

ΠΘΑΝΟΚΡΑΤΙΚΗ ΛΙΤΙΟΤΗΤΑ: ΠΛΑΙΣΙΑΚΗ ΟΜΟΦΩΝΙΑ ΚΑΙ ΤΑ ΠΑΡΑΔΟΞΑ ΤΟΥ SIMPSON

των Νίκου Μπισκετζή και Στάθη Ψύλλου*

1. ΕΙΣΑΓΩΓΗ

Τσως, η πλέον γνωστή προσέγγιση στην αιτιότητα είναι του David Hume (1748). Σύμφωνα με τον Hume το συμβάν¹ Α είναι αιτία του συμβάντος Ε όταν:

1. το Α προηγείται χρονικά του Ε
2. το Α και το Ε έχουν χωρική γειτνίαση
3. όλα τα συμβάντα τύπου Α ακολουθούνται από συμβάντα τύπου Ε².

Αυτός ο ορισμός, της αιτίας, ανάγει την αιτιότητα στη χωροχρονική γειτνίαση, στη διαδοχή της αιτίας από το αποτέλεσμα και στη σταθερή σύζευξη ανάμεσα σε αιτία και αποτέλεσμα. Δεν υπάρχει κάποια αναγκαία σύνδεση ανάμεσα στην αιτία και στο αποτέλεσμα, πέρα από την εμφανιζόμενη σχέση κανονικότητας.

Έχει καθιερωθεί οι θεωρίες αιτιότητας, οι οποίες υιοθετούν αυτό τον ορισμό, να ονομάζονται θεωρίες κανονικότητας.

2. Η ΛΙΤΙΟΤΗΤΑ ΩΣ ΚΑΝΟΝΙΚΟΤΗΤΑ

Οι θεωρίες κανονικότητας οφείλουν να αντιμετωπίσουν την ύπαρξη κανονικοτήτων που δεν εκφράζουν αιτιακές σχέσεις. Τέτοιου είδους είναι οι ατυχηματικές κανονικότητες (accidental regularities) και οι ύποπτες συσχέτισεις (spurious correlations).

* Ο Ν. ΜΠΙΣΚΕΤΖΗΣ είναι υπ. διδάκτορας του Τμήματος ΜΙΘΕ του Πανεπιστημίου Αθηνών.

* Ο Σ. ΨΥΛΛΟΣ είναι Αν. Καθηγητής του Τμήματος ΜΙΘΕ του Πανεπιστημίου Αθηνών. Ευχαριστούμε θερμά έναν ανώνυμο κριτή του περιοδικού ΝΕΥΣΙΣ για τις πολύτιμες υποδείξεις του.

¹ Μια περιεκτική εισαγωγή στην αιτιότητα αποτελεί το άρθρο του Σ. Ψύλλου (2005).

² Είναι καθιερωμένο στη βιβλιογραφία τα μικρά γράμματα να δηλώνουν συγκεκριμένες περιπτώσεις συμβάντων, ενώ τα κεφαλαία τύπους συμβάντων. Στο παρόν άρθρο αναφερόμαστε σχεδόν αποκλειστικά σε τύπους συμβάντων, οπότε για ευκολία χρησιμοποιούμε κεφαλαία γράμματα για το συμβολισμό. Επίσης αντί της εκφράσεως «το συμβάν Α κατά το οποίο τα άτομα χ χαρακτηρίζονται από τη κατοχή του παράγοντα ψ, αποτελεί αιτία για το συμβάν Β όπου τα άτομα χ χαρακτηρίζονται από τη κατοχή του παράγοντα z» χρησιμοποιούμε τη φράση «ο παράγων Ψ αποτελεί αιτία για το Z».

Στις ατυχηματικές κανονικότητες, ενώ ένα συμβάν Α ακολουθείται από ένα συμβάν Β, τα Α και Β δεν συνδέονται αιτιακά μεταξύ τους κατά κανέναν τρόπο. Για παράδειγμα, ο Α καθημερινά, πηγαίνοντας στην εργασία του, συναντά τον Β ο οποίος ακολουθώντας το καθημερινό πρόγραμμα του κάνει το πρωινό του τζόγκιν. Εντελώς διασθητικά, αντιλαμβάνεται κανείς ότι δεν υπάρχει αιτιακή σχέση ανάμεσα στα δύο συμβάντα.

