων χωρίς αιτιώδη ερμηνεία. Ο *Muthen* [37] έγραψε ‘Θα ήταν πιο υγιές αν ακόμη περισσότεροι ερευνητές εγκατέλειπαν τη χρήση των όρων αιτία και επίδραση’. Ο *Sobel* [58] δήλωσε πως η ερμηνεία των παραμέτρων των *SEM* μοντέλων ως επιδράσεις δεν είναι ορθή ακόμη και αν το μοντέλο είναι σωστά προσδιορισμένο. Συγκρίνοντας τα μοντέλα δομικών εξισώσεων με τα ενδεχόμενα αποτελέσματα ο *Sobel* [57] ισχυρίστηκε πως ακόμη και στις

τυχαιοποιημένες μελέτες, οι δομικές και οι αιτιώδεις παράμετροι δεν είναι ίσοι, υπονοώντας ότι οι δομικές παράμετροι δε θα πρέπει να ερμηνεύονται ως επιδράσεις. Η τυπική ανάλυση ωστόσο αποδεικνύει με εντυπωσιακό τρόπο ακριβώς το αντίθετο: οι δομικές και οι αιτιώδεις επιδράσεις είναι ακριβώς το ίδιο και θα πρέπει πάντοτε να ερμηνεύονται ως επιδράσεις (*Galles and Pearl* [17]). Ο *Paul Holland* [27], ένας ακόμη ερευνητής του ενδεχόμενου αποτελέσματος ξεκαθαρίζει τη σύγχυση όσον αναφορά την εξίσωση $\{y = \alpha + bx + \epsilon\}$. Η μόνη έννοια που έχει δοθεί από τον *Pearl* μέχρι στιγμής είναι ένας σύντομος τρόπος για την περιγραφή της υπό συνθήκης κατονομής της μεταβλητής y δεδομένης της x . Θα διαπιστώσουμε ότι η δομική ερμηνεία αυτής της εξίσωσης δεν έχει σχέση με την υπό συνθήκη κατανομή της μεταβλητής y δεδομένης της x , αλλά μεταφέρει μία αιτιώδη πληροφορία. Τα δομικά μοντέλα εξισώσεων παρέχουν στη πραγματικότητα τη μαθηματική βάση από την οποία το πιθανό αποτέλεσμα αντλεί τη νομιμότητά του. Αυτό μαζί με τον αποτελεσματικό μηχανισμό τους, εξηγεί γιατί τα *SEM* διατηρούν την ιδιότητά τους ως τη πρώτη γλώσσα για την αιτιώδη και *counterfactual* ανάλυση.

3.2 Η λογική των *SEM*

Ο *Byrne* [6] περιγράφει τα δομικά μοντέλα εξισώσεων ως μια στατιστική μέθοδο η οποία απαιτεί μια προσέγγιση επιβεβαίωσης (υποθετικό έλεγχο) για την ανάλυση της δομικής θεωρίας. Οπότε το υποθετικό μοντέλο να μπορεί να ελεγχθεί στατιστικά με ταυτόχρονη ανάλυση όλων των μεταβλητών. Εάν ο έλεγχος καλής προσαρμογής είναι επαρκής, τότε το μοντέλο τάσσεται της εγκυρότητας των υποθετικών σχέσεων μεταξύ των μεταβλητών. Λαμβάνοντας κυριολεκτικά αυτή τη προσέγγιση επιβεβαίωσης αντιμετωπίζουμε κάποιες λογικές δυσκολίες. Ας θεωρήσουμε για παράδειγμα το υποθετικό μοντέλο:

$$M = \text{'Η Σταχτοπούτα είναι τρομοκράτης.'}$$

Μολονότι οποιοδήποτε μοντέλο καλής προσαρμογής θα αποτύγχανε να αποκαλύψει την αντίφαση σε αυτό το υποθετικό μοντέλο, θα ήταν περίεργο να υποστηρίξουμε πως κάτι τέτοιο είναι εφικτό. Οι απόπειρες για τη διόρθωση αυτής της αντίφασης επιμένοντας στην άποψη ότι το M είναι διαψεύσιμο, και ότι περιλαμβάνει μόνο μετρήσιμες μεταβλητές, δε δίνει λύση. Επιλέγοντας το υποθετικό μοντέλο:

$M = \text{'Οι μετρήσεις του βαρόμετρου προκαλούν βροχές και ο μέσος όρος ηλικίας στο Los Angeles είναι μεγαλύτερος από 3.'}$

θα αντιμετωπίζαμε την ίδια αντίφαση; Στη περίπτωση αυτή το M είναι επίσης διαψεύσιμο αλλά οι μεταβλητές του είναι μετρήσιμες. Η ικανότητα προσαρμογής των δεδομένων δεν μας δίνει καμία πληροφορία για τις αιτιώδεις σχέσεις μεταξύ της βροχής και του βαρόμετρου.

Ο μόνος τρόπος για την αποφυγή αυτού του παράδοξου είναι να επιμείνουμε πως το στοιχείο ελέγχου του M (ο μέσος όρος ηλικίας είναι μεγαλύτερος από 3), σχετίζεται λογικά με τους ισχυρισμούς (το βαρομετρικό προκαλεί βροχή). Κάτι τέτοιο όμως αντιβαίνει στη φιλοσοφία της προσέγγισης επιβεβαίωσης, σύμφωνα με την οποία το υποθετικό μοντέλο υποβάλλεται σε έναν έλεγχο όλων των μεταβλητών του συστήματος, ανεξάρτητα εάν το ελεγχόμενο μέρος του φέρει οποιαδήποτε σχέση με τους ισχυρισμούς που προκύπτουν.

Μέσα από αυτό το απλό παράδειγμα προκύπτει το βασικό μειονέκτημα αυτής της προσέγγισης βεβαίωσης. Υπογραμμίζεται επίσης το εμπειρικό περιεχόμενο των υποθέσεων που βασίζονται στο υποθετικό μοντέλο, οι ισχυρισμοί που απορρέουν από αυτό και ο βαθμός στον οποίο τεκμηριώνονται αυτοί οι ισχυρισμοί.

Σύμφωνα με αυτή την ερμηνεία, τα SEM είναι μια μέθοδος συμπερασματολογίας η οποία δέχεται τρία εισερχόμενα και παράγει τρία αποτελέσματα. Τα εισερχόμενα είναι:

E-1. Ένα σύνολο A ποιοτικών αιτιωδών υποθέσεων και ένα μοντέλο M_A που κωδικοποιεί αυτές τις υποθέσεις. (Το μοντέλο M_A παίρνει τη μορφή ενός διαγράμματος μονοπατιού ή τη μορφή ενός συνόλου δομικών εξισώσεων με ελεύθερες παράμετρους.)

E-2. Ένα σύνολο Q ερωτήσεων που αφορούν αιτιώδεις και *counterfactual* σχέσεις μεταξύ των μεταβλητών ενδιαφέροντος. Το Q αφορά τα μεγέθη των δομικών συντελεστών, αλλά σε γενικά μοντέλα καθορίζει τις αιτιώδεις σχέσεις πιο άμεσα, για παράδειγμα:

Q_1 : Ποια είναι η επίδραση της θεραπείας X στο αποτέλεσμα Y ;

Q_2 : Είναι αυτός ο εργοδότης υπόλογος για διάπραξη φυλετικής διάκρισης;

Θεωρητικά κάθε $Q_i \in Q$ πρέπει να υπολογίζεται από το πλήρως προσδιορισμένο μοντέλο M στο οποίο όλες οι συναρτησιακές σχέσεις είναι δεδομένες. Μη υπολογίσιμα ερωτήματα είναι μη αποδεκτά.

E-3. Ένα σύνολο D πειραματικών και μη δεδομένων, που διέπονται από μία από κοινού συνάρτηση πιθανότητας P παραγόμενη από μια διαδικασία σχετική με το σύνολο A .

Τα αποτελέσματα είναι:

A-1.: Ένα σύνολο A^* προτάσεων που είναι λογικά συμπεράσματα του συνόλου A , διαχωρισμένο από τα δεδομένα.

A-2.: Ένα σύνολο C ισχυρισμών σχετικά με τη σημασία των ερωτήσεων στο σύνολο Q , καθένα εξαρτώμενο από το A . Το C μπορεί λοιπόν να περιέχει το εκτιμημένο μέσο και τη διακύμανση μιας δομικής παράμετρου, ή την αναμενόμενη επίδραση μιας δεδομένης παρέμβασης. Τα *SEM* επίσης παράγουν ένα εκτιμητή $Q_i(P)$ για κάθε ερώτημα του συνόλου Q

A-3.: Μια κατάσταση ελέγξιμων στατιστικών συμπερασμάτων του A , και ο βαθμός $g(T_i), T_i$, στον οποίο τα δεδομένα βρίσκονται σε συμφωνία με κάθε ένα από τα συμπεράσματα. $\in T$, στον οποίο τα δεδομένα βρίσκονται σε συμφωνία με κάθε ένα από τα συμπεράσματα.

Τα δομικά μοντέλα εξισώσεων παρουσιάζονται μέσα από ένα στατιστικό μοντέλο όπου εξωγενείς μεταβλητές (επεξηγηματικές) μπορούν δυνητικά να επηρεάσουν ενδογενείς μεταβλητές (μεταβλητές απόκρισης) τόσο άμεσα όσο και έμμεσα μέσω παρεμβατικών μεταβλητών. Στο κεφάλαιο αυτό θα παρουσιάσουμε :

- Τα εργαλεία που θεωρούνται απαραίτητα για την ανάγνωση και την επεξήγηση αιτιωδών υποθέσεων που ενσωματώνονται στα δομικά μοντέλα εξισώσεων καθώς και τις υποθέσεις που στηρίζουν κάθε ανεξάρτητο αιτιώδη ισχυρισμό.
- Μεθόδους που προσδιορίζουν συνεπαγωγές των παραπάνω υποθέσεων, και τρόπους για την εκτίμηση όχι μόνο του μοντέλου, αλλά και επαγωγές των υποθέσεων που αναφέρονται σε κάθε ανεξάρτητο αιτιώδη ισχυρισμό.
- Μεθόδους που μας βοηθούν να αποφασίσουμε ποιες μετρήσεις πρέπει να ληφθούν υπόψη και ποιες μεροληπτούν εις βάρος των ποσοτήτων που επιθυμούμε.
- Μεθόδους για την εκπόνηση κρίσιμων στατιστικών ελέγχων σύμφωνα με τα οποίους δύο θεωρίες μπορούν να διακριθούν.
- Μεθόδους για τον μαθηματικό καθορισμό εάν οι αιτιώδεις σχέσεις μπορούν να εκτιμηθούν από τα δεδομένα και αν όχι ποιες επιπρόσθετες υποθέσεις, μετρήσεις ή πειραματικές μελέτες, τις καθιστούν εκτιμήσιμες.

- Μεθόδους αναγνώρισης και δημιουργίας ισοδύναμων μοντέλων που παγιώνουν, επεκτείνουν και τροποποιούν τις εμπειρικές μεθόδους του *Stelzl* [60] και των *Lee and Hershberger* [30].
- Γενίκευση των *SEM* σε κατηγορικά δεδομένα και μη γραμμικές αλληλεπιδράσεις, συμπεριλαμβανομένη της λύσης που καλείται πρόβλημα μεσολάβησης.

3.3 Η αιτιώδης ερμηνεία των δομικών μοντέλων εξισώσεων

Η πρώτη προσπάθεια για τη διατύπωση μιας μαθηματικής σχέσης του ισχυρισμού ότι ένα σύμπτωμα δεν προκαλεί την ασθένεια, πραγματοποιήθηκε το 1920 από το γενετιστή *Sewall Wright* [64]. Ο *Wright* χρησιμοποίησε το συνδυασμό εξισώσεων και γραφημάτων για την ερμηνεία των αιτιωδών σχέσεων. Ας θεωρήσουμε τη γραμμική δομική εξίσωση :

$$y = \beta x + u_Y, \quad x = u_X \quad (3.1)$$

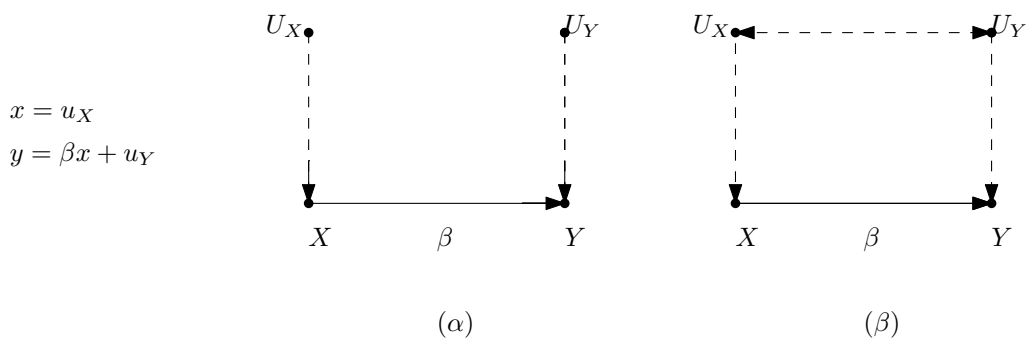
όπου το x συμβολίζει το επίπεδο της ασθένειας, το y συμβολίζει το επίπεδο του συμπτώματος και το u_Y συμβολίζει όλους τους άλλους παράγοντες εκτός από την ασθένεια, που θα μπορούσαν να επηρεάσουν το Y όταν το X διατηρείται σταθερό. Παρόμοια για να δηλώσουμε την ύπαρξη της ασθένειας X με την εξίσωση $x = u_X$, όπου η μεταβλητή u_X συμβολίζει τους παράγοντες που επηρεάζουν τη μεταβλητή X . Επισχύσουμε την εξίσωση κατασκευάζοντας ένα διάγραμμα μονοπατιού όπου τα βέλη σχεδιάζονται από τις αιτίες στις επιδράσεις τους, όπως στο Σχήμα 3.1. Η απουσία του βέλους που δείχνει από το Y στο X , παριστά τον ισχυρισμό ότι το σύμπτωμα Y δεν είναι ένας από τους παράγοντες U_X που επιδρούν στην ασθένεια X . Με την βοήθεια του διαγράμματος μπορούμε να συμπεράνουμε τη πιθανή ύπαρξη αιτιώδους επιρροής του X στο Y και την απουσία αιτιωδούς επιρροής από το Y στο X . Η παράμετρος β καλείται συντελεστής μονοπατιού και προσδιορίζει ποσοτικά την αιτιώδη επίδραση του X στο Y .

Οι μεταβλητές U_X και U_Y καλούνται εξωγενείς μεταβλητές και αντιπροσωπεύουν παράγοντες που επηρεάζουν αλλά δεν επηρεάζονται από άλλες μεταβλητές του μοντέλου που καλούνται ενδογενείς. Μία ενδογενής μεταβλητή είναι μία μεταβλητή που προσδιορίζεται μέσα στο σύστημα, ή με άλλα λόγια, προσδιορίζεται από κοινού με την y , δηλαδή μία μεταβλητή που υπόκειται σε αμφίδρομη αιτιότητα. Οι μη παρατηρήσιμες εξωγενείς μεταβλητές

συχνά καλούνται ‘διαταραχές’ ή ‘σφάλματα’. Στο παράδειγμά μας η μεταβλητή U_X εκπροσωπεί τους παράγοντες που επηρεάζουν την αθένεια X και πιθανόν ή όχι να συσχετίζεται με τη μεταβλητή U_Y . Οι εξωγενείς μεταβλητές για παράδειγμα μπορεί να περιγράφουν γενετικούς παράγοντες και κοινωνικές – οικονομικές καταστάσεις. Συμπεριφερόμαστε, στις εξωγενείς μεταβλητές όπως σε όλες τις άλλες του μοντέλου, αν και δε μπορούμε να εκτιμήσουμε ακριβώς τις τιμές τους και έτσι δεχόμαστε απλώς την ύπαρξή τους αξιολογώντας ποιοτικά τη σύνδεσή τους με τις υπόλοιπες.

Στην περίπτωση που η συσχέτιση μεταξύ των μεταβλητών θεωρείται πιθανή, τότε συνηθίζεται να τις συνδέουμε με ένα διπλό διακεκομμένο βέλος, όπως τις U_Y και U_X στο Σχήμα 3.1(β). Εστιάζοντας στις αιτιώδεις σχέσεις μεταξύ των παρατηρήσιμων μεταβλητών θεωρείται σημαντικό να διακρίνουμε τα συσχετιζόμενα και μη σφάλματα. Όπως αναφέραμε και προηγουμένως τα σφάλματα είναι μη παρατηρήσιμες εξωγενείς μεταβλητές. Στην περίπτωση που οι όροι σφαλμάτων δεν συσχετίζονται είναι ευκολότερο να απαλειφθούν από το διάγραμμα.

Για την μελέτη των διαγραμμάτων μονοπατιών χρησιμοποιούμε λέξεις συγγενικών σχέσεων, όπως γονέας, τέκνο, πρόγονος και απόγονος. Για παράδειγμα το βέλος $X \rightarrow Y$, ορίζει το X ως γονέα του Y και το Y ως τέκνο του X . Ένα μονοπάτι είναι μία διαδοχική σειρά ακμών, συνεχών ή διακεκομμένων. Στο Σχήμα 3.1(β) μπορούμε να παρατηρήσουμε πως υπάρχουν δύο μονοπάτια μεταξύ του X και του Y , το πρώτο δίδεται απευθείας με το βέλος $X \rightarrow Y$, ενώ το δεύτερο ακολουθώντας τους κόμβους X, U_X, U_Y και Y .



Σχήμα 3.1: Παράδειγμα ενός απλού δομικού μοντέλου εξισώσεων με τα συσχετιζόμενα διαγράμματα που παρουσιάζουν (α) τις ανεξάρτητες μη παρατηρήσιμες εξωγενείς μεταβλητές (συνδέονται με ένα διακεκομμένο βέλος), (β) τις εξαρτημένες εξωγενείς μεταβλητές.

Η μεγαλύτερη συνεισφορά του *Wright* στην αιτιώδη ανάλυση, όπως είδαμε και στο Κεφάλαιο 1, έπειτα από τη θέσπιση της γλώσσας των διαγραμμάτων μονοπατιών, υπήρξε

η ανάπτυξη των γραφικών κανόνων για τη συνδιακύμανση οποιοδήποτε ζεύγους παρατηρηθέντων μεταβλητών σε όρους συντελεστών διαγραμμάτων και της συνδιακύμανσης που υφίσταται ανάμεσα στους όρους σφαλμάτων. Στο παράδειγμά μας ισχύουν κατά συνέπεια οι σχέσεις

$$Cov(X, Y) = \beta$$

για το Σχήμα 3.1(α)

$$Cov(X, Y) = \beta + Cov(U_Y, U_X)$$

για το Σχήμα 3.1(β)

Η παρουσία ενός βέλους ανάμεσα σε δύο μεταβλητές υποδηλώνει την ύπαρξη αιτιώδους σχέσης, της οποίας η ισχύς αναμένεται να προσδιοριστεί από τα δεδομένα. Η απουσία βέλους μεταξύ των μεταβλητών δηλώνει τον ισχυρισμό μηδενικής επιρροής, ενώ η απουσία ενός διπλού βέλους δηλώνει τον ισχυρισμό της μηδενικής συνδιασποράς. Στο παράδειγμα 3.1(α), οι υποθέσεις που μας επιτρέπουν να προσδιορίσουμε την άμεση επίδραση β , κωδικοποιούνται μέσω της απουσίας του διπλού βέλους μεταξύ των μεταβλητών U_Y , U_X . Μαζί με την απουσία του βέλους από το Y στο X , υποδεικνύουν ότι $Cov(U_Y, U_X) = 0$. Στη περίπτωση όπου υπήρχε το διπλό βέλος ανάμεσα στις μεταβλητές U_Y , U_X , ή το βέλος από το Y στο X , δεν θα είχαμε τη δυνατότητα να προσδιορίσουμε την άμεση επίδραση β . Η πρόσθεση αυτών των αιτιωδών σχέσεων θα ισοδυναμούσε με την ασθενέστερη υπόθεση $Cov(U_Y, U_X) = 0$, ή την ασθενέστερη υπόθεση ότι το Y δεν επηρεάζει το X .

Αξίζει να επισημάνουμε πως και οι δύο υποθέσεις είναι αιτιώδεις, οι οποίες δε μπορούν να εξεταστούν μεμονωμένα, σε αντίθεση με το σύνολο όλων των αιτιωδών υποθέσεων του μοντέλου που έχουν αυτή τη δυνατότητα. Για παράδειγμα στο Σχήμα 3.2(α), κωδικοποιούνται επτά αιτιώδεις υποθέσεις, όπου η κάθε μία αντιστοιχεί σε ένα βέλος, ή σε ένα διπλό βέλος μεταξύ ενός ζεύγους μεταβλητών, που λείπει. Καμία από αυτές τις υποθέσεις δε μπορεί να ελεγχθεί μεμονωμένα, ωστόσο από το σύνολό τους συνεπάγεται πως η μεταβλητή Z δεν συσχετίζεται με τη Y σε καμία περίπτωση μέσω της X . Τέτοια συμπεράσματα μπορούν να συναχθούν από τα διαγράμματα μονοπατιών χρησιμοποιώντας ένα γραφικό κριτήριο γνωστό ως d -διαχωρισμό το οποίο θα παρουσιάσουμε στη πορεία.

3.3.1 Ντετερμινιστικό αιτιώδες μοντέλο

Ορισμός 1. (Αιτιώδες μοντέλο, ([41, p. 203]) Ένα αιτιώδες μοντέλο δίδεται από το σύνολο

$$M = \langle \mathbf{U}, \mathbf{V}, \mathbf{F} \rangle$$

όπου:

1. το \mathbf{U} είναι ένα σύνολο *background* μεταβλητών (μεταβλητές που θεωρούνται αμελητέες, συχνά καλούνται εξωγενείς), που προσδιορίζονται από παράγοντες εκτός του μοντέλου,
2. το \mathbf{V} είναι ένα σύνολο $\{V_1, V_2, \dots, V_n\}$ μεταβλητών που καλούνται ενδογενείς, και ορίζονται από μεταβλητές του μοντέλου – δηλαδή μεταβλητές από το σύνολο $\mathbf{U} \cup \mathbf{V}$, και
3. το \mathbf{F} σύνολο συναρτήσεων $\{f_1, f_2, \dots, f_n\}$ τέτοιο ώστε κάθε συνάρτηση f_i να είναι μια απεικόνιση από το $\mathbf{U} \cup (\mathbf{V} \setminus V_i)$, τέτοιο ώστε ολόκληρο το σύνολο \mathbf{F} να αποτελεί μια απεικόνιση από το \mathbf{U} στο \mathbf{V} . Δηλαδή, κάθε συνάρτηση f_i μας δίνει την τιμή του συνόλου V_i , δεδομένων όλων των τιμών όλων των υπολοίπων μεταβλητών του συνόλου $\mathbf{U} \cup \mathbf{V}$, οπότε ολόκληρο το σύνολο \mathbf{F} έχει μοναδική λύση $\mathbf{V}(\mathbf{u})$. Για λόγους συμβολισμού το σύνολο των εξισώσεων \mathbf{F} μπορεί να παρασταθεί από τη σχέση:

$$v_i := f_i(\text{ρα}(v_i), u_i), \quad i = 1, \dots, n,$$

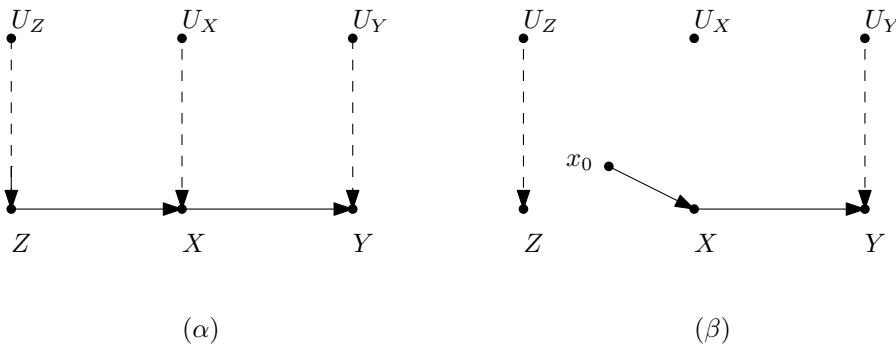
όπου ο όρος $\text{ρα}(v_i)$ είναι η πραγματοποίηση του ελάχιστου συνόλου $PA(V_i)$ στο σύνολο $\mathbf{V} \setminus V_i$ που επαρκεί για την αντιπροσώπευση του f_i . Ομοίως το σύνολο $\mathbf{U}_i \subseteq \mathbf{U}$ συμβολίζει το ελάχιστο σύνολο των μεταβλητών στο \mathbf{U} που επαρκεί για την αντιπροσώπευση του f_i .

3.3.2 Αιτιώδεις υποθέσεις σε μη παραμετρικά μοντέλα

Τα δομικά μοντέλα εξισώσεων έχουν υπάρξει το κύριο μέσο για την ανάλυση των επιδράσεων στις Οικονομικές και Κοινωνικές Επιστήμες. Ωστόσο το μεγαλύτερο μέρος της μεθοδολογίας των *SEM* αναπτύχθηκε για την γραμμική ανάλυση. Μέχρι πρόσφατα δεν έχει επινοηθεί καμία άλλη μέθοδος για την επέκταση των δυνατοτήτων της γραμμικής ανάλυσης, σε μοντέλα που περιλαμβάνουν διχοτομημένες μεταβλητές και μη γραμμικές εξαρτήσεις. Για μια

τέτοια επέκταση απαιτείται να διαχωριστεί η έννοια της επίδρασης από την αλγεβρική της παράσταση ως συντελεστή σε μία εξίσωση και να την ορίσουμε ξανά προσθέτοντάς της την κύρια ιδιότητα να διαβιβάζει τις μεταβολές μεταξύ των μεταβλητών.

Μια τέτοια επέκταση έχει ως κεντρική ιδέα την αξιοποίηση των αναλλοίωτων χαρακτηριστικών των δομικών εξισώσεων χωρίς τη δέσμευση κάποιας συγκεκριμένης συναρτησιακής μορφής. Για το λόγο αυτό θα μελετήσουμε τη **μη παραμετρική ερμηνεία** του διαγράμματος του Σχήματος 3.2(α) το οποίο αντιστοιχεί σε ένα σύνολο τριών συναρτήσεων όπου κάθε μία αντιστοιχεί σε μία από τις παρατηρήσιμες μεταβλητές:



Σχήμα 3.2: Το διάγραμμα (α) αναφέρεται στο δομικό μοντέλο της εξίσωσης (3.2) και το (β) στο τροποποιημένο μοντέλο της εξίσωσης (3.3) παρουσιάζοντας την παρέμβαση $do(X = x_0)$.

$$\begin{aligned}
 z &= f_Z(u_Z) \\
 x &= f_X(z, u_X) \\
 y &= f_Y(x, u_Y)
 \end{aligned}
 \tag{3.2}$$

όπου η z είναι μια ντετερμινιστική συνάρτηση της u_Z , η x μια ντετερμινιστική συνάρτηση των z και u_X , και η y μια ντετερμινιστική συνάρτηση των x και u_Y . Ολόκληρο το μοντέλο καλείται M .

Το μοντέλο M είναι μη-παραμετρικό (με την έννοια ότι περιλαμβάνει τεχνικές που δεν υποθέτουν ότι το μοντέλο έχει σταθερή δομή), οπότε δεν γίνεται καμία υπόθεση για τις f_Z , f_X , f_Y , U_Z , U_X και U_Y . Σημειώνεται ότι οποιαδήποτε εξάρτηση μεταξύ των εξωγενών μεταβλητών u , πρέπει να διαμορφωθεί με σαφή τρόπο με επιπρόσθετες εξισώσεις (ή πιθανώς μέσω του επαναπροσδιορισμού των u μεταβλητών). Για παράδειγμα, μπορούμε να γράψουμε το προηγούμενο μοντέλο περιλαμβάνοντας τις πιθανές εξαρτήσεις μεταξύ των U_Z και U_X :

$$\begin{aligned}
 u_Z &= f_{u_Z}(u_{ZX}) \\
 u_X &= f_{u_X}(u_{ZX}) \\
 z &= f_Z(u_Z) \\
 x &= f_X(z, u_X) \\
 y &= f_Y(x, u_Y)
 \end{aligned}$$

όμως προκειμένου να διατηρηθούν οι μεταβλητές u ως εξωγενείς απλώς γράφουμε το μοντέλο που ακολουθεί:

$$\begin{aligned}
 z &= f_Z(u_{ZX}) \\
 x &= f_X(z, u_{ZX}) \\
 y &= f_Y(x, u_Y)
 \end{aligned}$$

στο οποίο οι ορισμοί των f_Z και f_X έχουν αλλάξει και το γεγονός ότι έχει συμπεριληφθεί η u_{ZX} και στις δυο, δείχνει ένα συγκεκριμένο τρόπο εξάρτησης μεταξύ αυτών των μεταβλητών. Η απουσία μιας επιπλέον μεταβλητής και μια διακριτή μεταβλητή u στις συναρτήσεις f_Z , f_X και f_Y συνεπάγεται μια μορφή δεσμευμένης ανεξαρτησίας.

Επιστρέφοντας στο αρχικό μοντέλο των μεταβλητών (3.2) μπορούμε λοιπόν να πούμε ότι οι U_Z , U_X και U_Y θεωρούνται από κοινού ανεξάρτητες και τυχαία κατανοημένες. Κάθε μία από τις συναρτήσεις αποτελεί έναν αιτιώδη μηχανισμό που προσδιορίζει τη τιμή της αριστερής μεταβλητής με τη βοήθεια των δεξιών μεταβλητών. Η απουσία της μεταβλητής Z από τις παραμέτρους της f_Y δίδει τον ισχυρισμό, πως η μεταβολή της μεταβλητής Z δεν επηρεάζει τη μεταβλητή Y , όσο οι μεταβλητές U_Y και X παραμένουν σταθερές. Ένα σύστημα αποτελούμενο από τέτοιου είδους συναρτήσεις θεωρείται ότι είναι δομικό εάν οι συναρτήσεις αυτές θεωρηθούν ότι λειτουργούν αυτόνομα, δηλαδή κάθε συνάρτηση παραμένει αναλλοίωτη σε πιθανές αλλαγές της μορφής των υπολοίπων συναρτήσεων.

3.3.3 Παρουσιάζοντας τις παρεμβάσεις και τις αιτιώδεις επιδράσεις

Ας συμβολίσουμε με U το σύνολο των εξωγενών μεταβλητών $\{U_X, U_Y, U_Z\}$ και με u τη πραγματοποίηση της U . Για λόγους απλότητας, υποθέτουμε ότι το μοντέλο είναι αναδρομικό

(δηλαδή δεν περιλαμβάνει αιτιώδεις κύκλους, μόνο μιας κατεύθυνσης επιδράσεις) στο οποίο όλες οι ενδογενείς μεταβλητές είναι σε θέση να οριστούν από μια συνάρτηση των εξωγενών μεταβλητών. Για παράδειγμα το μοντέλο των μεταβλητών (3.2), η μεταβλητή y μπορεί να γραφεί ως ως συνάρτηση όλων των υπολοίπων μεταβλητών:

$$\begin{aligned} z &= f_Z(u_Z) \\ x &= f_X(z, u_X) \\ y &= f_Y(x, u_Y) \\ y &= f_Y(f_X(f_Z(u_Z), u_X), u_Y) \end{aligned}$$

Ως εκ τούτου, η $Y_M(u)$ δηλώνει τη μοναδική τιμή της Y που παράγεται από το μοντέλο M δεδομένου του u

Ας θεωρήσουμε τώρα τη παρέμβαση σε ένα σύστημα προκείμενου να καθορίσουμε τη τιμή X ίση με μια συγκεκριμένη τιμή x , χωρίς την άμεση διαταραχή των υπολοίπων μεταβλητών του συστήματος. Αυτό συνεπάγεται τη δημιουργία ενός υπομοντέλου στο οποίο η εξίσωση της X έχει αφαιρεθεί.

Ορισμός 2. (Αιτιώδες μοντέλο, ([41, p. 204]) Υποθέτουμε ότι το M είναι ένα αιτιώδες μοντέλο, το \mathbf{X} το σύνολο των μεταβλητών στο \mathbf{V} και η x μια συγκεκριμένη πραγματοποίηση του \mathbf{X} . Το υπομοντέλο M_x του M είναι το αιτιώδες μοντέλο

$$M_x = \langle \mathbf{U}, \mathbf{V}, F_x \rangle$$

όπου

$$\mathbf{F}_x = \{f_i : V_i \notin \mathbf{X}\} \cup \{\mathbf{X} = x\}$$

Συμβολίζουμε αυτή τη παρέμβαση με $do(x)$ και το αποτέλεσμα της είναι η δημιουργία ενός νέου συνόλου δομικών εξισώσεων στο οποίο η εξίσωση του x λαμβάνει κάποια τιμή. Τα αναλλοίωτα χαρακτηριστικά των δομικών εξισώσεων μας επιτρέπουν να χρησιμοποιήσουμε τις δομικές εξισώσεις ως βάση για τη μοντελοποίηση των αιτιωδών επιδράσεων και των *counterfactuals*. Όπως αναφέραμε και προηγουμένως, αυτό πραγματοποιείται με τη βοήθεια του μαθηματικού τελεστή $do(x)$, ο οποίος προσομοιώνει τις φυσικές παρεμβάσεις με τη διαγραφή ορισμένων συναρτήσεων από το μοντέλο, τις οποίες αντικαθιστά με μία σταθερά $X = x$, καθώς διατηρεί το υπόλοιπο μοντέλο αμετάβλητο. Για παράδειγμα, για την

αναπαράσταση της παρέμβασης $do(x)$ που διατηρεί το X σταθερό, όπως στο μοντέλο M του Σχήματος 3.2(α), αντικαθιστούμε την εξίσωση x και λαμβάνουμε το νέο μοντέλο M_{x_0} :

$$\begin{aligned} z &= f_Z(u_Z) \\ x &= x_0 \\ y &= f_Y(x, u_Y) \end{aligned} \tag{3.3}$$

η γραφική περιγραφή του οποίου παρουσιάζεται στο Σχήμα 3.2(β).

Στα πλαίσια ενός υπομοντέλου μπορούμε να ορίσουμε τα ενδεχόμενα αποτελέσματα με ανάλογο τρόπο όπως και στο RCM με την επίλυση ενός συνόλου εξισώσεων για τη μεταβλητή Y βάσει του υπομοντέλου.

Ορισμός 3. (Ενδεχόμενα αποτελέσματα ([41, p. 204])» Θεωρούμε τα \mathbf{X} και \mathbf{Y} δυο υποσύνολα μεταβλητών του \mathbf{V} . Η ενδεχόμενη απόκριση της \mathbf{Y} από τη παρέμβαση $do(\mathbf{X} = x)$, συμβολίζεται με $\mathbf{Y}_x(u)$ είναι η λύση της \mathbf{Y} του συνόλου των εξισώσεων F_x .

Ως εκ τούτου συμβολίζουμε με $Y_x(u)$ τη μοναδική τιμή της Y που απορρέει από μια συγκεκριμένη πραγματοποίηση της τιμής u του U σε ένα μοντέλο M_x . Αν η u είναι παρατηρήσιμη τότε το μοντέλο M και το υπομοντέλο M_x θα ορίζουν αιτιώδεις επιδράσεις μοναδιαίου επιπέδου, ανάλογες των αιτιωδών επιδράσεων του RCM . Θεωρούμε το ακόλουθο μη παραμετρικό παράδειγμα.

Υποθέτουμε ότι μια ιατρική μελέτη αναθέτει τυχαία δυο θεραπείες σε ασθενείς που έχουν υποστεί εγκεφαλικό επεισόδιο, στη θεραπεία ελέγχου (γενόσημο) και την ενεργή θεραπεία (φάρμακο αραιώσης αίματος). Ύστερα από μια χρονική περίοδο παρατηρεί εάν οι ασθενείς είναι ζωντανοί ή όχι. Υπάρχουν εγκεφαλικά δυο τύπων, πληροφορία άγνωστη στον ερευνητή. Κάποιοι από αυτούς υποφέρουν από θρόμβωση και κάποιοι από αιμορραγία. Επομένως η χορήγηση του φαρμάκου για τη αραιώση αίματος θα έχει διαφορετικές επιδράσεις στις δυο ομάδες ασθενών. Αν συμβολίσουμε με X την ανάθεση θεραπείας, Y τη κατάσταση του ασθενή, με U_1 να αποτελεί τον εξωγενή μηχανισμό τυχαιοποιημένης θεραπείας και η U_2 την εξωγενή μεταβλητή που περιγράφει την ενδεχόμενη απόκριση κάθε ασθενή στη θεραπεία. Τότε το παραπάνω σενάριο μπορεί να αποτυπωθεί στο πλαίσιο ενός μη παραμετρικού μοντέλου SEM με τις ακόλουθες συναρτήσεις:

$$\begin{aligned} x &= f_X(u_1) \\ y &= f_Y(x, u_2) \end{aligned}$$

Έχοντας μια δυαδική μεταβλητή θεραπείας και αποτελέσματος το πεδίο οριμού της U_2 (οπότε και των ασθενών με εκγεφαλικό επεισόδιο) μπορεί να διαχωριστεί σε τέσσερις ομάδες που περιγράφουν τα ενδεχόμενα αποτελέσματα της θεραπείας: Σε εκείνους που η κατάστασή τους θεωρείται ήπια και που θα είχαν ζήσει ανεξάρτητα από το ποια θεραπεία λαμβάνουν (*always – takers*). Σε εκείνους που η κατάστασή τους είναι τόσο σημαντική που δε καταφέρνουν να ζήσουν ανεξάρτητα με το ποια θεραπεία τους ανατίθεται (*never – takers*). Σε εκείνους με τη θεραπεύσιμη κατάσταση θρόμβωσης που θα μείνουν ζωντανοί με την αραιώση αίματος αλλά δε θα ζήσουν χωρίς αυτή (*compliers*). Και τέλος, εκείνους με την ήπια κατάσταση αιμορραγίας που θα πεθάνουν με την αραιώση αίματος αλλά θα ζήσουν με το γενόσημο φάρμακο (*non – compliers*).

Στο παράδειγμα αυτό, όπως και στα περισσότερα, η u δεν παρατηρείται με αποτέλεσμα να μην παρατηρούνται και οι αιτιώδεις επιδράσεις μοναδιαίου επιπέδου. Η λύση είναι να γίνει αλλαγή της έμφασης συμπερασματολογίας στις πληθυσμιακές αιτιώδεις επιδράσεις, όπως και στο *RCM*.

Η από κοινού συνάρτηση κατανομής που συσχετίζεται με το τροποποιημένο μοντέλο, συμβολίζεται με $P(z, y|do(x_0))$ και περιγράφει τη μετά τη παρέμβαση κατανομή των μεταβλητών Y και Z . Η συνάρτηση κατανομής πριν τη παρέμβαση είναι η $P(x, y, z)$ και συσχετίζεται με το μοντέλο της εξίσωσης (3.2). Οι δυο κατανομές διαφέρουν μεταξύ τους. Για παράδειγμα, αν το X εκπροσωπεί μία μεταβλητή θεραπείας, το Y αποτελεί μια μεταβλητή απόκρισης και το Z μία συμμεταβλητή που επηρεάζει τη ποσότητα της λαμβάνουσας θεραπείας, τότε η κατανομή $P(z, y|do(x_0))$ δίνει το ποσοστό των ατόμων που ανταποκρίθηκαν στη θεραπεία, στο επίπεδο απόκρισης $Y = y$ και στο επίπεδο της συμμεταβλητής $Z = z$, υπό την υποθετική κατάσταση στην οποία η θεραπεία $X = x_0$ χορηγείται ομοιόμορφα στο πληθυσμό.