Ως «ύποπτη συσχέτιση» νοείται η συσχέτιση δύο συμβάντων για τα οποία όμως ισχύει ότι το ένα δεν αποτελεί κάποιας μορφής αιτία για το άλλο. Συσχετίσεις αυτής της μορφής είναι δυνατόν να εξηγηθούν με αναφορά σε κοινό αίτιο, δηλαδή, με την παραδοχή ότι στο παρελθόν των δύο συμβάντων υπάρχει ένα συμβάν το οποίο συνιστά αιτία και των δύο. Κλασικό παράδειγμα στη βιβλιογραφία είναι το εξής: Η απότομη μείωση της ένδειξης του βαρομέτρου συνήθως ακολουθείται από καταιγίδα. Τα δύο αυτά συμβάντα εμφανίζονται συσχετισμένα. Εάν όμως ληφθεί υπόψη ότι και η πτώση της ένδειξης του βαρομέτρου και η καταιγίδα είναι αποτελέσματα της ίδιας αιτίας, δηλαδή του επερχόμενου βαρομετρικού χαμηλού, η συσχέτιση αυτή εξαλείφεται.

Μια θεωρία αιτιότητας οφείλει να διακρίνει τα προαναφερθέντα είδη κανονικοτήτων. Αν η διάκριση δεν είναι εφικτή, τότε θα αποδίδεται αιτιακή σχέση σε συμβάντα τα οποία είναι αιτιακά ανεξάρτητα. Πρέπει να παρατηρηθεί ότι ο χαρακτηρισμός μιας εμφανιζόμενης κανονικότητας ως ατυχηματικής ή ύποπτης δεν είναι πάντα εύκολο να γίνει.

Σοβαρότερο πρόβλημα αποτελούν οι ονομαζόμενες ατελείς κανονικότητες. Στις ατελείς κανονικότητες το συμβάν Α δεν ακολουθείται πάντα από το συμβάν Β, αλλά κάποιιο ποσοστό των εμφανίσεων του Α ακολουθείται από την εμφάνιση του Β. Οι κανιστές σε μεγάλο ποσοστό εμφανίζουν καρκίνο των πνευμόνων. Φαίνεται ότι υπάρχει αιτιακή σχέση ανάμεσα στους παράγοντες: κάπνισμα και καρκίνο των πνευμόνων. Η έλλειψη όμως σταθερής σύνδεσης ανάμεσα τους καθιστά αδύνατη την αιτιακή συσχέτιση των δύο παραγόντων, με βάση τον ορισμό του Hume.

Στο πλαίσιο της κλασικής αιτιότητας, ο Mackie (1974) ανέπτυξε μια περισσότερο εκλεπτυσμένη θεωρία κανονικότητας. Υιοθετώντας μέρος των θέσεων του John Stuart Mill, δέχεται ότι η πραγματοποίηση ενός συμβάντος στις περισσότερες περιπτώσεις δεν έχει ως αιτία ένα μοναδικό παράγοντα. Δέχεται, μάλιστα, ότι ένα συμβάν μπορεί να πραγματοποιηθεί από μια πλειάδα αιτιών. Δέχεται επίσης ότι για την πραγματοποίηση ενός συμβάντος απαιτείται η ταυτόχρονη παρουσία αριθμού παραγόντων. Με άλλα λόγια η έννοια της αιτίας, υπό τη θεώρηση του Hume, υποκαθίσταται από την έννοια της ολικής αιτίας, η οποία είναι η σύζευξη απλών παραγόντων. Κάθε απλός παράγων είναι μη αναγώγιμος και συσχετισμένος με το αποτέλεσμα, αλλά από μόνος του δεν είναι επαρκής για αυτό. Η συσχέτισή του με το αποτέλεσμα συνίσταται στο ότι αποτελεί τμήμα μιας ευρύτερης σύζευξης παραγόντων, η οποία είναι επαρκής για την πραγματοποίηση του αποτελέσματος, όχι όμως αναγκαία. Ένας απλός παράγων ονομάζεται *inipus* συνθήκη (*insufficient but necessary part of an unecessary but sufficient condition*). Για παράδειγμα, η πυρκαγιά ενός οικήματος μπορεί να οφείλεται σε πολλά διαφορετικά συμπλέγματα παραγόντων. Ένα τέτοιο σύμπλεγμα θα μπορούσε να είναι ένα ηλεκτρικό βραχυκύκλωμα σε συνδυασμό με την παρουσία εύφλεκτων υλικών και την παρουσία οξυγόνου. Αυτή η σύζευξη παραγόντων έχει ως αποτέλεσμα την εκδήλωση της

πυρκαγιάς, αποτελεί δηλαδή μια επαρκή συνθήκη. Δεν αποτελεί όμως και αναγκαία συνθήκη, γιατί η πυρκαγιά μπορεί να εκδηλωθεί και από διαφορετικό σύμπλεγμα παραγόντων, όπως είναι η παρουσία εμπρηστή σε συνδυασμό με τη χρήση πετρελαίου, παρουσία οξυγόνου κ.λπ. Ένας απλός παράγοντας μπορεί να μην καθιστά αναγκαίο το αποτέλεσμα, αλλά υπάρχει πάντα μια σύζευξη παραγόντων η οποία καθιστά την εμφάνιση του αποτελέσματος αναγκαία.