Γενικά είναι δυνατόν να ορίσουμε τη μετά τη παρέμβαση κατανομή σύμφωνα με την εξίσωση

$$P_M(y|do(x)) = P_{M_x}(y) \quad (3.4)$$

ειδικότερα όσο αναφορά το μοντέλο M , η μετά την παρέμβαση κατανομή του αποτελέσματος Y ορίζεται ως η πιθανότητα το μοντέλο M_x να αποδίδει σε κάθε αποτέλεσμα το επίπεδο $Y = y$. Μέσω της κατανομής αυτής είμαστε σε θέση να εκτιμήσουμε μία αποτελεσματική θεραπεία, έπειτα από τη σύγκριση αυτής της κατανομής στα διαφορετικά επίπεδα του x_0 . Ένα κοινό μέτρο σύγκρισης της αποτελεσματικότητας είναι η διαφορά των μέσων όρων

$$E(Y|do(x'_0)) - E(Y|do(x_0))$$

όπου οι μεταβλητές x'_0 και x_0 αποτελούν δύο τύπους θεραπείας που έχουν επιλεγεί για σύγκριση. Ένα ακόμη μέτρο σύγκρισης είναι η αναλογία κινδύνου

$$E(Y|do(x'_0))/E(Y|do(x_0))$$

Το κεντρικό ζήτημα στην ανάλυση των αιτιωδών επιδράσεων είναι το πρόβλημα της αναγνώρισης: Υπάρχει η δυνατότητα η μετά την παρέμβαση κατανομή $P(Y = y|do(x))$ να εκτιμηθεί από τα δεδομένα που διέπουν τη πριν τη παρέμβαση κατανομή $P(z, x, y)$; Στα παραμετρικά γραμμικά μοντέλα το πρόβλημα του προσδιορισμού μειώνει το ερώτημα για το αν κάποια παράμετρος του μοντέλου, έστω β , έχει μοναδική λύση σε όρους των παραμέτρων της P . Ενώ στα μη παραμετρικά μοντέλα η εκφράση ‘υπάρχει μοναδική λύση’, δεν έχει άμεση εφαρμογή στις αιτιώδεις ποσότητες όπως η $Q(M) = P(y|do(x))$, εφόσον δεν ορίζονται παραμετρικά αλλά διαδικαστικά, προσομοιώνοντας μία παρέμβαση σε ένα αιτιώδες μοντέλο M , όπως στην εξίσωση (3.4). Ο επόμενος Ορισμός συμβάλλει στο να ξεπεραστούν τέτοιου είδους προβλήματα.

Ορισμός 4. (Αναγνωρισιμότητα, ([41, p. 204])

Μία ποσότητα $Q(M)$ καθίσταται αναγνωρίσιμη σε ένα δοσμένο σύνολο υποθέσεων A , εάν για δύο οποιαδήποτε μοντέλα M_1 και M_2 που ικανοποιούν το A ισχύει

$$P(M_1) = P(M_2) \Rightarrow Q(M_1) = Q(M_2) \quad (3.5)$$

Αναλυτικότερα θα λέγαμε πως τα χαρακτηριστικά των M_1 και M_2 δεν παίζουν τόσο σημαντικό ρόλο, όσο οι υποθέσεις του συνόλου A , οι οποίες περιορίζουν τη μεταβλητότητα αυτών των χαρακτηριστικών με τέτοιο τρόπο ούτως ώστε η ισότητα των P να συνεπάγεται την ισότητα των Q . Όταν κάτι τέτοιο ισχύει, τότε το Q εξαρτάται μόνο από το P και εκφράζεται σε όρους του P .

3.3.4 Πληθυσμιακό αιτιώδες μοντέλο

Παράγουμε ένα πληθυσμιακό μοντέλο για το ντετερμινιστικό αιτιώδες μοντέλο της υποενότητας 3.2.1 θεωρώντας μια κατανομή στο U .

Ορισμός 5. (Πιθανολογικό αιτιώδες μοντέλο ([41, p. 205])» Ένα πιθανολογικό αιτιώδες μοντέλο είναι το σύνολο

$$\langle M, P(u) \rangle$$

όπου το M αποτελεί το αιτιώδες μοντέλο και η $P(u)$ είναι μια συνάρτηση πιθανότητας ορισμένη στο πεδίο ορισμού του U .

Το μοντέλο M μαζί με την υπόθεση ως προς τη κατανομή U παράγει την μετά τη παρέμβαση κατανομή $P(U, V)$. Δεδομένης της υπόθεσης ενός αναδρομικού μοντέλου αυτή η κοινή κατανομή ορίζεται μοναδικά. Είμαστε σε θέση να εκτιμήσουμε τη περιθώρια κατανομή των παρατηρήσιμων μεταβλητών στο V , άμεσα από μια μελέτη παρατήρησης, χωρίς την κατασκευή μη ελέγξιμων καταστάσεων. Συνεχίζοντας το παράδειγμα με τους ασθενείς των εγκεφαλικών επεισοδίων, της προηγούμενης ενότητας, ο διακριτός χαρακτήρας του συνόλου U επιτρέπει την ερμηνεία της $P(u)$ ως πληθυσμιακά ποσοστά των ατόμων. Ως εκ τούτου, η $P(u)$ περιγράφει τα ποσοστά των (*always-takers*, *never-takers*, *compliers*, *non-compliers*) ατόμων του πληθυσμού. Ακόμη, αυτές οι κατανομές σε συνδυασμό με το αιτιώδες μοντέλο ορίζουν τα ποσοστά των (ζωντανών, νεκρών) ατόμων του πληθυσμού.

Αν υποθέσουμε ότι η $P(u)$ παραμένει αμετάβλητη για το υπομοντέλο M_x , τότε μπορούμε να αναρωτηθούμε ποια είναι η πιθανότητα κατανομής των $Y_x(u)$ για την τυχαία εξαγωγή του u από τη πληθυσμιακή κατανομή της U . Με άλλα λόγια, ποια είναι η κατανομή της U του πληθυσμού έπειτα από τη παρέμβαση $do(x)$. Η ποσότητα αυτή συμβολίζεται με $P(y|do(x))$ και καλείται μετά τη παρέμβαση κατανομή της U . Η μετά τη παρέμβαση κατανομή δεν μπορεί να εκτιμηθεί άμεσα από τα παρατηρήσιμα δεδομένα χωρίς την ύπαρξη μη ελέγξιμων αιτιωδών επιδράσεων. Με τις κατανομές πιθανότητας που ορίζονται στις μετά τη παρέμβαση κατανομές, είναι εύκολο να παραχθούν οι αναμενόμενες τιμές και οι μέσες αιτιώδεις επιδράσεις:

$$E[Y|do(x)] \equiv E[Y_x]$$

και

$$ACE \equiv E[Y_1] - E[Y_0] = E[Y|do(1)] - E[Y|do(0)]$$

Παρόμοια με το *RCM* θέλουμε να δημιουργήσουμε τις συνθήκες στις οποίες η παρατηρήσιμη πριν τη παρέμβαση κατανομή προσδιορίζει τους μέσους όρους των μη παρατηρήσιμων

μετά τη παρέμβαση κατανομή. Αυτό πραγματοποιείται με τη παρουσίαση των μη παραμετρικών *SEM* ως *DAG* (*Directed Acyclic Graphs*)

3.3.5 Τα μη παραμετρικά *SEM* ως *DAGs*

Ένα γράφημα $G = \{\mathbf{V}, \mathbf{E}\}$ είναι ένα σύνολο κορυφών $V \in \mathbf{V}$ και ακμών $E \in \mathbf{E}$. Κάθε ακμή ενώνεται με δύο κορυφές. Κάθε ακμή θεωρείται άμεση με την έννοια ότι αντιπροσωπεύει μια ασύμμετρη σχέση (για κάθε a και b που ανήκουν X , αν το a σχετίζεται με το b τότε το b δεν σχετίζεται με το a) μεταξύ των κορυφών που ενώνει. Μια άμεση ακμή από την κορυφή V_1 στη V_2 συμβολίζεται ως $V_1 \rightarrow V_2$. Σε ότι ακολουθεί, θεωρούμε τις κορυφές ως μεταβλητές και τις ακμές ως τις αιτιώδεις σχέσεις μεταξύ τους. Ένα μονοπάτι είναι μια διαδοχική ακολουθία κορυφών E_1, E_2, \dots, E_k , στο οποίο το τέλος της κορυφής E_i είναι η αρχή της κορυφής E_{i+1} $i = 1, \dots, (k - 1)$. Η διεύθυνση των ακμών δεν έχει καμία σημασία. Για παράδειγμα, ένα μονοπάτι από τη V_1 στη V_3 υπάρχει σε κάθε ένα από τα παρακάτω μονοπάτια.

$$V_1 \rightarrow V_2 \rightarrow V_3 \quad (3.6)$$

$$V_1 \leftarrow V_2 \rightarrow V_3 \quad (3.7)$$

$$V_1 \rightarrow V_2 \leftarrow V_3 \quad (3.8)$$

$$V_1 \leftarrow V_2 \leftarrow V_3 \quad (3.9)$$

Ένα κατευθυνόμενο μονοπάτι μεταξύ των ακμών από το V_1 στο V_k προϋποθέτει ότι όλες οι ακμές δείχνουν προς τη V_k κατά μήκος του μονοπατιού. Για παράδειγμα το μονοπάτι (3.6) απεικονίζει ένα κατευθυνόμενο μονοπάτι από τη V_1 στη V_3 . Ένας κόμβος σύγκρουσης είναι μία κορυφή για την οποία δύο ακμές σε ένα μονοπάτι, δείχνουν και οι δυο σε αυτήν. Για παράδειγμα στο μονοπάτι (3.8) η V_2 αποτελεί ένα κόμβο σύγκρουσης. Όμως στις (3.6), (3.7), (3.9) δεν αποτελεί κόμβο σύγκρουσης.

Οι γονείς της V_i συμβολίζονται με $pa(V_i)$, και όπως έχουμε ξαναπεί, είναι το σύνολο των κορυφών από τις οποίες προέρχονται τα βέλη που δείχνουν άμεσα στη V_i . Οι απόγονοι της V_i συμβολίζονται με $an(V_i)$ και είναι το σύνολο των κορυφών από τις οποίες υπάρχει ένα κατευθυνόμενο μονοπάτι στη V_i . Παρόμοια τα παιδιά της V_i συμβολίζονται με $ch(V_i)$ και είναι το σύνολο των κορυφών στο οποίο ένα βέλος δείχνει από τη V_i . Τέλος οι απόγονοι συμβολίζονται με $de(V_i)$ και είναι το σύνολο των κορυφών στο οποίο ένα κατευθυνόμενο μονοπάτι υπάρχει από τη V_i .

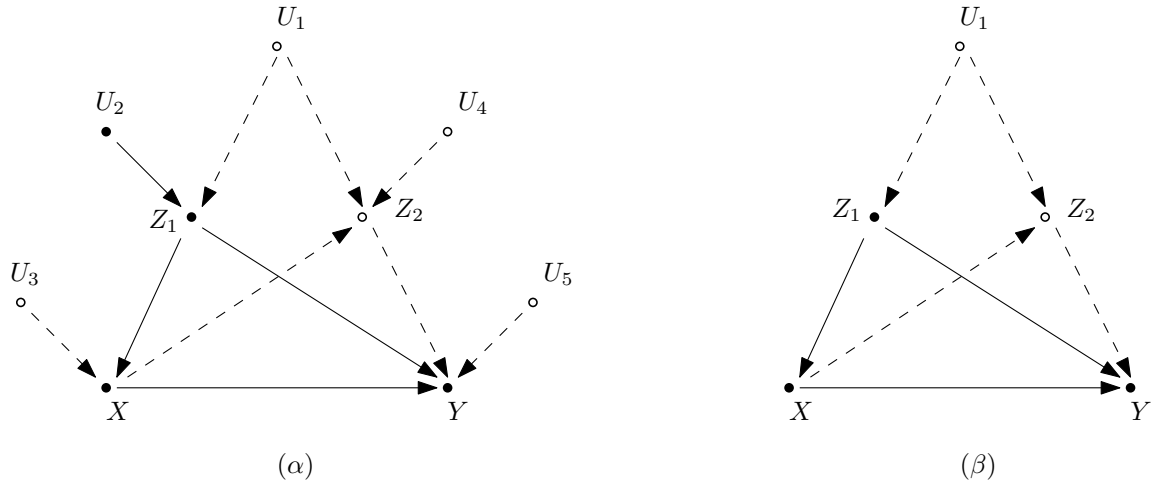
Ένα γράφημα στο οποίο όλες οι ακμές είναι κατευθυνόμενες (δηλαδή είναι βέλη με μία κεφαλή) και καμία ακμή της μορφής $V_i \rightarrow V_i$ δεν υφίσταται, λέμε ότι είναι *κατευθυνόμενο γράφημα*. Ένα κατευθυνόμενο γράφημα που δεν έχει κύκλους (ενδεχομένως με τη συμμετοχή μεγάλων σε μήκος μονοπατιών) της μορφής $V_i \rightarrow V_j, V_i \leftarrow V_j$, λέγεται *κατευθυνόμενο άκυκλο γράφημα (DAG)*.

Χρησιμοποιούμε τις ακμές σε ένα *DAG* προκειμένου να παραστήσουμε τα εισερχόμενα των συναρτήσεων του αντίστοιχου μη παραμετρικού *SEM*. Σε αυτό το τύπο αιτιώδους μοντέλου, δεδομένου του $U = u$, μπορούμε να προσδιορίσουμε μοναδικά τις τιμές όλων των υπολοίπων μεταβλητών. Επιπροσθέτως, εάν καθορίσουμε μια κατανομή πιθανότητας του U , τότε η από κοινού κατανομή των παρατηρήσιμων μεταβλητών μπορούν επίσης να προσδιοριστούν μοναδικά. Οι κανόνες που διέπουν ένα γράφημα *DAG* G_M από ένα μη παραμετρικό *SEM* M είναι οι ακόλουθοι:

1. Να παριστάνουν κάθε μη παρατηρήσιμη μεταβλητή με μια ‘ανοιχτή’ κορυφή.
2. Να παριστάνουν κάθε παρατηρήσιμη μεταβλητή με μια ‘κλειστή’ κορυφή.
3. Για κάθε εξίσωση M , να σχεδιάζεται μια ακμή από κάθε μεταβλητή του δεξιού μέλους της εξίσωσης στη μεταβλητή του αριστερού μέλους, χρησιμοποιώντας μια συνεχή γραμμή όταν και οι δύο κορυφές είναι παρατηρήσιμες και μια διακεκομμένη γραμμή όταν η κορυφή είναι μη παρατηρήσιμη.

Μπορούμε να διαγράψουμε τους κόμβους και τις ακμές οι οποίες δεν επηρεάζουν το αποτέλεσμα των γραφικών ελέγχων. Για παράδειγμα, οι εξωγενείς μεταβλητές που δείχνουν σε μια μόνο ενδογενή μεταβλητή, μπορούν να αφαιρεθούν. Ωστόσο κάτι τέτοιο θα μπορούσε να δημιουργήσει ασαφείς υποθέσεις. Ειδικότερα, μια εξωγενής μεταβλητή που δείχνει σε μια μόνο ενδογενή μεταβλητή υποδηλώνει μια μορφή ανεξαρτησίας μεταξύ των εξωγενών μεταβλητών. Επίσης, η αφαίρεση των εξωγενών μεταβλητών από το γράφημα μπορεί να επισκιάσει το γεγονός ότι τα μη παραμετρικά *SEM* προσδιορίζουν ανεξάρτητες αιτιώδεις επιδράσεις.

Ένα απλό παράδειγμα καθιστά αυτή τη διαδικασία πιο ξεκάθαρη. Θεωρούμε το επόμενο δομικό μοντέλο M



Σχήμα 3.3: Το γράφημα (α) παρουσιάζει όλες τις εξωγενείς μεταβλητές και τις συσχετιζόμενες ακμές (β) οι περιττές εξωγενείς μεταβλητές έχουν αφαιρεθεί. Σημειώνεται ότι η δυνατότητα παρατήρησης των μεταβλητών είναι ούτε απαραίτητη ούτε επαρκής για να είναι μια μεταβλητή εξωγενής. Στο γράφημά μας οι U_1, \dots, U_5 είναι εξωγενείς μεταβλητές. Η U_2 είναι παρατηρήσιμη αλλά όλες οι άλλες U μεταβλητές δεν είναι. Επίσης η ενδογενής μεταβλητή Z_2 είναι μη παρατηρήσιμη ενώ οι υπόλοιπες ενδογενείς παρατηρούνται.

$$z_1 = f_{Z_1}(u_1, u_2) \tag{3.10}$$

$$x = f_X(z_1, u_3) \tag{3.11}$$

$$z_2 = f_{Z_2}(x, u_1, u_4) \tag{3.12}$$

$$y = f_Y(x, z_1, z_2, u_5) \tag{3.13}$$

Στο Σχήμα 3.3(α), το γράφημα G_M είναι κατασκευασμένο χρησιμοποιώντας του κανόνες 1-3. Το Σχήμα 3.3(β) παρουσιάζει το μοντέλο G_M όπου οι περιττές ακμές και κορυφές για την αναγνώριση της επίδρασης της X στη Y έχουν αφαιρεθεί.

3.3.6 Counterfactual ανάλυση στα δομικά μοντέλα εξισώσεων

Προκειμένου να αξιολογήσουμε τις επιπτώσεις μιας παρέμβασης, όπως παρουσιάστηκε πιο πριν, χρησιμοποιούμε την counterfactual ανάλυση. Ουσιαστικά η counterfactual ανάλυση είναι ένα υποθετικό σενάριο-κατάσταση, που εφαρμόζεται ούτως ώστε να συγκρίνουμε τι πραγματικά έχει συμβεί (με τη παρέμβαση), με το τι θα είχε συμβεί αν δεν είχε γίνει η

παρέμβαση. Ο Pearl [43] ορίζει τα *counterfactual* απευθείας σε όρους δομικών μοντέλων εξισώσεων. Ένα σύνολο εξισώσεων στο οποίο κάθε μεταβλητή είναι συνάρτηση των υπολοίπων μεταβλητών του συστήματος. Έχοντας ένα μοντέλο τέτοιου τύπου, η πρόταση ‘το Y θα είναι y δεδομένου ότι το X ήταν x όταν $U = u$ ’, ορίζεται ως τον ισχυρισμό: αν αντικαταστήσουμε την εξίσωση ορίζοντας $X = x$ και λύσουμε το σύνολο των εξισώσεων για τη μεταβλητή Y , θα προκύψει $Y = y$ όταν $U = u$. Συμβολίζουμε με

$$Y_x(u) = y$$

Η βασική ιδέα είναι να ερμηνεύσουμε τη φράση ‘όταν το X ήταν x ’ ως υπόδειξη για τον ελάχιστο μετασχηματισμό στο μοντέλο M , ο οποίος να είχε ενδεχομένως αναθέσει στο X μια διαφορετική τιμή έστω $X = x'$. Ο ελάχιστος αυτό μετασχηματισμός ισοδυναμεί με το να αντικαταστήσουμε την εξίσωση του X με τη σταθερά x , όπως ενηργήσαμε και στην εξίσωση (3.3). Η αντικατάσταση αυτή επιτρέπει στη σταθερά x να διαφέρει από τη πραγματική τιμή της μεταβλητής X (συγκεκριμένα την $f_X(z, u_X)$) χωρίς να καταστεί το σύστημα των εξισώσεων ασυμβίβαστο. Με αυτό τον τρόπο αποδίδεται μια επίσημη ερμηνεία των *counterfactuals* σε πολυσταδιακά μοντέλα όπου αν μία μεταβλητή είναι εξαρτημένη σε μία εξίσωση, η ίδια μεταβλητή είναι ανεξάρτητη σε μία άλλη.

Ορισμός 6. (Μοναδιαίου επιπέδου *Counterfactuals*, ([41, p. 98])

Ας υποθέσουμε πως το M είναι ένα πλήρως καθορισμένο δομικό μοντέλο και το M_x μία τροποποιημένη μορφή του μοντέλου M με την εξίσωση του X να έχει αντικατασταθεί με τη σχέση $X = x$ και η λύση της μεταβλητής Y στις εξισώσεις του M_x συμβολίζεται με $Y_{M_x}(u)$. Τότε το *counterfactual* $Y_x(u)$ δίδεται από τη σχέση

$$Y_x(u) = Y_{M_x}(u) \tag{3.14}$$

Παρατηρούμε ότι το μοναδιαίου επιπέδου *counterfactual* $Y_x(u)$, το οποίο στο *RCM* αντιμετωπίζεται ως προσδιορίσιμη ποσότητα είναι στη πραγματικότητα μια ποσότητα που μπορεί να υπολογιστεί στο πλαίσιο των δομικών μοντέλων εξισώσεων. Θεωρούμε τη λύση του τροποποιημένου μοντέλου M_{x_0} της εξίσωσης (3.3), της οποίας ο Ορισμός 6 δίνει το συμβολισμό $Y_{x_0}(u_X, u_Y, u_Z)$. Η ποσότητα αυτή έχει καθαρά *counterfactual* ερμηνεία διότι δείχνει τον τρόπο με τον οποίο ένα άτομο με χαρακτηριστικά (u_X, u_Y, u_Z) , θα είχε ανταποκριθεί αν η θεραπεία ήταν η x_0 , έναντι της θεραπείας $x = f_X(z, u_X)$ που στη πραγματικότητα έλαβε. Στο παράδειγμά μας, εφόσον η μεταβλητή Y δεν εξαρτάται από τις u_X

και u_Z μπορούμε να γράψουμε:

$$Y_{x_0}(u) = Y_{x_0}(u_Y, u_X, u_Z) = f_Y(x_0, u_Y)$$

με τον ίδιο τρόπο μπορούμε να λάβουμε

$$\begin{aligned} Y_{z_0} &= f_Y(f_X(z_0, u_X), u_Y) \\ X_{z_0, y_0}(u) &= f_X(z_0, u_X) \end{aligned}$$

και ούτω καθεξής. Μέσα από τα παραδείγματα αυτά αποκαλύπτεται η *counterfactual* ερμηνεία για κάθε ανεξάρτητη δομική εξίσωση του μοντέλου της εξίσωσης (3.2). Η εξίσωση $x = f_X(z, u_X)$ δίνει τον εμπειρικό σχεδιασμό ότι ανεξάρτητα με τις τιμές που έχουν ληφθεί από τις υπόλοιπες μεταβλητές του συστήματος, αν το Z ήταν z_0 , το X δεν θα είχε λάβει καμία άλλη τιμή εκτός από $x = f_X(z_0, u_X)$.