Σύμφωνα με τον Mackie, αποδίδοντας τον όρο «αιτία» σε έναν παράγοντα νοείται ότι ο παράγοντας αυτός αποτελεί τουλάχιστον μια *ipius* συνθήκη. Με αυτή την εκλέπτυνση, η ύπαρξη ατελών κανονικότητας αποδίδεται στην απουσία μιας ή περισσότερων *ipius* συνθηκών. Ενώ όμως αντιμετωπίζει επαρκώς τις ατελείς κανονικότητες, συναντά δυσκολίες διαφορετικής μορφής. Η διάκριση των ύποπτων συσχετίσεων δεν είναι πάντα δυνατή. Αν τα συμβάντα Α και Β είναι αποτελέσματα μιας κοινής αιτίας, μπορεί να δείχθει (Psillos 2002, 84-92) ότι το Α αποτελεί *ipius* συνθήκη για το Β χωρίς προφανώς να αποτελεί αιτία του. Επίσης η έλλειψη σαφήνειας για το χαρακτηρισμό ενός παράγοντα ως *ipius* συνθήκη έγινε αντικείμενο κριτικής από τον Jaegwon Kim (1971). Η πιο σημαντική αντίρρηση που προβάλλει ο Kim είναι ότι ο χαρακτηρισμός ενός παράγοντα ως *ipius* συνθήκης εξαρτάται από την περιγραφή του συμβάντος. Αυτό εισάγει την ανάγκη να καθορίζονται με μεγαλύτερη αυστηρότητα τα είδη των συμβάντων που μετέχουν σε μια αιτιακή σχέση.

3. Η ΠΙΘΑΝΟΚΡΑΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

Εναλλακτικό τρόπο αντιμετώπισης των ατελών κανονικότητων αποτελεί η πιθανοκρατική προσέγγιση της αιτιότητας. Με αυτή την προσέγγιση επιχειρείται η εδραίωση αιτιακής σχέσης, όταν ανάμεσα στην αιτία και στο αποτέλεσμα δεν υπάρχει σταθερή σύζευξη. Η αιτιακή σχέση στην περίπτωση αυτή εκφράζεται σε μαθηματική μορφή, με τη βοήθεια του λογισμού των πιθανοτήτων. Πρέπει να αναφερθεί ότι η υιοθέτηση αυτής της προσέγγισης δεν συνεπάγεται και την απόρριψη του ντετερμινισμού. Αυτός είναι απόλυτα συμβατός με την πιθανοκρατική αιτιότητα, αν ο πιθανοκρατικός χαρακτηρισμός της φύσης θεωρηθεί επιστημικής μορφής, δηλαδή οφειλόμενος στην ανθρώπινη άγνοια για τον κόσμο.

Στον πυρήνα όλων των θεωριών πιθανοκρατικής αιτιότητας βρίσκεται η θέση:

Η πιθανότητα πραγματοποίησης του αποτελέσματος αυξάνεται παρουσία της αιτίας του, δηλαδή:

- (1) Ο παράγοντας Α είναι αιτία του παράγοντα Ε εάν η πιθανότητα πραγματοποίησης του Ε δοθείσης της παρουσίας του Α είναι μεγαλύτερη από την πιθανότητα πραγματοποίησης του Ε στην απουσία του Α, δηλαδή:

$$P(E/A) > P(E/\sim A)$$

Η $P(E/A)$ δηλώνει την πιθανότητα πραγματοποίησης του Ε δοθείσης της παρουσίας

του Α και αντίστοιχα η $P(E/\sim A)$ δηλώνει την πιθανότητα πραγματοποίησης του Ε δοθείσης της απουσίας του Α. Ο συμβολισμός αυτός ακολουθείται και στη συνέχεια του κειμένου.

Σύμφωνα με αυτή τη διατύπωση, ακόμα και μικρή αύξηση της πιθανότητας είναι αρκετή για να χαρακτηρίσει το Α αιτία του Ε.

Όμως η σχέση (1) δεν αποτυπώνει την ασυμμετρία της αιτιακής σχέσης, δηλαδή το γεγονός ότι αν το Α είναι αιτία του Β, τότε το Β δεν μπορεί να είναι αιτία του Α. Αν το Α είναι αιτία του Β τότε ισχύει

$$P(B/A) > P(B/\sim A) \quad (1)$$

Σύμφωνα όμως με το λογισμό πιθανοτήτων, αν ισχύει η (1), τότε θα ισχύει και

$$P(A/B) > P(A/\sim B) \quad (2)$$

δηλαδή και το Β χαρακτηρίζεται ως αιτία του Α. Αυτό οφείλεται στο ότι η σχέση (1) είναι συμμετρική σε αντίθεση με τη σχέση αιτίας-αποτελέσματος. Το «σπάσιμο» αυτής της συμμετρίας επιτυγχάνεται, αν αποδεχθούμε ότι η αιτία προηγείται χρονικά του αποτελέσματος.