Είναι ξεκάθαρο πως η πιθανότητα $P(u_Y, u_X, u_Z)$ επάγει μία καλώς ορισμένη πιθανότητα για την *counterfactual* κατάσταση $Y_{x_0} = y$ όπως και για τις *counterfactual* καταστάσεις $Y_{x_0} = y$ και $Y_{x_1} = y'$. Κατά συνέπεια για να απαντήσουμε σε ερωτήσεις του τύπου για το αν το Y θα είναι y_1 στη περίπτωση που το X είναι x_0 , χρειάζεται να υπολογίσουμε την υπό συνθήκη πιθανότητα $P(Y_{x_1} = y_1 | Y = y_0, X = x_0)$. Μελετώντας τις γραμμικές εξισώσεις που περιγράφονται στο Σχήμα 3.1

$$x = u_X, \quad y = \beta x + u_Y$$

οι υπό συνθήκη καταστάσεις $Y = y_0$ και $X = x_0$, δίδουν $U_X = x_0$ και $U_Y = y_0 - \beta x_0$, και μπορούμε να συμπεράνουμε με πιθανότητα 1 ότι, η μεταβλητή Y_{x_1} πρέπει να πάρει την τιμή $Y_{x_1} = \beta x_1 + U_Y = \beta(x_1 - x_0) + y_0$. Με άλλα λόγια αν η μεταβλητή X ήταν x_1 αντί για x_0 , η μεταβλητή Y θα αύξανε β φορές τη διαφορά $(x_1 - x_0)$.

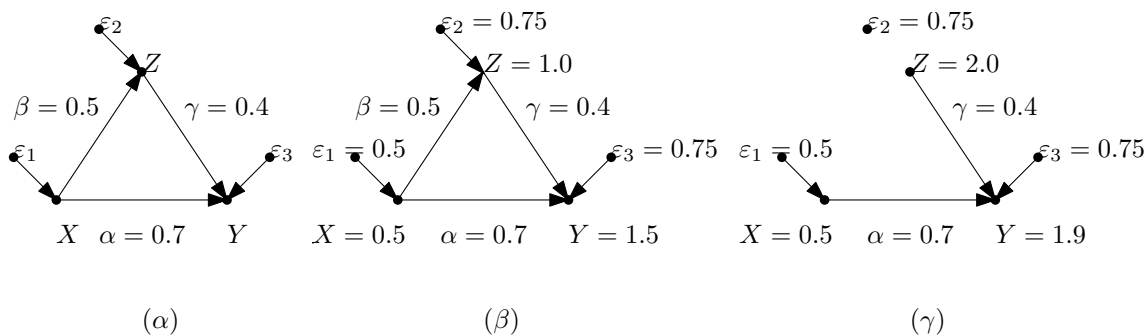
Παράδειγμα ενός *counterfactual*:

Οι δομικές εξισώσεις διαθέτουν την ικανότητα να κωδικοποιούν και να μεταφέρουν *counterfactual* πληροφορίες σε ατομικό επίπεδο αλλά και σε επίπεδο πληθυσμού. Με τη βοήθεια του παρακάτω παραδείγματος θα παρουσιάσουμε αυτή τη δυνατότητα, χρησιμοποιώντας τρεις γραμμικές μεταβλητές. Ας υποθέσουμε το μοντέλο του Σχήματος 3.4, όπου

- η μεταβλητή X συμβολίζει το βαθμό των μαθημάτων που ένας μαθητής έχει παρακολουθήσει,
- η μεταβλητή Z το ποσοστό του χρόνου που ο μαθητής αφιερώνει στο διάβασμα,
- και η μεταβλητή Y το αποτέλεσμα, δηλαδή την επίδοση του μαθητή στο διαγώνισμα.

Ξεκινώντας από ατομικό επίπεδο ανάλυσης και θεωρώντας ότι ο μαθητής ονομάζεται Ευάγγελος για τον οποίον έχουν γίνει οι μετρήσεις $X = 0.5, Z = 1, Y = 1.5$ και για τον οποίο κάνουμε την *counterfactual* ερώτηση:

Q_1 : Ποιό θα ήταν το αποτέλεσμα του Ευάγγελου στις εξετάσεις εάν είχε διπλασιάσει το χρόνο διαβάσματός του;



Σχήμα 3.4: Δομικά μοντέλα που χρησιμοποιούνται για να απαντήσουν στην *counterfactual* ερώτηση για τα ανεξάρτητα δεδομένα $u = (\epsilon_1, \epsilon_2, \epsilon_3)$ (α) Το γενικό μοντέλο (β) Το ειδικό μοντέλο u και (γ) το τροποποιημένο μοντέλο προκειμένου να ισχύει $Z = 2$ της *counterfactual* ερώτησης Q_1 .

Η παραπάνω ερώτηση ισοδυναμεί με την εκτίμηση της $Y_{Z=2}(u)$, με τη μεταβλητή u να αποτελεί τη μεταβλητή που περιγράφει τα διακριτά χαρακτηριστικά του Ευάγγελου, δηλαδή $u = (\epsilon_1, \epsilon_2, \epsilon_3)$, όπως συμπεραίνονται από τα παρατηρήσιμα δεδομένα ($X = 0.5, Z = 1, Y = 1.5$). Η απάντηση του προβλήματος αυτού δίδεται σε τρία βήματα :

- Χρησιμοποιούμε τα δεδομένα μας για να υπολογίσουμε τους εξωγενείς παράγοντες $\epsilon_1, \epsilon_2, \epsilon_3$. Οι παράγοντες αυτοί αποτελούν αναλλοίωτα χαρακτηριστικά του επιπέδου u και δεν μεταβάλλονται από παρεμβάσεις ή *counterfactual* υποθέσεις. Έτσι για το μοντέλο του Σχήματος 3.4(β) λαμβάνουμε

$$\epsilon_1 = 0.5$$

$$\epsilon_2 = 1 - 0.5 \times 0.5 = 0.75$$

$$\epsilon_3 = 1.5 - 0.5 \times 0.7 - 1 \times 0.4 = 0.75$$

- Τροποποιούμε το μοντέλο στη μορφή $M_{Z=2}$, όπου η μεταβλητή Z τίθεται ίση με τη τιμή 2 και έχουν αφαιρεθεί όλα τα βέλη προς τη Z .
- Υπολογίζουμε τη τιμή της μεταβλητής Y στο τροποποιημένο μοντέλο όπως έχει διαμορφωθεί στο δεύτερο βήμα, οπότε

$$Y_{Z=2} = 0.5 \times 0.7 + 2.0 \times 0.4 + 0.75 = 1.90$$

Το παράδειγμα αυτό παρουσιάζει την αναγκαιότητα του να τροποποιήσουμε το αρχικό μοντέλο στο οποίο ο συνδυασμός ($X = 1, \epsilon_2 = 0.75, Z = 2.0$) αποτελεί αντίφαση. Σε αυτό το σημείο ας κάνουμε μία υποθετική ερώτηση σχετικά με τον Ευάγγελο.

Q_2 : Ποιο θα ήταν το αποτέλεσμα του Ευάγγελου αν η παρακολούθηση μαθημάτων ισούταν με το μηδέν, και είχε μελετήσει στο ίδιο επίπεδο με το αν είχε παρακολουθήσει μαθήματα επιπέδου 1;

Η παραπάνω περίπλοκη ερώτηση, είναι η βάση για να ορίσουμε τη παρέμβαση. Η ζητούμενη ποσότητα δίδεται από την Y_{0,Z_1} , όπου η μεταβλητή Z_1 είναι η τιμή της Z που θα είχε επιτευχθεί αν το X ήταν ένα. Για να υπολογίσουμε αυτή τη ποσότητα χρειάζεται να ορίσουμε δύο τροποποιημένα μοντέλα. Το πρώτο παρουσιάζεται στο Σχήμα 3.5(α) και μας βοηθάει να υπολογίσουμε τη μεταβλητή Z_1

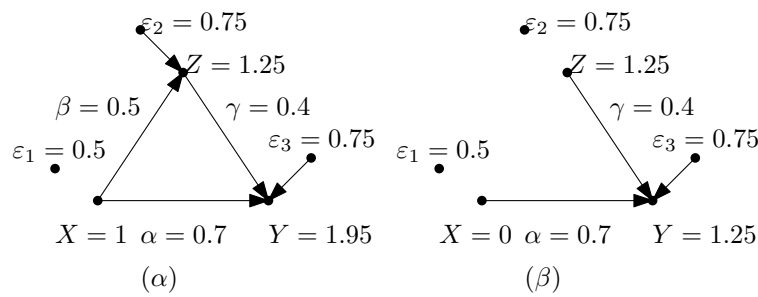
$$Z_1 = 1.0 \times 0.5 + 0.75 = 1.25$$

Ενώ το δεύτερο παρουσιάζεται στο Σχήμα 3.5(β) μας βοηθάει να υπολογίσουμε την Y_{0,Z_1} και έτσι να δώσουμε μία απάντηση στην ερώτηση Q_2 :

$$Y_{0,Z_1} = Y_{0,1.25} = 1.25 \times 0.4 + 0.75 = 1.25$$

Εάν συγκρίναμε το παραπάνω αποτέλεσμα με το αποτέλεσμα του Ευάγγελου στην περίπτωση που δεν έχει παρακολουθήσει κανένα μαθήμα, δηλαδή $Y_0 = 0.75 \times 0.4 + 0.75 = 1.05$, θα λαμβάναμε την αναμενόμενη διαφορά της έμμεσης επίδρασης της X στην Y που ισούται με $Y_{0,Z_1} - Y_0 = 0.20 = \beta \times \gamma$.

Το παράδειγμα αυτό ενδεχομένως να φαίνεται περιττό εφόσον μπορούμε να υπολογίσουμε την ποσότητα που επιθυμούμε κατευθείαν από το γινόμενο $\beta \times \gamma$. Το πλεονέκτημα των *counterfactual* θα φανεί στην επόμενη ενότητα όπου οι έμμεσες επιδράσεις θα οριστούν για διακριτές μεταβλητές και θα υπολογίζονται από τα δεδομένα χωρίς να υποθέτουμε παραμετρικές μορφές συναρτήσεων.



Σχήμα 3.5: Δομικά μοντέλα ειδικής μονάδας που χρησιμοποιούνται για την απάντηση της *counterfactual* ερώτησης Q_2 σχετικά με την έμμεση επίδραση της X στην Y . (α) Το τροποποιημένο μοντέλο που απαιτείται για τον υπολογισμό του Z_1 , (β) το τροποποιημένο μοντέλο που απαιτείται για τον υπολογισμό του Y_{0,Z_1} .

Προβλέποντας τα αποτελέσματα και τα ενδεχόμενα αποτελέσματα σε εμπειρικές μελέτες

Λαμβάνοντας σαν δεδομένο πως κάθε *counterfactual* ερώτηση μπορεί να απαντηθεί μέσω της (3.14), από ένα πλήρως καθορισμένο δομικό μοντέλο, μεταφερόμαστε σε πληθυσμιακό επίπεδο, θέτοντας μια ερώτηση σε ένα σύνολο 10 ατόμων, με τον Ευάγγελο να αποτελεί τον συμμετέχοντα 1. Κάθε ένας συμμετέχων έχει ένα διακριτό διάνυσμα χαρακτηριστικών $u_i = (\epsilon_1, \epsilon_2, \epsilon_3)$ όπως φαίνεται και από τη πρώτη στήλη του πίνακα 3.1. Για την τριάδα χαρακτηριστικών $(\epsilon_1, \epsilon_2, \epsilon_3)$, το μοντέλο του Σχήματος 3.4(α) μας δίνει τη δυνατότητα να συμπληρώσουμε πλήρως μια γραμμή, συμπεριλαμβανομένων των Y_0 και Y_1 , τα οποία συμβολίζουν τα ενδεχόμενα αποτελέσματα υπό τον έλεγχο ($X = 0$) και τη θεραπεία ($X = 1$) αντίστοιχα. Παρατηρούμε πως ένα απλό δομικό μοντέλο όπως αυτό του Σχήματος 3.4(α), κωδικοποιεί την επίδραση σε ένα σύνθετο πληθυσμό ατόμων μαζί με τη προβλεπόμενη συμπεριφορά τους υπό τις συνθήκες μελέτης αλλά και τις πειραματικές συνθήκες. Οι

Συμμετέχοντες	Χαρακτηριστικά συμμετεχόντων			Παρατηρήσιμα αποτελέσματα			Ενδεχόμενα αποτελέσματα πρόβλεψης				
	ϵ_1	ϵ_2	ϵ_3	X	Y	Z	Y_0	Y_1	Z_0	Z_1	Z_{00}
1	0.5	0.75	0.75	0.5	1.50	1.0	1.05	1.95	0.75	1.25	0.75
2	0.3	0.1	0.4	0.3	0.71	0.25	0.44	1.34	0.1	0.6	0.4
3	0.5	0.9	0.2	0.5	1.01	1.15	0.56	1.46	0.9	1.4	0.2
4	0.6	0.5	0.3	0.6	1.04	0.8	0.50	1.40	0.5	1.0	0.3
5	0.5	0.8	0.9	0.5	1.67	1.05	1.22	2.12	0.8	1.3	0.9
6	0.7	0.9	0.3	0.7	1.29	1.25	0.66	1.56	0.9	1.4	0.3
7	0.2	0.3	0.8	0.2	1.10	0.4	0.92	1.82	0.3	0.8	0.8
8	0.4	0.6	0.2	0.4	0.80	0.8	0.44	1.34	0.6	1.1	0.2
9	0.6	0.4	0.3	0.6	1.00	0.7	0.46	1.36	0.4	0.9	0.3
10	0.3	0.8	0.3	0.3	0.89	0.95	0.62	1.52	0.8	1.3	0.3

Πίνακας 3.1: Ενδεχόμενα και παρατηρήσιμα αποτελέσματα που προβλέπονται βάσει του δομικού μοντέλου του Σχήματος 3.4(α). Οι μονάδες επιλέχθηκαν τυχαία με κάθε ϵ_i να είναι ομοιόμορφα κατανομημένο στο διάστημα $[0,1]$

στήλες X, Y, Z προβλέπουν τα αποτελέσματα των μελετών παρατήρησης και οι στήλες Y_0, Y_1, Z_0, Z_1 προβλέπουν το υποθετικό αποτέλεσμα υπό τις δύο θεραπείες $X = 0$ και $X = 1$. Στη πραγματικότητα μπορούν να προβλεφθούν άπειρα αποτελέσματα, όπως για παράδειγμα το $Y_{X=0.5, Z_1=2.0}$, το οποίο υπολογίζετε από το Σχήμα 3.4(γ), αλλά και όλοι οι συνδυασμοί των μεταβλητών που έχουν συζητηθεί. Από τον σύνθετο πληθυσμό μπορεί κανείς να υπολογίσει την κατανομή οποιασδήποτε *counterfactual* ερώτησης για τις μεταβλητές X, Y, Z , συμπεριλαμβανομένων των εκ των υστέρων *counterfactuals*. Παράδειγμα αποτελεί η πιθανότητα που ένα άτομο το οποίο επιλέγεται τυχαία, θα είχε πετύχει στο διαγώνισμα έχοντας βοήθεια δεδομένου ότι στη πραγματικότητα, αυτό το άτομο θα είχε αποτύχει εάν δεν είχε λάβει βοήθεια. Η πιθανότητα αυτή εκφράζεται ως $P(Y_1|X = 0, Y = 0)$ και είναι γνωστή ως ‘πιθανότητα αιτίας’ [41], και επίσης ποσοτικοποιεί τις ‘αιτίες της επίδρασης’ σε αντίθεση με την ‘επίδραση των αιτιών’. Η τελευταία αφαιρείται από το σύνολο της ενδεχόμενης ανάλυσης αποτελεσμάτων.

Η πρόβλεψη αυτή διευκολύνεται φυσικώς με τη βοήθεια δυο μη ελέγξιμων πληροφοριών: (1) τη δομή του μοντέλου (η οποία περιλαμβάνει την υπόθεση των ανεξάρτητων όρων σφαλμάτων) και (2) τις τιμές των παραμέτρων του μοντέλου (που περιλαμβάνουν τη κατανομή κάθε μιας εξωγενούς μεταβλητής). Εφόσον η δεύτερη πληροφορία μπορεί εύκολα να συμ-

περαθεί από τα δεδομένα (βλέπε ενότητα 3.3), η δεύτερη εξαρτάται σε πολύ μεγάλο βαθμό από την επιστημονική κρίση.

Συμμετέχοντες	Ενδεχόμενα αποτελέσματα πρόβλεψης		Παρατηρήσιμα αποτελέσματα	
	Y_0	Y_1	Y_0	Y_1
1	1.05	1.95	1.05	■
2	0.44	1.34	■	1.34
3	0.56	1.46	■	1.46
4	0.50	1.40	■	1.40
5	1.22	2.12	1.22	■
6	0.66	1.56	0.66	■
7	0.92	1.82	■	1.82
8	0.44	1.34	0.44	■
9	0.46	1.36	■	1.36
10	0.62	1.52	0.62	■

Πραγματικός μέσος όρος επίδρασης: 0.90
Μέσος όρος επίδρασης έπειτα από μελέτη: 0.68

Πίνακας 3.2: Ενδεχόμενα και παρατηρήσιμα αποτελέσματα σε μια τυχαιοποιημένη κλινική μελέτη, όπου στη μεταβλητή X δίνονται τυχαία οι τιμές $X = 0$ και $X = 1$

Υποθέτουμε τώρα πως δεν έχουμε καμία απολύτως πληροφορία σχετικά με το υποκείμενο μοντέλο και το μόνο που γνωρίζουμε είναι οι μετρήσεις της μεταβλητής Y που έχουν ληφθεί από πειραματική μελέτη στην οποία η μεταβλητή X τυχαιοποιείται σε δυο επίπεδα $X = 0$ και $X = 1$.

Ο Πίνακας 3.2 δίνει τις αποκρίσεις των ίδιων 10 συμμετεχόντων από αυτές τις πειραματικές συνθήκες. Οι πρώτες δυο στήλες δίδουν το πραγματικό ενδεχόμενο αποτέλεσμα (το οποίο έχει ληφθεί από το πίνακα 3.1) ενώ οι δυο τελευταίες περιγράφουν τις πληροφορίες που είναι διαθέσιμες στον πειραματιστή. Το τεράγωνο αντί τιμών στο πίνακα υποδεικνύει ότι η απόκριση δεν μπορούσε να παρατηρηθεί. Η τυχαιοποίηση μας βεβαιώνει ότι, αν και τα μισά από τα ενδεχόμενα αποτελέσματα δεν παρατηρούνται, η διαφορά μεταξύ των παρατηρήσιμων μέσων όρων των ομάδων θεραπείας και ελέγχου, τείνει στο πραγματικό μέσο όρο της διαφοράς $E(Y_1 - Y_0) = 0.9$.

Στο μοντέλο μας από τη στιγμή που οι εξωγενείς μεταβλητές είναι ανεξάρτητες, η κλίση της παλινδρόμησης του Y στο X επίσης θα τείνει στο μέσο όρο της αιτιώδους επίδρασης.

Η μεροληψία θα εμφανιστεί αν η ϵ_1 συσχετίζεται είτε με την ϵ_2 είτε με την ϵ_3 . Ωστόσο τέτοια συσχέτιση δεν θα μεροληπτεί της αιτιώδους επίδρασης που εκτιμούμε στη πειραματική μελέτη.

3.3.7 Σύνδεση με το *framework* του ενδεχόμενου αποτελέσματος

Ο Ορισμός 6 αποτελεί τη ‘γέφυρα’ μεταξύ των *SEM* και το πλαίσιο που καλείται ενδεχόμενο αποτέλεσμα [44], το οποίο συχνά παρουσιάζεται και ως μια ‘περισσότερο εναλλακτική αρχή’ προς τα *SEM* [24][49][57][58]. Τέτοιες παρουσιάσεις είναι παραπλανητικές και αποτελούν παραπληροφόρηση. Τα δύο πλαίσια έχουν αποδειχτεί ότι είναι λογικώς ίσα και διαφέρουν μόνο στη γλώσσα με την οποία οι ερευνητές τους επιτρέπεται να εκφράζουν τις υποθέσεις. Το ένα Θεώρημα στο ένα πλαίσιο είναι Θεώρημα στο άλλο [41, p. 228-231], με τον Ορισμό 6 να αποτελεί βάση και για τα δυο.

Η ιδέα για το ενδεχόμενο αποτέλεσμα είναι απλή. Οι ερευνητές που αισθάνονται άβουλα να παρουσιάσουν τις υποθέσεις τους με διαγράμματα ή δομικές εξισώσεις, μπορούν να χρησιμοποιήσουν τις τυχαίοποιημένες δοκιμές ως μοντέλο απόφασης και ερμηνεύοντας την *counterfactual* $Y_x(u)$ ως ενδεχόμενο αποτέλεσμα της μονάδας u στην υποθετική θεραπεία $X = x$, αγνοώντας του μηχανισμούς που ισχύουν για το αποτέλεσμα. Το πρόβλημα της αιτιώδους συμπερασματολογίας θεωρείται ως ένα με ελλιπή δεδομένα, όπου τα ελλιπή δεδομένα είναι τα ενδεχόμενα αποτελέσματα $Y_x(u)$ υπό την θεραπεία που δεν λαμβάνεται. Καθώς τα παρατηρήσιμα δεδομένα είναι ενδεχόμενα αποτελέσματα όπως παρουσιάζονται στο Σχήμα 3.2.

Έτσι, η Y_x αποκτά το ρόλο της νέας λανθάνουσας μεταβλητής που αποκαλύπτει τη τιμή της μόνο όταν $X = x$, μέσω της σχέσης

$$X = x \Rightarrow Y_x = Y, \quad (3.15)$$

που για δυαδική μεταβλητή X γίνεται:

$$Y = xY_1 + (1 - x)Y_0$$

Πέραν αυτής της σχέσης ο ερευνητής μπορεί να παραλείψει τη πραγματικότητα ότι η Y_x είναι η ίδια η Y , μόνο μετρημένη κάτω από διαφορετικές συνθήκες (όπως στο Σχήμα 3.4(γ)) και προβαίνει στην εκτίμηση του μέσου όρου της αιτιώδους επίδρασης, $E(Y_{x'}) - E(Y_x)$, με τις

τεχνικές που έχουν αναπτυχθεί από τη Στατιστική για τα ελλειπή δεδομένα. Επιπροσθέτως εφόσον η εξίσωση (3.15), αποτελεί επίσης ένα Θεώρημα για τη λογική των *counterfactuals*, οι ερευνητές τους εγγυώνται να μην επιτυγχάνουν αποτελέσματα που έρχονται σε αντίθεση με όσα προκύπτουν από το δομικό πλαίσιο.

Η αδυναμία αυτής της προσέγγισης αντιμετωπίζεται στο πρόβλημα της φάσης διαμόρφωσης, όπου στερείται διαγραμμάτων και δομικών εξισώσεων και οι ερευνητές αναγκάζονται να εκφράσουν την υπόθεση του συνόλου A σε μια γλώσσα που έχει τελείως μετακινηθεί από την επιστημονική γνώση.

Για παράδειγμα, για να εκφράσουμε το γεγονός ότι σε μια τυχαιοποιημένη μελέτη, η μεταβλητή X είναι ανεξάρτητη και από τις δυο εξωγενείς μεταβλητές ϵ_2, ϵ_3 (Σχήμα 3.4), Θα χρειαζόταν να γράψουμε την ‘ισχυρής αγνοησιμότητας’ έκφραση $X \perp\!\!\!\perp \{Z - 1, Z_0, Y_{00}, Y_{01}, Y_{11}\}$. Προκειμένου να ξεπεραστεί αυτό το εμπόδιο ο Pearl επινόησε έναν τρόπο που συνδυάζει τα βέλτιστα χαρακτηριστικά των προσεγγίσεων. Στηρίζεται στη κωδικοποίηση των αιτιωδών υποθέσεων στη γλώσσα των διαγραμμάτων ή των δομικών εξισώσεων, μεταφράζοντας τα *counterfactuals* χρησιμοποιώντας αξιωματισμούς που επινοούνται από την (3.15), και τελικώς ερμηνεύοντας τα αποτελέσματα σε μια απλή αιτιώδη γλώσσα. Το πρόβλημα της μεσολάβησης που παρουσιάζεται στην Ενότητα 3.5 δείχνει τον τρόπο με τον οποίο μια τέτοια ‘συμβίωση’ αποσαφηνίζει την εννοιολόγηση και την εκτίμηση των άμεσων και έμμεσων επιδράσεων.

3.4 Συμπεράσματα δομικών εξισώσεων προς έλεγχο

Η αιτιώδης ανάλυση στα γραφικά μοντέλα ξεκινά με τη διαπίστωση ότι όλες οι αιτιώδεις επιδράσεις είναι αναγνωρίσιμες όταν το μοντέλο είναι Μαρκοβιανό, δηλαδή το γράφημα είναι άκυκλο (δεν περιέχει κατευθυνόμενους κύκλους) και όλοι οι όροι σφάλματος είναι κοινώς ανεξάρτητοι. Τα μη Μαρκοβιανά μοντέλα που περιλαμβάνουν συσχετιζόμενα σφάλματα (απορρέουν από τους μη μετρήσιμους συγχυτικούς παράγοντες), επιτρέπουν την αναγνώριση μόνο κάτω από συγκεκριμένες καταστάσεις οι οποίες μπορούν να προσδιοριστούν με τη βοήθεια των γραφημάτων.

Δεδομένου ενός μη παραμετρικού *SEM*, M με k μεταβλητές (παρατηρήσιμες και μη), μετανομάζουμε τις μεταβλητές ως V_1, \dots, V_k προκειμένου να καταστεί δυνατή μια γενικότερη διατύπωση. Χρησιμοποιώντας τη διαδικασία της 3.2.4, σχηματίζουμε το γράφημα *DAG* G_M με τις κορυφές V_1, \dots, V_k . Η αιτιώδης *Markov* συνθήκη ισχύει [41, p. 30] οπότε μπορούμε να παραγοντοποιήσουμε την από κοινού κατανομή των μεταβλητών σύμφωνα με το Θεώρημα

που ακολουθεί.

Θεώρημα 1. (Αιτιώδης Markov συνθήκη) Κάθε κατανομή που παράγεται από ένα Markovιανό μοντέλο M , μπορεί να παραγοντοποιηθεί ως

$$P(v_1, v_2, \dots, v_k) = \prod_i^k P(v_i | pa_i(v_i)) \quad (3.16)$$

όπου οι μεταβλητές V_1, V_2, \dots, V_k είναι οι ενδογενείς μεταβλητές του μοντέλου M και οι όροι pa_i είναι οι ενδογενείς 'γονείς' των V_i , στο αιτιώδες γράφημα που συσχετίζεται με το μοντέλο M .

Για παράδειγμα η κατανομή που συσχετίζεται με το μοντέλο του Σχήματος 3.2(α) μπορεί να παραγοντοποιηθεί ως εξής:

$$P(z, y, x) = P(z)P(x|z)P(y|x) \quad (3.17)$$

εφόσον η μεταβλητή X είναι ενδογενής και 'γονέας' της Y , η Z είναι 'γονέας' της X και η Z δεν έχει 'γονέα'.

Πόρισμα 1. (Truncated παραγοντοποίηση) Για κάθε Markovιανό μοντέλο, η κατανομή που παράγεται από τη παρέμβαση $do(X = x_0)$ σε ένα σύνολο ενδογενών μεταβλητών δίδεται από την κατατμημένη παραγοντοποίηση

$$P(v_1, v_2, \dots, v_k | do(x_0)) = \prod_{i|V_i \notin X} P(v_i | pa_i)|_{x=x_0}$$

όπου οι $P(v_i | pa_i)$ είναι οι πριν τη παρέμβαση δεσμευμένες πιθανότητες.

Το παραπάνω Πόρισμα μας οδηγεί στο να αφαιρέσουμε από το γινόμενο της εξίσωσης (3.16) όλους τους παράγοντες που σχετίζονται με τις μεταβλητές παρέμβασης. Αυτό απορρέει από το γεγονός πως το μετά τη παρέμβαση μοντέλο είναι επίσης Markovιανό, και έτσι σύμφωνα με το Θεώρημα 1, πρέπει να παράγει μια κατανομή που παραγοντοποιείται σύμφωνα με το τροποποιημένο μοντέλο, αποδίδοντας το γινόμενο του Πορίσματος 1. Στο Σχήμα 3.2(β), η κατανομή $P(z, y | do(x_0))$ συσχετιζόμενη με το τροποποιημένο μοντέλο είναι

$$P(z, y | do(x_0)) = P(z)P(y|x_0)$$

όπου οι $P(z)$ και $P(y|x_0)$ είναι πανομοιότυπες με εκείνες που συσχετίζονται με την πριν τη παρέμβαση κατανομή της εξίσωσης (3.17). Όπως ήταν αναμενόμενο, η κατανομή της Z δεν επηρεάζεται από την παρέμβαση εφόσον

$$P(z|do(x_0)) = \sum_y P(z, y|do(x_0)) = \sum_y P(z)P(y|x_0) = P(z)$$

ενώ η κατανομή της Y δίδεται από την

$$P(y|do(x_0)) = \sum_z P(z, y|do(x_0)) = \sum_z P(z)P(y|x_0) = P(y|x_0)$$

Το παράδειγμα αυτό αποδεικνύει πως οι αιτιώδεις υποθέσεις που ενσωματώνονται στο μοντέλο M , μας επιτρέπουν να προβλέψουμε τη μετά τη παρέμβαση κατανομή από την πριν τη παρέμβαση κατανομή. Επιπλέον, μπορούμε να εκτιμήσουμε την αιτιώδη επίδραση της X στη Y , σε μη πειραματικά μοντέλα, εφόσον η $P(y|x_0)$ μπορεί να εκτιμηθεί από τα δεδομένα.

3.4.1 Το κριτήριο του d -διαχωρισμού

Στην ενότητα αυτή θα μελετήσουμε τα συμπεράσματα που συνεπάγονται από τα δομικά μοντέλα, πως μπορούν να συναχθούν και με ποιον τρόπο απορρέουν από τα γραφήματα. Για το λόγο αυτό παραθέτουμε τον παρακάτω χρήσιμο Ορισμό γνωστό ως d -διαχωρισμό [39].

Ορισμός 7. (d -διαχωρισμός) Ένα σύνολο S κόμβων θεωρείται ότι εμποδίζει ένα μονοπάτι p αν ισχύει τουλάχιστον ένα από τα ακόλουθα:

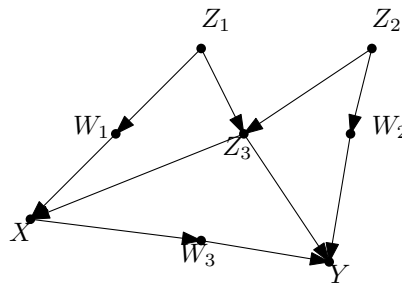
1. το μονοπάτι p περιέχει τουλάχιστον ένα εκπεμπόμενο βέλος από έναν κόμβο ο οποίος περιέχεται στο S , ή
2. το μονοπάτι p περιέχει τουλάχιστον ένα κόμβο σύγκρουσης ο οποίος δεν περιέχεται στο S και δεν έχει απογόνους σε αυτό.

Εάν το σύνολο S εμποδίζει όλα τα μονοπάτια από τη μεταβλητή X στη Y , λέμε ότι d -διαχωρίζει το X από το Y , και έπεται ότι οι μεταβλητές X, Y είναι ανεξάρτητες δεδομένου του S . Συμβολίζουμε $X \perp\!\!\!\perp Y | S$.

Αναλυτικότερα, θεωρούμε το μονοπάτι $U_Z \rightarrow Z \rightarrow X \rightarrow Y$ του Σχήματος 3.2(α), το οποίο εμποδίζεται από το σύνολο $S = \{Z\}$ και από το σύνολο $S = \{X\}$, εφόσον κάθε ένα εκπέμπει ένα βέλος κατά μήκος του. Κατά συνέπεια μπορούμε να συμπεράνουμε πως οι υπό

συνθήκη ανεξαρτησίας $U_Z \perp\!\!\!\perp Y|Z$ και $U_Z \perp\!\!\!\perp Y|X$, θα πρέπει να πληρούνται σε κάθε συνάρτηση πιθανότητας που παράγεται από το μοντέλο. Επίσης το μονοπάτι $U_Z \rightarrow Z \rightarrow X \leftarrow U_X$ εμποδίζεται από το κενό σύνολο $\{\emptyset\}$, αλλά δεν εμποδίζεται από το σύνολο $S = \{Y\}$, εφόσον το Y είναι απόγονος της μεταβλητής σύγκρουσης X . Οπότε η ανεξαρτησία $U_Z \perp\!\!\!\perp U_X$ θα ισχύει για την κατανομή, αλλά δεν γνωρίζουμε με βεβαιότητα αν θα ισχύει η $U_Z \perp\!\!\!\perp U_X|Y$.

Παρατηρώντας το γράφημα G του Σχήματος 3.6 διαπιστώνουμε πως κάθε d -διαχωρισμός που υφίσταται σε αυτό αντιστοιχεί σε έναν υπό συνθήκη ανεξάρτητο έλεγχο, ο οποίος μπορεί να εκτελεστεί με βάση τα δεδομένα, προκειμένου να αποδειχθεί η εγκυρότητα του μοντέλου M . Τα συμπεράσματα προς έλεγχο απαριθμούνται εύκολα παρατηρώντας κάθε ακμή που λείπει από το γράφημα και επιλέγοντας το σύνολο των μεταβλητών που d -διαχωρίζει το ζεύγος των μεταβλητών που αντιστοιχεί στην ακμή εκείνη που λείπει. Για παράδειγμα στο Σχήμα 3.6, τρεις από τις ακμές που λείπουν είναι οι $Z_1 - Z_2$, $Z_1 - Y$ και $Z_2 - X$, με τα σύνολα διαχωρισμού $\{\emptyset\}$, $\{X, Z_2, Z_3\}$ και $\{Z_1, Z_3\}$, αντίστοιχα. Επομένως τα συμπεράσματα προς έλεγχο του μοντέλου M περιλαμβάνουν τις $Z_1 \perp\!\!\!\perp Z_2$, $Z_1 \perp\!\!\!\perp Y|\{X, Z_2, Z_3\}$ και $Z_2 \perp\!\!\!\perp X|\{Z_1, Z_3\}$.



Σχήμα 3.6: Ο d -διαχωρισμός με βάση το Μαρκοβιανό μοντέλο. Τα τυχαία σφάλματα θεωρούνται κοινώς ανεξάρτητα και δεν παρουσιάζονται με σαφήνεια.

Στα γραμμικά μοντέλα αυτές οι υπό συνθήκη ανεξάρτητοι περιορισμοί μεταφράζονται σε μηδενικές συσχετίσεις ή μηδενικούς συντελεστές στις αντίστοιχες εξισώσεις παλινδρόμησης. Ως εκ τούτου οι τρεις παραπάνω συνεπαγωγές μεταφράζονται στους ακόλουθους περιορισμούς:

$$r_{Z_1 Z_2} = 0, \quad r_{Z_1 Y \cdot X Z_2 Z_3} = 0, \quad \text{και} \quad r_{Z_2 X \cdot Z_1 Z_3} = 0.$$

Τέτοιοι έλεγχοι μπορούν εύκολα να διεξαχθούν με συνηθισμένες τεχνικές παλινδρόμησης, παρέχοντάς μας πολύτιμες πληροφορίες στη περίπτωση αποτυχίας του μοντέλου σε ενδεχόμενη τροποποίησή του. Όταν το μοντέλο είναι Μαρκοβιανό (άκυκλα και με ασυσχέτιστα σφάλματα), οι όροι που δημιουργούνται από τον d -διαχωρισμό, είναι τα μόνα συμπεράσματα

προς έλεγχο που προκύπτουν. Αν στο μοντέλο περιλαμβάνονται συσχετιζόμενα σφάλματα, τότε επιβάλλονται επιπρόσθετοι περιορισμοί που καλούνται ‘λανθάνουσες ανεξαρτησίες’.

Παρουσιάζουμε ένα ακόμη απλό παράδειγμα από το οποίο μπορεί να φανεί γιατί d -διαχωρισμός σχετίζεται με την υπό συνθήκη ανεξαρτησία στα μη παραμετρικά SEM . Υποθέτουμε ότι το γρασίδι του κήπου μας είναι είτε βρεγμένο ($Y = 1$) είτε στεγνό ($Y = 0$), και το γρασίδι είτε είναι βρεγμένο είτε όταν βρέχει ($X = 1$) είτε όταν το αυτόματο πότισμα είναι ανοιχτό ($Z = 1$). Επίσης υποθέτουμε ότι αποφασίζουμε αν θα ανοίξουμε ή όχι το αυτόματο πότισμα καθημερινώς ρίχνοντας ένα αμερόληπτο νόμισμα. Λόγω της τυχαίας ανάθεσης του μηχανισμού ποτίσματος, οι μεταβλητές X και Z είναι οριακώς ανεξάρτητες μεταξύ τους. Ωστόσο η κατάσταση της Y μειώνει την ανεξαρτησία (αν γνωρίζουμε ότι το γρασίδι είναι βρεγμένο και γνωρίζουμε επίσης ότι το αυτόματο πότισμα δεν ήταν ανοιχτό, τότε ξέρουμε ότι θα πρέπει να είχε βρέξει). Οπότε οι X και Z είναι ανεξάρτητες υπό συνθήκη στη Y . Το μη παραμετρικό SEM που σχετίζεται με αυτό το παράδειγμα είναι το ακόλουθο:

$$\begin{aligned} z &= f_Z(u_1) \\ z &= f_X(u_2) \\ y &= f_Y(x, z) \end{aligned}$$

όπου η u_1 παριστάνει το αποτέλεσμα του αμερόληπτου νομίσματος που ρίχνουμε και η u_2 το αποτέλεσμα αν θα βρέξει ή όχι. Σε αυτό το παράδειγμα, η y εξαρτάται εξαρτάται ντετερμινιστικά από το z και x , οπότε η ύπαρξη της u_3 δεν κρίνεται απαραίτητη. Επίσης το μοντέλο μπορεί να παρουσιαστεί από το ακόλουθο DAG :

$$U_1 \rightarrow Z \rightarrow Y \leftarrow X \leftarrow U_2$$

Ακολουθώντας τους κανόνες του d -διαχωρισμού παρατηρούμε ότι το μονοπάτι του Z στο X εμποδίζεται από το κόμβο σύγκρουσης Y καθώς δεν θέτουμε κάποια συνθήκη στους απογόνους του Y . Οπότε οι Z και η X είναι d -διαχωρισμένες υπό το κενό σύνολο. Εάν επιμείνουμε για να θέσουμε κάποια συνθήκη στη Y τότε θα έχουμε θέσει κάποια συνθήκη και στους απογόνους του Y , με αποτέλεσμα το μονοπάτι από το Z στο X να μην εμποδίζεται πλέον. Έτσι οι κανόνες του d -διαχωρισμού ανακάπτουν τις υπό συνθήκη ανεξαρτησία συνθήκες.

3.4.2 Οι κανόνες του *do*-λογισμού

Χρησιμοποιώντας τα μη παραμετρικά *SEM* και το *d*-διαχωρισμό όπως έχουν παρουσιαστεί στις προηγούμενες υποενότητες, ο *Pearl* [41] παρέχει γραφικά κριτήρια για την γενικότερη αναγνώριση των πληθυσμιακών αιτιωδών επιδράσεων.

Ορισμός 8. (Κανόνες του *do*-λογισμού ([41, p. 85]) Υποθέτουμε ότι το G είναι ένα άμεσο κατευθυνόμενο γράφημα που σχετίζεται με το αιτιώδες μοντέλο του Ορισμού 1, και η $P(\cdot)$ συμβολίζει την κατανομή πιθανότητας που συνεπάγεται από αυτό το μοντέλο. Για κάθε αναξάρτητα υποσύνολο των μεταβλητών X, Y, Z και W . Συμβολίζουμε επίσης με $G_{\bar{X}}$ το γράφημα που λαμβάνεται αν αφαιρέσουμε από το G όλα τα βέλη που δείχνουν στο κόμβο του X . Ακόμη συμβολίζουμε με $G_{\underline{X}}$ το γράφημα που λαμβάνεται αν αφαιρέσουμε από το G όλα τα βέλη που εκπέμπονται από κόμβο του X . Για να παρουσιάσουμε την ταυτόχρονη διαγραφή και των δύο τύπων βελών, χρησιμοποιούμε το συμβολισμό $G_{\bar{X}, \underline{Z}}$. Τέλος η παράσταση $P(y|do(x), z) \equiv P(y, z|do(x))/P(z|do(x))$ συμβολίζει τη πιθανότητα της $Y = y$ δεδομένου ότι η X διατηρείται σταθερή και ίση με x , με αποτέλεσμα η $Z = z$ να παρατηρείται. Ακολουθούν οι παρακάτω κανόνες:

1. (Εισαγωγή/Διαγραφή παρατηρήσεων)

$$P(y|do(x), z, w) = P(y|do(x), w) \quad \text{εάν} \quad (Y \perp\!\!\!\perp Z | X, W)_{G_{\bar{X}}}$$

2. (Ενέργεια/Παρατήρηση ανταλλαγής)

$$P(y|do(x), do(z), w) = P(y|do(x), z, w) \quad \text{εάν} \quad (Y \perp\!\!\!\perp Z | X, W)_{G_{\bar{X}, \underline{Z}}}$$

3. (Εισαγωγή/Διαγραφή των ενεργειών)

$$P(y|do(x), do(z), w) = P(y|do(x), w) \quad \text{εάν} \quad (Y \perp\!\!\!\perp Z | X, W)_{G_{\bar{X}, \overline{Z(W)}}}$$

όπου η $Z(W)$ είναι το σύνολο Z κόμβων που δεν είναι πρόγονοι κανενός W κόμβου στο $G_{\bar{X}}$

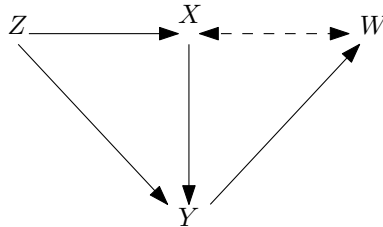
Ο πρώτος κανόνας περιγράφει τη κατάσταση όπου η προσαρμογή επιπρόσθετων μεταβλητών δεν κρίνεται απαραίτητη για την παρέμβαση. Ο δεύτερος περιγράφει τη κατάσταση ανάλογη στην υπό συνθήκη αγνοσιμότητα όπου η πρότυπες μέθοδοι προσαρμογής θα είναι σε θέση να αναγνωρίσουν τις αιτιώδεις επιδράσεις. Ο *Pearl* [41] αναφέρεται σε αυτό το

κανόνα ως *back-door* κριτήριο, το οποίο θα παρουσιάσουμε στη πορεία. Ο τρίτος κανόνας περιγράφει τις καταστάσεις όπου η παρέμβαση σε μια μεταβλητή δεν επηρεάζει το αποτέλεσμα που μας ενδιαφέρει. Εφαρμόζοντας επαναληπτικά αυτούς τους τρεις κανόνες, μπορούμε ορισμένες φορές να αναγνωρίσουμε τους μέσους όρους των αιτιωδών επιδράσεων, παρουσιάζοντας τη μη παρατηρήσιμη μετά τη παρέμβαση κατανομή με τη παρατηρήσιμη μετά τη παρέμβαση κατανομή.

3.4.3 Ισοδύναμα μοντέλα

Το πρόβλημα των ισοδύναμων μοντέλων απασχολεί τους ερευνητές από την αρχή της δημιουργίας των *SEM*. Ισοδύναμα μοντέλα θεωρούνται εκείνα που δίδουν τα ίδια σύνολα στατιστικών δείκτων προσαρμογής (κριτήρια *chi-square* και *p-value*), όπως και τα υποθετικά μοντέλα. Ο *Stelzl* [60] επισημοποίησε την έννοια των ισοδύναμων μοντέλων, παρουσιάζοντας βασικούς κανόνες για την εύρεσή τους. Οι κανόνες αυτοί γενικεύτηκαν από τους *Lee* και *Hershberger*[30]. Ο *d*-διαχωρισμός ορίζει συνθήκες για την ισοδυναμία των μοντέλων που μπορούν να προσδιοριστούν τόσο στα Μαρκοβιανά όσο και στα Ημι-Μαρκοβιανά μοντέλα. Αυτές οι συνθήκες θα πρέπει να τροποποιούν τους περιορισμούς που ισχύουν στη μελέτη των εξισώσεων των δομικών μοντέλων. Ο βασικός κανόνας προκειμένου να διατηρηθεί η ισοδυναμία είναι, σε κάθε τροποποίηση που υφίσταται το μοντέλο, να μην δημιουργηθεί ή καταστραφεί οποιαδήποτε συνθήκη *d*-διαχωρισμού.

Για παράδειγμα ας θεωρήσουμε το μοντέλο του Σχήματος 3.7, όπου σύμφωνα με το κριτήριο αντικατάστασης των *Lee* και *Hershberger* μπορούμε να αντικαταστήσουμε το βέλος $X \rightarrow Y$ με το διπλό βέλος $X \leftrightarrow Y$, όταν όλοι οι προγνωστικοί παράγοντες (Z), της μεταβλητής επίδρασης (Y) είναι οι ίδιοι που αντιστοιχούν και στη μεταβλητή (X). Δυστυχώς από το μοντέλο που έχει προκύψει έπειτα από την αντικατάσταση συνεπάγει το περιορισμό $r_{WZ.Y} = 0$, το οποίο δε μπορεί να συναχθεί από το μοντέλο που είχαμε πριν την αντικατάσταση. Αυτό μπορεί να παρατηρηθεί και από το γεγονός ότι το μονοπάτι $Z \rightarrow Y \leftarrow X \leftrightarrow W$, εξαρτώμενο από τη μεταβλητή Y , δεν εμποδίζεται αλλά μπορεί να εμποδιστεί αν αντικατασταθεί από το μονοπάτι $Z \rightarrow Y \leftrightarrow X \leftrightarrow W$. Το ίδιο ισχύει και για το μονοπάτι $Z \rightarrow X \leftrightarrow W$, εφόσον η μεταβλητή Y παύει να είναι απόγονος της X .



Σχήμα 3.7: Παρουσιάζεται η διαφορά μεταξύ του κανόνα αντικατάστασης των *Lee* και *Hershberger* και του d -διαχωρισμού, που δεν επιτρέπει την αντικατάσταση $X \rightarrow Y$ από τη $X \leftrightarrow Y$.

3.5 Η χρήση γραφημάτων στο ‘*Back – Door*’ κριτήριο

Θεωρούμε μια μελέτη παρατήρησης όπου στόχος μας είναι να εντοπίσουμε την επίδραση της X στη Y . Υποθέτουμε ότι οι παράγοντες που κρίνονται σχετικοί του μοντέλου, είναι δομημένοι όπως στο Σχήμα 3.6. Ορισμένοι από αυτούς τους παράγοντες επηρεάζουν την απόκριση, ορισμένοι τη θεραπεία και ορισμένοι και τα δύο. Επίσης κάποιοι από αυτούς είναι ποσοτικές μεταβλητές και κάποιες ποιοτικές μεταβλητές. Το πρόβλημα συνοψίζεται στο αν το ερώτημα $Q = P(y|do(x))$ είναι αναγνωρίσιμο, δεδομένου του μοντέλου, και αν είναι, να εξάγουμε έναν εκτιμητή $Q(P)$ που θα μας οδηγήσει στην εκτίμηση του Q .

Το παραπάνω πρόβλημα λύνεται με προσαρμογή, δηλαδή επιλέγοντας ένα υπόσυνολο των παραπάνω παραγόντων για μέτρηση, στα οποία αν συγκρίνουμε τα θεραπευμένα με τα μη θεραπευμένα άτομα, έχοντας τις ίδιες τιμές στους επιλεγμένους παράγοντες, μας δίνουν την σωστή επίδραση της θεραπείας. Τέτοια σύνολα παραγόντων καλούνται ‘επαρκή σύνολα’ ή ‘αποδεκτά σύνολα’ για προσαρμογή. Το πρόβλημα για την εύρεση ενός αποδεκτού συνόλου έχει προβληματίσει για δεκαετίες τους επιδημιολόγους και τους επιστήμονες της κοινωνιολογίας.

Το ακόλουθο κριτήριο που καλείται ‘*back – door*’ [40] καθιστά εφικτή τη λύση αυτού του προβλήματος, παρέχοντάς μας μια γραφική μέθοδο για την επιλογή αποδεκτών συνόλων για προσαρμογή. Ακόμη αποδεικνύει ότι μη παραμετρικά ερωτήματα όπως $Q = P(y|do(x))$, κάποιες φορές μπορούν να προσδιοριστούν χωρίς τη γνώση της συναρτησιακής μορφής των εξισώσεων ή των κατανομών που αφορούν τις λανθάνουσες μεταβλητές του μοντέλου M .

Ορισμός 9. (Αποδεκτά σύνολα – Το *back-door* κριτήριο) Ένα σύνολο S είναι αποδεκτό (ή ‘επαρκές’) εάν ισχύουν και οι δυο συνθήκες:

1. Κανένα στοιχείο του συνόλου S δεν είναι απόγονος της X .

2. Τα στοιχεία του συνόλου S εμποδίζουν όλα τα 'back – door' μονοπάτια από τη X στη Y , δηλαδή όλα τα μονοπάτια τα οποία ολοκληρώνονται με ένα βέλος το οποίο δείχνει τη X .

Στο παραπάνω κριτήριο η έννοια της λέξης 'εμποδίζεται' ερμηνεύεται όπως και στον Ορισμό 7. Βασιζόμενοι σε αυτό το κριτήριο μπορούμε να παρατηρήσουμε βάση του Σχήματος 3.6 πως καθένα από τα σύνολα $\{Z_1, Z_2, Z_3\}$, $\{Z_1, Z_3\}$, $\{W_1, Z_3\}$ και $\{W_2, Z_3\}$ είναι αποδεκτό για προσαρμογή, εφόσον καθένα εμποδίζει όλα τα back – door μονοπάτια μεταξύ της X και της Y . Το σύνολο $\{Z_3\}$ ωστόσο δεν είναι αποδεκτό για προσαρμογή διότι δεν εμποδίζει το μονοπάτι $X \leftarrow W_1 \leftarrow Z_1 \rightarrow Z_3 \leftarrow Z_2 \rightarrow W_2 \rightarrow Y$.

Η διαίσθησή μας βάσει το back – door κριτήριο είναι η ακόλουθη. Τα back – door μονοπάτια στο διάγραμμα μεταφέρουν ψευδής συσχετισμούς από τη X στη Y , καθώς τα μονοπάτια κατά μήκος των βελών από τη X στη Y μεταφέρουν αιτιώδεις συσχετισμούς. Εμποδίζοντας τα back – door μονοπάτια εξασφαλίζουμε πως ο υπολογισμός της συσχέτισης μεταξύ της X και της Y είναι καθαρά αιτιώδης, δηλαδή εκπροσωπούν επάξια τη ποσότητα στόχο: την αιτιώδη επίδραση της X στη Y .

3.5.1 Αναγνώριση παραμέτρων και αιτιωδών επιδράσεων

Το back – door κριτήριο παρέχει απλές λύσεις σε πολλά προβλήματα προσδιορισμού, σε γραμμικά και μη γραμμικά μοντέλα, και συνοψίζεται στο επόμενο Θεώρημα.

Θεώρημα 2. (Αναγνώριση αιτιωδών επιδράσεων) Για οποιαδήποτε δύο ανεξάρτητα σύνολα μεταβλητών, X και Y σε ένα αιτιώδες διάγραμμα G , η αιτιώδης επίδραση της X στη Y δίδεται από τη σχέση

$$P(Y = y|do(X = x)) = \sum_s P(Y = y|X = x, S = s)P(S = s) \quad (3.18)$$

όπου S είναι οποιοδήποτε σύνολο συμμεταβλητών που ικανοποιούν τη back – door συνθήκη του Ορισμού 9.

Εφόσον όλοι οι παράγοντες της δεξιάς πλευράς της εξίσωσης μπορούν να εκτιμηθούν (με παλινδρόμηση) από τα πριν τη παρέμβαση δεδομένα, η αιτιώδης επίδραση μπορεί επίσης να εκτιμηθεί αμερόληπτα από αυτά τα δεδομένα χωρίς μεροληψία.

Στα γραμμικά συστήματα, η αναγνώριση των αιτιωδών επιδράσεων όπως στην εξίσωση (3.18) έχει ως σκοπό να μειώσει τα αθροίσματα και τα γινόμενα, των αμερόληπτων συντελεστών παλινδρόμησης. Για παράδειγμα εάν επιθυμούμε να εκτιμήσουμε την ολική επίδραση τ_{XY} της X στη Y στη γραμμική μορφή του Σχήματος 3.6, απλώς επιλέγουμε το συντελεστή παλινδρόμησης της Y στη X , αμερόληπτα σε οποιοδήποτε επαρκές σύνολο S , δίνοντας

$$\tau_{XY} = r_{YX \cdot S} = r_{YX \cdot Z_1, Z_3} = r_{YX \cdot W_1, Z_3} = \dots$$

Οι σημερινές μέθοδοι *SEM* δεν εκμεταλλεύονται τα πλεονεκτήματα αυτής της ικανότητας για να αποδώσουν το προσδιορισμό γραφικά, πριν την λήψη δεδομένων, και να εκτιμήσουν τις προσδιοριστικές ποσότητες άμεσα, διατηρώντας τα αποδεκτά σύνολα. Τα θεωρήματα του d - διαχωρισμού και το *back - door* κριτήριο μας επιτρέπουν να επικεντρωθούμε στην αναγνώριση των ποσοτήτων που έχουμε ορίσει ως στόχο και να εξάγουμε ένα προσδιοριστέο εκτιμητή.

3.5.2 Η σημασία των συμμεταβλητών

Η επιλογή ποιων μεταβλητών να προσαρμόσουμε όταν επιχειρούμε να εξάγουμε αιτιώδη συμπεράσματα από παρατηρήσιμα δεδομένα, είναι μια από τις πιο σημαντικές αποφάσεις που πρέπει να ληφθούν. Δυστυχώς έχει δοθεί μικρότερη προσοχή από το ζήτημα του πώς να εκτελέσουμε μια προσαρμογή, δεδομένων των κατάλληλων συμμεταβλητών.

Η συμβουλή που δίνεται όταν ένα γραμμικό μοντέλο είναι κατάλληλο για την αιτιώδη συμπερασματολογία και δεν είμαστε σίγουροι αν μια πριν τη θεραπεία μεταβλητή που έχει μετρηθεί, πρέπει να συμπεριληφθεί στο μοντέλο ή όχι, τότε είναι προτιμότερο να σφάλουμε προς τη πλευρά να περιληφθεί μια ενδεχόμενη άνευ σημασίας συμμεταβλητή. Αυτό το επιχείρημα οδηγεί τους ερευνητές στην προσαρμογή πολύ μεγάλου αριθμού μεταβλητών που έχουν μετρηθεί πριν τη θεραπεία. Δεν σκέφτονται όμως αν αυτές σχετίζονται ή όχι με το αποτέλεσμα, τη θεραπεία ανάθεσης και μεταξύ τους, αιτιωδώς.

Αρκετοί επιδημιολόγοι έχουν μια διαφορετική άποψη. Προκειμένου να εκτιμήσουν αν ένα συγκεκριμένο σύνολο πριν τη θεραπεία μεταβλητών Z λειτουργεί ως συγχυτικός παράγοντας (δηλαδή η αποτυχία της προσαρμογής του Z θα μεροληπούσε στις αιτιώδεις επιδράσεις των εκτιμήσεων), τείνουν να εστιάζουν στο αν οι διάφορες συσχετίσεις δεν είναι μηδενικές στη παρατηρήσιμη πριν τη παρέμβαση κατανομή. Συγκεκριμένα κρίνουν την επίδραση της X στη Y να συγχέονται από το σύνολο Z αν:

1. $X \not\perp Z$ και

2. $Y \not\perp Z|X$

Λόγω του ότι οι άνευ σημασίας συσχετίσεις είναι άμεσα παρατηρήσιμες από τα παρατηρήσιμα δεδομένα ένα τέτοιο κριτήριο έχει μεγάλη απήχηση. Με το κριτήριο μπορεί εύκολα να χρησιμοποιηθεί και αν το Z κριθεί ως συγχυτικός παράγοντας, τότε μπορεί να προσαρμοστεί με ορισμένους τρόπους, όπως η παλινδρόμηση, η στρωματοποίηση, η αντιστοίχιση.

Καμία όμως από τις παραπάνω προσεγγίσεις δεν είναι έγκυρη για την αναγνώριση συνόλων για προσαρμογή μεταβλητών που είναι επαρκής για τον έλεγχο των συγχυτικών παραγόντων. Η συμβολή του RCM συνεισφέρει σε μεγάλο βαθμό διότι εστιάζει τη προσοχή του ερευνητή αν είναι ορθή η υπό συνθήκη αγνοσιμότητα (μια αιτιώδης υπόθεση), για ένα δεδομένο σύνολο προσαρμοσμένων μεταβλητών. Το μεγαλύτερο πλεονέκτημα είναι ότι εάν η υπό συνθήκη αγνοσιμότητα ισχύει δεδομένου του Z , τότε μετά τη προσαρμογή του Z εγγυάται του να είναι επαρκής στον έλεγχο των συγχυτικών παραγόντων. Το μειονέκτημα όμως είναι πως από μόνο του το RCM παρέχει λίγη καθοδήγηση στην επιλογή των συνόλων *background* μεταβλητών που είναι πιθανόν να παρέχουν υπό συνθήκη αγνοσιμότητα. Η υπό συνθήκη αγνοσιμότητα είναι μια υπόθεση που ορίζεται για τα ενδεχόμενα αποτελέσματα και επομένως δεν είναι απόλυτα ελέγξιμη. Με την εξάρτηση του ενδεχόμενου αποτελέσματος και της σχετιζόμενης έλλειψης ελέγχου το RCM δεν είναι κατάλληλο. Ωστόσο το ντετερμινιστικό μοντέλο των δομικών εξισώσεων του *Pearl* προσφέρει στους ερευνητές πρόσθετα μέσα για την αξιολόγηση της επάρκειας των στατηγικών προσαρμογής.

Οι κανόνες του *do*-λογισμού που έχουν παρουσιαστεί παραπάνω δημιουργούν μια άλλη μορφή του γνωστού *back-door* κριτηρίου, που ελέγχει αν το σύνολο Z που έχει δοθεί είναι επαρκής για τον έλεγχο μεροληψίας των συγχυτικών παραγόντων.

Ορισμός 10. (*back-door κριτήριο* ([41, p. 79])) Δεδομένου ενός μοντέλου M και του σχετιζόμενου γραφήματός του G_M , ένα σύνολο A συμμεταβλητών Z ικανοποιούν το *back-door* κριτήριο για μια αιτιώδη μεταβλητή X και το αποτέλεσμα Y αν:

1. κανένα στοιχείο του Z είναι απόγονος του X και
2. το X είναι d -διαχωρισμένο από το Y με το Z στο γράφημα $G_{\underline{X}}$ με την αφαίρεση όλων των ακμών του X από το G_M

Αν το Z ικανοποιεί το *back-door* κριτήριο τότε η κατανομή των ενδεχόμενων μπορεί να υπολογιστεί χρησιμοποιώντας τη προσαρμογή στρωματοποίησης [7][45]:

$$P(Y_x = y) = \sum_z P(y|x, z)P(z)$$

όπου το z μπορεί να είναι πολυμεταβλητή. Ο *Pearl* αναφέρεται στη παραπάνω εξίσωση ως *back-door* προσαρμογή. Εφόσον το Z ικανοποιεί το *back-door* κριτήριο η τυπική προσαρμογή στρωματοποίησης ενδείκνυται, συνεπάγεται ότι η αντιστοίχιση ή η στρωματοποίηση του $P(x|z)$ (η βαθμολογία τάσης δεδομένης μιας πραγματοποιήσης μεταβλητής z του Z), μαζί με τις σχετιζόμενες προσαρμογές που χρησιμοποιούν την υπό συνθήκη αγνοησιμότητα, θα ενδείκνυται και αυτές. Αυτό ισχύει ανεξάρτητα αν όλες οι μεταβλητές που επηρεάζουν τη θεραπεία ανάθεσης ανήκουν στο Z .

3.5.3 Παραμετρικός προσδιορισμός στα γραμμικά *SEM*

Ένα πρόβλημα που απασχολήσει τους ερευνητές για δεκαετίες όσο αναφορά την αναγνώριση στα γραμμικά *SEM*, είναι οι συνθήκες υπό τις οποίες ο συντελεστής μονοπατιού β_{XY} μπορεί να εκτιμηθεί, και ποιες μεταβλητές χρησιμεύουν ως μεταβλητές παλινδρόμησης. Τη λύση έρχεται να δώσει ένα στενά συνδεδεμένο κριτήριο με το *back – door*, που καλείται ‘*single – door*’ [41].

Πόρισμα 2. (Το *single – door* κριτήριο) *Ας υποθέσουμε πως ο παράγοντας β_{XY} αποτελεί ένα δομικό συντελεστή του βέλους $X \rightarrow Y$ και ο παράγοντας $r_{YX.S}$ αποτελεί τον συντελεστή της X , στην παλινδρόμηση του Y στην X και στο σύνολο S , δηλαδή ισχύει $r_{YX.S} = \frac{\partial}{\partial x} E(Y|x, s)$. Η ισότητα $\beta_{XY} = r_{YX.S}$ ισχύει εάν*

1. το σύνολο S δεν περιέχει κανένα απόγονο της μεταβλητής Y και
2. αν το S εμποδίζει όλα τα μονοπάτια μεταξύ της X και της Y , εκτός από το απευθείας μονοπάτι $X \rightarrow Y$.

Στο Σχήμα 3.7 για παράδειγμα ο συντελεστής β_{XY} ισούται με τον $r_{YX.Z}$ ή με το συντελεστή b_1 της παλινδρόμησης με εξίσωση $Y = b_1X + b_2Z + \epsilon$, καθώς ο συντελεστής β_{YW} που αντιστοιχεί στο βέλος $Y \rightarrow W$, ισούται με τον συντελεστή $r_{WY.XZ}$. Παρόμοια μπορούμε να παρατηρήσουμε τις ισότητες $\beta_{ZY} = r_{YZ.X}$ και $\beta_{ZX} = r_{XZ.Y}$. Στην περίπτωση που κανένα σύνολο S δεν ικανοποιεί τις προϋποθέσεις του πορίσματος 2, τότε ο συντελεστής β_{XY} δε μπορεί να απλοποιηθεί σε έναν απλό συντελεστή παλινδρόμησης, οπότε επικαλούμαστε άλλες τεχνικές προσδιορισμού όπως οι βοηθητικές μεταβλητές [5].

3.5.4 Αναγνωρίζοντας τις βοηθητικές μεταβλητές

Οι βοηθητικές μεταβλητές αποτελούν τη πιο παλιά μέθοδο αναγνώρισης που έχει σχεδιαστεί για γραμμικά συστήματα. Η μέθοδος στηρίζεται στην εύρεση μιας μεταβλητής Z που συσχετίζεται με την X , και θεωρείται πως δεν συσχετίζεται με τους όρους σφαλμάτων της εξίσωσης. Ενώ κανένας στατιστικός έλεγχος δεν έχει τη δυνατότητα να πιστοποιήσει μία μεταβλητή ως βοηθητική (*instrumental*), το κριτήριο του d -διαχωρισμού μας επιτρέπει να αναγνωρίσουμε αυτές τις μεταβλητές σε ένα αιτιώδες γράφημα και να τις χρησιμοποιήσουμε για να προσδιορίσουμε τις παραμέτρους που δεν ικανοποιούν τους όρους του πορίσματος 2. Επίσης το αιτιώδες γράφημα, μας υποδεικνύει πως μπορούμε να μετατρέψουμε τις μεταβλητές σε βοηθητικές. Στο Σχήμα 3.6 για παράδειγμα, η μεταβλητή Z_1 δεν θεωρείται βοηθητική μεταβλητή για την επίδραση της Z_3 στην Y , διότι υπάρχει ένα απευθείας μονοπάτι από τη Z_3 στη Y μέσω της W_1 και της X . Εξάγοντας κάποια συμπεράσματα για τον παράγοντα X , δεν θα βελτιωθεί η κατάσταση, διότι η X καθώς είναι απόγονος του Z_3 δε θα εμποδίζεται το μονοπάτι $Z_1 \rightarrow Z_3 \leftarrow Z_2 \rightarrow W_2 \rightarrow Y$. Ωστόσο εξάγοντας κάποια συμπεράσματα για τον παράγοντα W_1 , καταστούμε τον παράγοντα Z_1 μία νόμιμη βοηθητική μεταβλητή, εφόσον όλα τα μονοπάτια που ενώνουν τη Z_1 με τη Y περνούν από τη Z_3 .

Το γενικό κριτήριο δίνεται από το Θεώρημα που ακολουθεί

Θεώρημα 3. (Αναγνώριση με τη χρήση των βοηθητικών μεταβλητών) *Ας υποθέσουμε πως ο παράγοντας β_{XY} αντιπροσωπεύει το συντελεστή του μονοπατιού του βέλους $X \rightarrow Y$ σε ένα αιτιώδες γράφημα G . Η παράμετρος β_{XY} είναι αναγνωρίσιμη, εάν υπάρχει ένα ζεύγος (Z, W) , όπου το Z αποτελεί έναν κόμβο (*single*) στο G (χωρίς να εξαιρούμε το $Z = X$), και το W είναι ένα (πιθανότατα κενό) σύνολο κόμβων στο G , τέτοιο ώστε :*

1. Το W δεν αποτελείται από απόγονους του Y ,
2. Το W προκαλεί τον d -διαχωρισμό του Z από το Y στο γράφημα G_{XY} που δημιουργείται αφού απομακρύνουμε το $X \rightarrow Y$ από το γράφημα G ,
3. Το Z και το X είναι d -συνδεδεμένα, δεδομένου του W , στο G_{XY} και επιπλέον ο εκτιμητής που επάγεται από το ζεύγος (Z, W) δίνεται από τον τύπο

$$\beta_{XY} = \frac{\text{cov}(Y, Z|W)}{\text{cov}(X, Z|W)}. \quad (3.19)$$

3.6 Μεσολάβηση: άμεσες και έμμεσες επιδράσεις

3.6.1 Διάσπαση

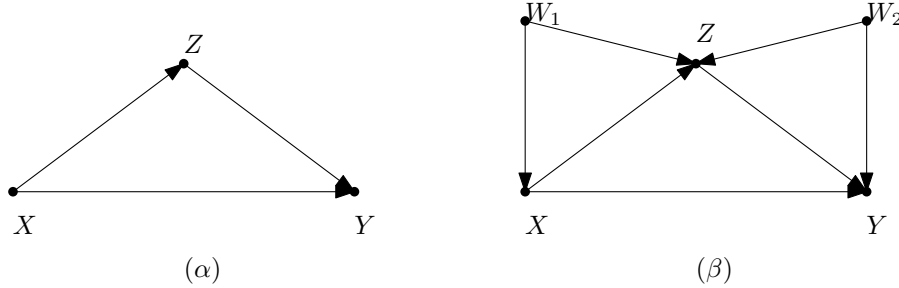
Η αιτιώδης επίδραση $P(y|do(x))$ που έχουμε αναλύσει μέχρι στιγμής, υπολογίζει την ολική επίδραση μιας μεταβλητής X σε μια μεταβλητή απόκρισης Y . Σε πολλές περιπτώσεις η ποσότητα αυτή δεν εκπροσωπεί επαρκώς το στόχο της έρευνάς μας και επικεντρώνει τη προσοχή μας στην άμεση επίδραση της X στην Y . Ο όρος άμεση επίδραση έχει σκοπό να προσδιορίσει ποσοτικά μια επίδραση η οποία δεν διαμεσολαβείται από άλλες μεταβλητές του μοντέλου.

Η διάσπαση των επιδράσεων σε άμεσες και έμμεσες, φέρει μία θεωρητική σημαντικότητα και ως εκ τούτου μας δίνεται η δυνατότητα να προβλέψουμε τη συμπεριφορά των επιδράσεων κάτω από μια ποικιλία καταστάσεων και παρεμβάσεων. Για παράδειγμα ένας ερευνητής ενδεχομένως να ενδιαφέρεται να εκτιμήσει το βαθμό στον οποίο η επίδραση μίας δεδομένης μεταβλητής μπορεί να μειωθεί, αποδυναμώνοντας μια ενδιάμεση διαδικασία, η οποία βρίσκεται ανάμεσα σε αυτή τη μεταβλητή και το αποτέλεσμα.

Τα δομικά μοντέλα εξισώσεων παρέχουν μία φυσική γλώσσα για την ανάλυση αυτών των ειδικών επιδράσεων μονοπατιών και μια αξιοσημείωτη βιβλιογραφία άμεσων, έμμεσων και συνολικών επιδράσεων, η οποία έχει συνταχθεί από τους ερευνητές των *SEM* [43]. Η ανάλυση αυτή συνήθως περιλαμβάνει ποσά ισχύος που αφορούν συντελεστές πινάκων, όπου κάθε Πίνακας αντιπροσωπεύει το συντελεστή μονοπατιού που συσχετίζεται με τις δομικές εξισώσεις.

Παρόλη τη γενικευμένη ύπαρξή της, η ανάλυση της μεσολάβησης αποτελεί ένα ακανθώδες ζήτημα στις κοινωνικές επιστήμες κατά κύριο λόγο διότι η διαφορά μεταξύ των αιτιωδών παραμέτρων και των παλινδρομικών τους ερμηνειών συχνά συγχέεται. Καθώς οι απαιτήσεις αυξήθηκαν, με αποτέλεσμα οι ερευνητές δεν ήταν σε θέση πλέον να προσδιορίσουν τις άμεσες και έμμεσες επιδράσεις σε όρους δομικών ή παλινδρομικών συντελεστών. Αυτό είχε ως αποτέλεσμα όλες οι προσπάθειες που αποσκοπούσαν στο να επεκταθούν τα γραμμικά παραδείγματα των διασπάσεων των επιδράσεων σε μη γραμμικά συστήματα, να παράγουν αλλοιωμένα αποτελέσματα. Η *counterfactual* ερμηνεία της εξίσωσης (3.14) δίνει τη δυνατότητα να προσδιοριστούν ξανά και να αναλυθούν οι άμεσες και έμμεσες επιδράσεις.

Άμεσες επιδράσεις



Σχήμα 3.8: Ένα γενικό μοντέλο που απεικονίζει την μεσολάβηση μέσω της μεταβλητής Z (α) χωρίς συγχυτικούς παράγοντες και (β) με δύο συγχυτικούς παράγοντες, τους W_1 και W_2 .

Εννοιολογικά, μπορούμε να προσδιορίσουμε την άμεση επίδραση $DE_{x,x'}(Y)$ ως την αναμενόμενη μεταβολή της μεταβολής Y που προκαλείται από τη μεταβολή της X , από x σε x' , καθώς διατηρούμε όλους τους άλλους παράγοντες σταθερούς, σε οποιοδήποτε επίπεδο είχαν παρατηρηθεί, στο πλαίσιο της παρέμβασης $do(x)$. Προσδιορίζουμε την άμεση επίδραση χρησιμοποιώντας τον *counterfactual* συμβολισμό

$$DE_{x,x'}(Y) = E(Y_{x',z_x}) - E(Y_x) \tag{3.20}$$

Η μεταβλητή Y_{x',z_x} αντιπροσωπεύει τη τιμή που θα είχε επιτευχθεί για την Y υπό τον καθορισμό της X σε x' , καθώς ταυτόχρονα καθορίζει τη μεταβλητή Z στην όποια τιμή θα είχε ληφθεί υπό τον καθορισμό $X = x$. Δεδομένων ορισμένων υποθέσεων χωρίς συγχυτικούς παράγοντες, είναι πιθανό να αποδειχθεί πως η άμεση επίδραση μπορεί να μετατραπεί σε μία ανηγμένη *do* έκφραση,

$$DE_{x,x'}(Y) = \sum_{zw} [E(Y|do(x', z), w) - E(Y|do(x, z), w)]P(z|do(x), w)P(w) \tag{3.21}$$

Όπου η μεταβλητή W πληροί το κριτήριο *back – door* σχετικά τόσο του $X \rightarrow Z$ όσο και του $(X, Z) \rightarrow Y$.

Η παραπάνω έκφραση είναι έγκυρη και προσδιορίσιμη ειδικότερα σε Μαρκοβιανά μοντέλα, όπου το δεξί μέλος μπορεί να ανηγθεί σε μία ελεύθερη από το *do* έκφραση χρησιμοποιώντας την εξίσωση (3.18), και έπειτα να εκτιμηθεί μέσω παλινδρόμησης.

Για παράδειγμα για το μοντέλο του Σχήματος 3.8(β) η εξίσωση (3.21) έχει ως εξής

$$DE_{x,x'}(Y) = \sum_z \sum_{w_2} P(w_2)[E(Y|(x', z, w_2)) - E(Y|x, z, w_2)] \sum_{w_1} P(z|x, w_1, w_2)P(w_1) \quad (3.22)$$

Ενώ για το μοντέλο χωρίς συγχυτικούς παράγοντες του Σχήματος 3.21(α) έχουμε

$$DE_{x,x'}(Y) = \sum_z [E(Y|x', z) - E(Y|x, z)]P(z|x) \quad (3.23)$$

Και οι δύο τελευταίες εξισώσεις μπορούν να εκτιμηθούν με δύο βήματα παλινδρόμησης.

Έμμεσες επιδράσεις

Ο ορισμός της άμεσης επίδρασης της εξίσωσης (3.20), μπορεί να διαφοροποιηθεί και να παράσχει έναν λειτουργικό ορισμό για την έμμεση επίδραση. Η έμμεση επίδραση, (συμβολίζεται ως IE από το *indirect effect*), της μετάβασης από το x στο x' , ορίζεται ως την αναμενόμενη μεταβολή της μεταβλητής Y , επηρεαζόμενη καθώς διατηρούμε την μεταβλητή X σταθερή, στο επίπεδο $X = x$, και μεταβάλλοντας την Z σε οποιαδήποτε τιμή θα είχε επιτευχθεί καθώς η X είχε καθοριστεί ως $X = x'$. Επίσημα αυτό αναφέρεται ως

$$IE_{x,x'}(Y) = E[(Y_{x,Z_{x'}}) - E(Y_x)] \quad (3.24)$$

Το οποίο είναι σχεδόν όμοιο με την άμεση επίδραση της εξίσωσης (3.20), εκτός από την ανταλλαγή των μεταβλητών x, x' στο πρώτο όρο.

Πράγματι μπορεί να αποδειχθεί γενικότερα πως η συνολική επίδραση TE (*total effect*) από την αλλαγή-μετάβαση είναι ίση με τη διαφορά ανάμεσα στην άμεση επίδραση αυτής της μετάβασης και της έμμεσης επίδρασης της αντίστροφης μετάβασης. Επίσημα αυτό δίδεται ως εξής:

$$TE_{x,x'}(Y) = E(Y_{x'} - Y_x) = DE_{x,x'}(Y) - IE_{x,x'}(Y) \quad (3.25)$$

Στα γραμμικά συστήματα, όπου η αντιστροφή των μεταβατικών καταστάσεων αντιστοιχεί στην εξουδετέρωση των ενδείξεων των επιδράσεων, έχουμε το καθορισμένο τύπο

$$TE_{x,x'}(Y) = DE_{x,x'}(Y) - IE_{x,x'}(Y) \quad (3.26)$$

Εφόσον κάθε όρος της παραπάνω σχέσης βασίζεται σε έναν ανεξάρτητο λειτουργικό ορισμό, η ισότητα συνιστά μία επίσημη αιτιολογία για τον προσθετικό τύπο που χρησιμοποιείται κατά κόρον στα γραμμικά συστήματα.

3.6.2 Ο Τύπος της μεσολάβησης: Μία απλή λύση για το ακανθώδες πρόβλημα

Στο σημείο αυτό θα προσπαθήσουμε να παρουσιάσουμε πως η λύση που μας παρείχαν οι εξισώσεις (3.23), (3.26), μπορούν να εφαρμοστούν στην εκτίμηση των επιδράσεων της μεσολάβησης σε μη γραμμικά μοντέλα. Θα χρησιμοποιήσουμε το απλό μοντέλο μεσολάβησης του Σχήματος 3.8(α), όπου όλοι οι όροι σφαλμάτων θεωρούνται κοινώς ανεξάρτητοι, με την προϋπόθεση πως η προσαρμογή κατάλληλων συνόλων συμμεταβλητών W ενδέχεται να είναι απαραίτητη για να επιτευχθεί η ανεξαρτησία. Συνδυάζοντας τις εξισώσεις (3.23), (3.25) η έκφραση για την έμμεση επίδραση, IE , μετατρέπεται

$$IE_{x,x'}(Y) = \sum_z E(Y|x, z)[P(z|x') + P(z|x)] \quad (3.27)$$

η οποία μας παρέχει ένα γενικό τύπο για τη μεσολάβηση των επιδράσεων, που ισχύει σε οποιοδήποτε μη γραμμικό σύστημα, οποιαδήποτε κατανομή, και οποιοδήποτε τύπο μεταβλητών. Επιπλέον ο τύπος είναι εύκολα εκτιμώμενος μέσω παλινδρόμησης. Λόγω της γενικότητας αυτού του τύπου αναφερόμαστε σε αυτήν την έκφραση ως 'τύπος μεσολάβησης'.

Ο συγκεκριμένος τύπος μεσολάβησης παρουσιάζει τη μέση αύξηση του αποτελέσματος Y ώστε η μετάβαση από το $X = x$ στο $X = x'$, να αναμένεται να παράγει την οποιαδήποτε απουσία άμεσης επίδρασης της Y στη X . Μία σκέψη που βασίστηκε σε σταθερές αιτιώδεις αρχές και ενσωματώνει μη αιτιώδεις υποθέσεις, πλην της γενικής μεσολάβησης όπως είναι δομημένη στο Σχήμα 3.8(α). Στην περίπτωση που το αποτέλεσμα Y είναι δυαδικής μορφής η αναλογία $(1 - IE/TE)$ παρουσιάζει το ποσοστό της ανταπόκρισης των ατόμων που οφείλουν την απόκρισή τους στα άμεσα μονοπάτια, καθώς η αναλογία $(1 - DE/TE)$ παρουσιάζει το ποσοστό εκείνων που οφείλουν την απόκρισή τους στα Z - μεσολαβούμενα μονοπάτια.

Ο τύπος μεσολάβησης μας βεβαιώνει πως η έμμεση επίδραση IE εξαρτάται μόνο από την αναμενόμενη τιμή του *counterfactual* Y_{xz} , και όχι από τη συναρτησιακή του μορφή $f_Y(x, z, u_Y)$ ή της συνάρτησης κατανομής του $P(Y_{xz} = y)$. Ως εκ τούτου χρησιμοποιείται για δύο βημάτων παλινδρόμηση, η οποία μπορεί να πραγματοποιηθεί μη παραμετρικά. Στο πρώτο βήμα παλινδρομούμε το Y στο X και Z , και λαμβάνουμε την εκτίμηση

$$g(x, z) = E(Y|x, z) \quad (3.28)$$

για κάθε ζεύγος (x, z) . Στο δεύτερο βήμα σταθεροποιούμε το x και θεωρούμε την εκτίμηση $g(x, z)$ ως συνάρτηση $g_x(z)$ του Z . Έπειτα υπολογίζουμε την δεσμευμένη αναμονή της $g_x(z)$ υπό τις συνθήκες $X = x'$ και $X = x$, αντίστοιχα, και παίρνουμε τη διαφορά

$$IE_{x,x'}(Y) = E_{Z|X}[g_x(z)|x'] - E_{Z|X}[g_x(z)|x] \quad (3.29)$$

Οι μη παραμετρικές εκτιμήσεις δεν είναι πάντοτε πρακτικές-χρήσιμες. Όταν το Z , είναι φορέας πολλών μεσολαβητών, η διάσταση ενός προβλήματος ενδεχομένως να καθιστά αδύνατη την εκτίμηση του $E(Y|x, z)$ για κάθε ζεύγος (x, z) , και προκύπτει η ανάγκη του να χρησιμοποιήσουμε παραμετρική προσέγγιση. Τότε μπορούμε να επιλέξουμε οποιαδήποτε κατάλληλο παραμετρικό τύπο για τον μέσο $E(Y|x, z)$, ούτως ώστε να εκτιμήσουμε τις παραμέτρους ξεχωριστά, εισάγοντας την παραμετρική προσέγγιση στην (3.27) και υπολογίζοντας τις δύο δεσμευμένες αναμονές για να λάβουμε την επίδραση της μεσολάβησης.

Ας προσπαθήσουμε να εξετάσουμε τι αποδίδει ο τύπος μεσολάβησης όταν εφαρμόζεται στην γραμμική μορφή του Σχήματος 3.8(α) για την οποία ισχύει

$$\begin{aligned} x &= u_X \\ z &= b_0 + \beta x + u_Z \\ y &= c_0 + \alpha x + \gamma z + u_Y \end{aligned} \quad (3.30)$$

με τα u_X, u_Y και u_Z να είναι ασυσχέτιστα και τους όρους σφαλμάτων να έχουν μηδενική μέση τιμή. Υπολογίζοντας την δεσμευμένη αναμονή της εξίσωσης (3.28) έχουμε,

$$E(Y|x, z) = E(c_0 + \alpha x + \gamma z + u_Y) = c_0 + \alpha x + \gamma z$$

δίνοντας

$$\begin{aligned} IE_{x,x'}(Y) &= \sum_z (\alpha x + \gamma z)[P(z|x') - P(z|x)] \\ &= \gamma[E(Z|x') - E(Z|x)] \end{aligned} \quad (3.31)$$

$$= (x' - x)(\beta\gamma) \quad (3.32)$$

$$= (x' - x)(\tau - \alpha) \quad (3.33)$$

όπου το τ αποτελεί την κλίση της συνολικής επίδρασης, για το οποίο ισχύει

$$\tau = ((E(Y|x') - E(Y|x)) / (x' - x)) = \alpha + \beta\gamma$$

Έτσι έχουμε λάβει τις τυπικές εκφράσεις για τις έμμεσες επιδράσεις στα γραμμικά συστήματα, οι οποίες μπορούν να εκτιμηθούν είτε από τη διαφορά $\tau - \alpha$, που αποτελείται από δύο συμμεταβλητές παλινδρόμησης όπως στην (3.33), είτε από το γινόμενο $\beta\gamma$, το οποίο αποτελείται επίσης από δύο συμμεταβλητές παλινδρόμησης και παρουσιάζεται στην (3.32). Οι δύο αυτές μέθοδοι δεν μπορούν να γενικευτούν για όλα τα μη γραμμικά συστήματα.

Προκειμένου να κατανοήσουμε τη δυσκολία, ας υποθέσουμε πως το κατάλληλο μοντέλο περιλαμβάνει το γινόμενο δ_{xz} , το οποίο εάν το συμπεριλάβουμε στην (3.30) μας δίνει

$$y = c_0 + \alpha x + \gamma z + \delta_{xz} + u \quad (3.34)$$

επίσης υποθέτουμε πως ορθά προσθέσαμε αυτόν τον όρο και μέσω της ανάλυσης παλινδρόμησης αποκτούμε ακριβείς εκτιμήσεις για όλες τις παραμέτρους του μοντέλου. Δεν είναι ακόμη βέβαια σαφές ποιοι συνδυασμοί παραμέτρων υπολογίζουν τις άμεσες και έμμεσες επιδράσεις του X στο Y ή πιο συγκεκριμένα πως θα αξιολογηθεί το κλάσμα - αναλογία της συνολικής επίδρασης που εξηγείται από τη μεσολάβηση και το κλάσμα - αναλογία που οφείλεται σε αυτή τη μεσολάβηση. Στα γραμμικά συστήματα το πρώτο κλάσμα - αναλογία αποτυπώνεται μέσω του γινομένου $\beta\gamma/\tau$, και το δεύτερο από τη διαφορά $(\tau - \alpha)/\tau$. Οι δύο αυτές ποσότητες συμπίπτουν, ωστόσο όπως υπαγορεύεται από το τύπο διαμεσολάβησης, λόγω αλληλεπιδράσεων κάθε κλάσμα - αναλογία απαιτεί ξεχωριστή ανάλυση.

Αντικαθιστώντας τη μη γραμμική εξίσωση στις (3.23), (3.26) και (3.27) και υποθέτοντας ότι $x = 0$ και $x' = 1$, έχουμε την ακόλουθη διάσπαση της επίδρασης

$$\begin{aligned} DE &= \alpha + b_0\delta \\ IE &= \beta\gamma \\ TE &= \alpha + b_0\delta + \beta(\gamma + \delta) \\ &= DE + IE + \beta\gamma \end{aligned}$$

Οπότε μπορούμε να συμπεράνουμε ότι μέρος του αποτελέσματος αλλάζει, για όποια μεσολάβηση θεωρείται επαρκής όταν

$$IE = \beta\gamma$$

ενώ για το μέρος για το οποίο η μεσολάβηση θεωρείται απαραίτητη ισχύει

$$TE - DE = \beta\gamma + \beta\delta$$

Επισημαίνουμε πως λόγω της αλληλεπίδρασης, μία άμεση επίδραση μπορεί να διατηρηθεί ακόμη και όταν η παράμετρος α εξαλειφτεί. Γενικότερα για μία συνολική επίδραση ισχύει ότι μπορεί να διατηρηθεί ακόμη και αν και η άμεσες και έμμεσες επιδράσεις εξαλειφτούν. Αυτό επεξηγεί το γεγονός ότι οι παράμετροι οι οποίοι έχουν εκτιμηθεί, μεμονωμένοι μπορούν να δώσουν λίγες πληροφορίες για την επίδραση της μεσολάβησης.

Η κύρια δυνατότητα του τύπου μεσολάβησης φαίνεται σε μελέτες που συμπεριλαμβάνουν κατηγορηματικές μεταβλητές, ειδικότερα αν δεν έχουμε κάποιο παραμετρικό μοντέλο που προκύπτει από τα δεδομένα. Η χαμηλή διάσταση της περίπτωσης δυαδικών μεταβλητών, επιτρέπει μια μη παραμετρική λύση και μια σαφή επίδειξη για το πώς η μεσολάβηση μπορεί να εκτιμηθεί απευθείας από τα δεδομένα.

Ας υποθέσουμε πως το μοντέλο του Σχήματος 3.8(α) είναι ορθό και έτσι τα παρατηρήσιμα δεδομένα δίνονται από το Πίνακα 1. Οι παράγοντες $E(Y|x, z)$ και $P(Z|x)$ μπορούν άμεσα να υπολογισθούν όπως φαίνεται από τις δύο τελευταίες στήλες του Πίνακα 1, και αντικαθιστώντας στις εξισώσεις 3.23, 3.26, 3.27 λαμβάνουμε

$$DE = (g_{10} - g_{00})(1 - h_0) + (g_{11} - g_{01})h_0 \quad (3.35)$$

$$IE = (h_1 - h_0)(g_{01} - g_{00}) \quad (3.36)$$

$$TE = g_{11}h_1 + g_{10}(1 - h_1) - [g_{01}h_0 + g_{00}(1 - h_0)] \quad (3.37)$$

Παρατηρούμε πως η λογιστική ή η *probit* παλινδρόμηση δεν είναι αναγκαία. Απλές μαθηματικές πράξεις αρκούν για να μας δώσουν μία γενική λύση για οποιοδήποτε πιθανά δεδομένα ανεξάρτητα από τα δεδομένα που προκύπτουν από τη διαδικασία παραγωγής.

Ένα αριθμητικό παράδειγμα

Υποθέτουμε πως με $X = 1$ συμβολίζουμε τη θεραπεία ενός φάρμακου, με $Y = 1$ την ανάρρωσή του και με $Z = 1$ την παρουσία ενός ορισμένου ενζύμου στο αίμα του ασθενή, το οποίο δείχνει να διεγείρεται από τη θεραπεία. Ας υποθέσουμε ακόμη πως τα δεδομένα που περιγράφονται στους Πίνακες 3.4 και 3.5 ελήφθησαν από μία τυχαία κλινική δοκιμή και ότι το ερώτημα της έρευνάς μας είναι είτε εάν το Z μεσολαβεί για τη δράση του X στο Y , είτε λειτουργεί απλώς ως καταλύτης που επιταχύνει τη δράση του X στο Y .

αριθμός δειγμάτων	X	Y	Z	$E(Y x, z) = g_{xz}$	$E(Z x) = h_x$
n_1	0	0	0	$\frac{n_2}{n_1+n_2} = g_{00}$	$\frac{n_3+n_4}{n_1+n_2+n_3+n_4} = h_0$
n_2	0	0	1		
n_3	0	1	0	$\frac{n_4}{n_3+n_4} = g_{01}$	
n_4	0	1	1		
n_5	1	0	0	$\frac{n_6}{n_5+n_6} = g_{10}$	$\frac{n_7+n_8}{n_5+n_6+n_7+n_8} = h_1$
n_6	1	0	1		
n_7	1	1	0	$\frac{n_8}{n_7+n_8} = g_{11}$	
n_8	1	1	1		

Πίνακας 3.3:

θεραπεία X	παρουσία ενζύμου Z	ποσοστό που έχει θεραπευτεί $g_{xz} = E(Y x, z)$
YES	YES	$g_{11} = 80\%$
YES	NO	$g_{10} = 40\%$
NO	YES	$g_{01} = 30\%$
NO	NO	$g_{00} = 20\%$

Πίνακας 3.4:

θεραπεία X	ποσοστό παρουσίας του Z
NO	$h_0 = 40\%$
YES	$h_1 = 75\%$

Πίνακας 3.5:

Αντικαθιστώντας τα παραπάνω δεδομένα στις εξισώσεις (3.35)-(3.37) λαμβάνουμε

$$DE = (0.40 - 0.20)(1 - 0.40) + (0.80 - 0.30)0.40 = 0.32$$

$$IE = (0.75 - 0.40)(0.30 - 0.20) = 0.035$$

$$TE = 0.80 \times 0.75 + 0.40 \times 0.25 - (0.30 \times 0.40 + 0.20 \times 0.60) = 0.46$$

$$IE/TE = 0.07 \quad DE/TE = 0.696 \quad 1 - DE/TE = 0.304$$

Καταλήγουμε στο συμπέρασμα πως το 30.4% όλων των ατόμων που έχουν αναρρώσει οφείλεται στην ικανότητα που έχει η θεραπεία να ενισχύει την έκκριση του ενζύμου, ενώ μόνο το 7% των ανθρώπων που έχουν αναρρώσει έχουν υποστεί μόνο αύξηση του ενζύμου. Μία τέτοια μελέτη θα έχει ως αποτέλεσμα να καταβληθούν προσπάθειες να αναπτυχθεί ένα φθηνότερο φάρμακο, όμοιο με αυτό που έχει μελετηθεί, χωρίς όμως να διαθέτει τη δυνατότητα να ενισχύσει την έκκριση του ενζύμου, που θα έχει ως αποτέλεσμα μία μείωση της τάξεως του 30.4% στις περιπτώσεις ανάρρωσης. Προτάσεις για ένα υποκατάστατο φαρμάκου που μιμείται τη διεγερτική δράση στο Z αλλά δεν έχει άμεση επίδραση στο Y , είναι βέβαιο πως θα αποτύχουν. Το φάρμακο προφανώς από μόνο του έχει μία ευεργετική επίδραση στην ανάρρωση που είναι ανεξάρτητη από την ενίσχυση του φαρμάκου με διέγερση.

Συγκρίνοντας αυτά τα αποτελέσματα με εκείνα που έχουν ληφθεί από τη κλασική ανάλυση της μεσολάβησης, θα πρέπει να παρατηρήσουμε ότι η κλασική μέθοδος δεν προσδιορίζει τις άμεσες και έμμεσες επιδράσεις σε ένα σύνολο όπου η υποκείμενη διαδικασία είναι άγνωστη.

Βιβλιογραφία

- [1] Aalen O. (1987). Dynamic modelling and causality. *Scand. Actuarial J.* **13**, 177–190.
- [2] Angrist J.D., Imbens G.W., Rubin D.B. (1996). Identification of causal effects using instrumental variables (with discussion and rejoinder). *The Journal of the American Statistical Association* **91**, 444–472.
- [3] Barnard G.A. (1982). Causation. *Encyclopedia of statistical science* **1**, 387–389.
- [4] Box G.E.P. (1966). Use and abuse of regression. *Technometrics* **8**, 625–629.
- [5] Brito C., Pearl J. (2002a). Generalized instrumental variables. in uncertainty in artificial intelligence, proceedings of the eighteenth conference e (a. darwiche and n. friedman, eds.). *Morgan Kaufmann, San Francisco* , 85–93.
- [6] Byrne B. (2006). Structural equation modeling with EQS: Basic concepts, applications, and programming. *2nd ed. Routledge, New York.* .
- [7] Cochran W.G. (1938). The planning of observational studies in human populations (with discussion). *J.R. Statist. Soc. A* **128**, 234–265.
- [8] Cox D.R. & Snell E.J. (1981). Applied statistics. *London: Chapman and Hall* .
- [9] Ettner S.L. (1996). The time of preventive services for women and children: the effect of having a usual source of care. *Am. J. Public Health* **86**, 1748–1754.
- [10] Fisher R.A. (1918). The causes of human variability. *The Eugenics Review* **10**, 213–220.
- [11] Fisher R.A. (1925). Statistical methods for research workers. *Oliver and Boyd, Edinburgh* .
- [12] Fisher R.A. (1926). The arrangement of field experiments. *J. Ministry of Agric.* **33**, 503–513.
- [13] Fisher R.A. (1935). Design of experiments. *Edinburgh: Oliver and Boyd. And subsequent editions* .
- [14] Folkman J. (1996). Tumor angiogenesis and tissue factor. *Nature Med.* **2**, 167–168.

-
- [15] Fox J. (2002). Structural equations models <http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-sems.pdf>.
- [16] Freedman D. (1987). As others see us: A case study in path analysis (with discussion). *Journal of Educational Statistics* **12**, 101–223.
- [17] Galles D., Pearl J. (1998). An axiomatic characterization of causal counterfactuals. *Foundation of Science* **3**, 151–182.
- [18] Good I.J. (1961). A causal calculus, I. *British J. Philosophy of Science* **11**, 305–318.
- [19] Good I.J. (1962). A causal calculus, II. *British J. Philosophy of Science* **12**, 43–51.
- [20] Haavelmo T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica* **11**, 1–12.
- [21] Heckman J.J. (2005). The scientific model of causality. *Sociological methodology* **35**, 1–97.
- [22] Hill A.B. (1965). The environment and disease: association or causation. *Proc. R. Soc. Medicine* **58**, 295–300.
- [23] Holland P.W. (1986). Statistics and causal inference. *The Journal of the American Statistical Association* **81**, 945–970.
- [24] Holland P.W. (1988). Causal inference, path analysis, and recursive structural equations models. In *Sociological Methodology* (C. Clogg, ed.). American Sociological Association, Washington D.C., 449–484.
- [25] Holland R.W., Rubin D.B. (1980). Causal inference in prospective and retrospective studies. *Address given at the Jerome cornfield memorial session of the American statistical Association annual meeting* .
- [26] Holland R.W., Rubin D.B. (1983). On Lord’s paradox. in *Principal of modern psychological measurement* .
- [27] Holland R.W., Rubin D.B. (1995). Some reflections on Freedman’s critiques. *Foundations of Science* **1**, 50–57.

-
- [28] Imbens G.W., Rubin D.B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *Ann. Stat.* **25**, 305–327.
- [29] Kenny D.A., Kashy D.A., Bolger N. (1998). Data analysis in Social Psychology. *The Handbook of Social Psychology* , 223–265.
- [30] Lee S., Hershberger S. (1990). A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research* **25**, 313–334.
- [31] Lewis K.B. (1973). Counterfactuals. *Oxford: Blackwell publishers and cambridge* .
- [32] Little R.J., Rubin D.B. (1987). Statistical analysis with missing data. *New York: Wiley & Sons* .
- [33] Little R.J., Rubin D.B. (2000). Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu. Rev. Public Health* **21**, 121–145.
- [34] Lord F.M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin* **68**, 304–305.
- [35] Mackie J.L. (1974). The cement of the universe. *Oxford: Oxford University Press* .
- [36] Mill J.S. (2012). A system of logic. Collected works of John Stuart Mill. *University of Toronto Press, Toronto* , vol.7.
- [37] Muthen B. (1987). Response to freedman’s critique of path analysis: Improve credibility by better methodological training. *Journal of Educational Statistics* **12**, 178–184.
- [38] Neyman J. (1923). On the application of probability theory to agricultural experiments. English translation from the polish original plus commentary. *Statistical Science* **5**, 465–480.
- [39] Pearl J. (1988). Probabilistic reasoning in intelligent systems. *Morgan Kaufmann, San Mateo* , CA.
- [40] Pearl J. (1993). Comment: Graphical models, causality, and intervention. *Statistical Science* **8**, 266–269.

-
- [41] Pearl J. (2000). Causality: Models, reasoning, and inference. *Cambridge University Press*, New York.
- [42] Pearl J. (2009). Causal inference in statistics: An overview. *Statistics Surveys* **3**, 96–146.
- [43] Pearl J. (2012). The causal foundations of structural equation modeling. *Handbook of Structural Equation Modeling*. New York: Guilford Press. Chapter 5, 68–91.
- [44] Rubin D.B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology* **66**, 688–701.
- [45] Rubin D.B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics* **2**, 1–26.
- [46] Rubin D.B. (1977). Statistical inference. *The International Encyclopedia of Neurology, Psychiatry, Psychoanalysis, and Psychology* **X**, 445–447.
- [47] Rubin D.B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics* **6**, 34–58.
- [48] Rubin D.B. (2004). Basic concepts of statistical inference for causal effects in experiments and observational studies <http://www.stat.columbia.edu/~cook/qr33.pdf>.
- [49] Rubin D.B. (2004). Direct and indirect causal effects via potential outcomes. *Scandinavian Journal of Statistics* **31**, 161–170.
- [50] Rubin D.B. (2008). Statistical inference for causal effects, with emphasis on applications in epidemiology and medical statistics. *Handbook of Statistics* **27**.
- [51] Rubin D.B., Schenker N. (1986). Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *The Journal of the American Statistical Association* **81**, 366–374.
- [52] Rubin D.B., Schenker N. (1991). Multiple imputation in health care databases: an overview and some applications. *The Journal of the American Statistical Association* **10**, 585–598.

- [53] Schweder T. (1970). Composable markov processes. *J. Appl. Prob.* **7**, 400–410.
- [54] Sheiner L.B., Rubin D.B. (1994). Intention-to-treat analysis and the goals of clinical trials. *Clin. Pharm. Ther.* **1**, 6–15.
- [55] Simon H. (1953). Causal ordering and identifiability. *Studies in Econometric Method*, New York: Wiley.
- [56] Smith R.J. (1980). Government says cancer rate is increasing. *Science* **227**, 998–1002.
- [57] Sobel M. (1996). An introduction to causal inference. *Sociological Methods & Research* **24**, 353–379.
- [58] Sobel M. (2008). Identification of causal parameters in randomized studies with mediating variables. *Journal of Educational and Behavioral Statistics* **33**, 230–231.
- [59] Sommer A., Zeger S. (1991). On estimating efficacy from clinical trials. *Stat. Med.* **10**, 45–52.
- [60] Stelzl I. (1986). Changing a causal hypothesis without changing the fit: Some rules for generating equivalent path models. *Multivariate Behavioral Research* **21**, 309–331.
- [61] Suppes P. (1970). A probabilistic theory of causation. *Amsterdam: North Holland*.
- [62] Tanner M.A., Wong W.H. (1987). The calculation of posterior distributions by data augmentation. *The Journal of the American Statistical Association* **82**, 528–550.
- [63] van der Laan M.J. & Robins J.M (2002). Unified methods for censored longitudinal data and causality. *New York: Springer*.
- [64] Wright S. (1921). Correlation and causation. *Journal of Agricultural Research* **20**, 557–585.
- [65] Wright S. (1934). The method of path coefficients. *Ann. Math. Statist.* **5**, 161–215.
- [66] Yule G.U. (1900). On the association of attributes in statistics. *Phil. Trans. Roy. Soc. (London) A* **194**, 257–319.