Οι εργασίες των Reichenbach (1956, 157-166) και Suppes (1970, 12-47) καθίστουν δυνατή τη σύλληψη των ύποπτων συσχετίσεων με την εισαγωγή της αρχής του κοινού αιτίου. Ο τυπικός ορισμός της αρχής του κοινού αιτίου (Salmon 1984, 159-168) είναι η εξής: Αν δύο συμβάντα είναι στατιστικώς συσχετισμένα, τότε είτε το ένα εξ αυτών συνιστά αιτία του άλλου είτε και τα δύο συμβάντα συνιστούν αποτέλεσμα ενός κοινού αιτίου στο παρελθόν τους.



Σχήμα 1. Αιτιακή δομή του κοινού αιτίου. Ο Γ αποτελεί κοινή αιτία των Α και Β.

Είναι επίσης δυνατός ένας διαφορετικός ορισμός της αρχής του κοινού αιτίου με τη βοήθεια της έννοιας της απόκρυψης (screening off). Ένας παράγοντας Γ αποκρύπτει τον Α από τον Β, όταν:

$$P(B/\Gamma \& A) = P(B/\Gamma \& \sim A)$$

δηλαδή όταν η πιθανότητα πραγματοποίησης του Β παρουσία των παραγόντων Γ και Α είναι ίση με την πιθανότητα πραγματοποίησης του Β παρουσία μόνο του Γ.

Αποτελέσματα του τρόπου με τον οποίο επιδρά ο καθορισμός του πληθυσμού στην τιμή της πιθανότητας εμφάνισης του αποτελέσματος είναι τα παρακάτω του Simpson. Μια οποιαδήποτε σχέση ανάμεσα σε δύο παράγοντες είναι δυνατόν να αντιστραφεί σε κάποιο ειδικό τμήμα του πληθυσμού, βρίσκοντας κάποιον τρίτο παράγοντα, ο οποίος να είναι συσχετισμένος με τους άλλους δύο.

4. ΤΟ ΠΑΡΑΔΟΞΟ ΤΟΥ SIMPSON

Στη στατιστική η εμφανιζόμενη συσχέτιση ανάμεσα σε δύο παράγοντες είναι δυνατόν να αντιστραφεί ως προς κάποιο τμήμα του στατιστικού δείγματος, λόγω της συσχέτισης με έναν τρίτο παράγοντα. Το φαινόμενο αυτό επισημάνθηκε από τον Karl Pearson το 1899 και έγινε γνωστό ως παράδοξο Simpson από την εργασία του τελευταίου «The Interpretation of Interaction in Contingency Tables» το 1951 (Simpson 1951).

Ένα κλασικό παράδειγμα που αναδεικνύει το παράδοξο βασίζεται σε μια αγωγή που κατετέθη κατά του Πανεπιστημίου του Berkeley το 1973, για μεροληψία υπέρ των ανδρών κατά την επιλογή των υποψηφίων.

Στη συνέχεια εξετάζεται μια απλοποιημένη εκδοχή του παραδείγματος. Εξετάζονται δύο μόνο τμήματα, ενώ στα αριθμητικά μεγέθη έχουν δοθεί μικρές τιμές για καλύτερη κατανόηση. Αυτές οι απλοποιήσεις δεν αλλάζουν την ουσία του παραδείγματος. Επίσης γίνεται η παραδοχή ότι οι άνδρες και οι γυναίκες υποψήφιοι είναι ισοδύναμοι ως προς τις γνώσεις, ευφυΐα, προσόντα κ.λπ. Έστω λοιπόν ότι κατά την επιλογή στο Πανεπιστήμιο του Berkeley, η κατανομή των επιτυχόντων σε σχέση με το φύλο των υποψηφίων δίνεται από τον πίνακα 1.

Πίνακας 1: Επιτυχόντες κατά φύλο

Φύλο	Επιτυχόντες	Υποψήφιοι	Ε/Υ	Παρατηρήσεις
Ανδρες	32	57	0,56	Μεγαλύτερη πιθανότητα επιτυχίας των ανδρών.
Γυναίκες	28	57	0,49	

Σύμφωνα με τα δεδομένα του πίνακα 1 φαίνεται ότι το πανεπιστήμιο μεροληπτεί υπέρ των ανδρών υποψηφίων, οι άνδρες έχουν πιθανότητα επιτυχίας 56% ενώ οι γυναίκες 49%. Είναι όμως δυνατόν ο λόγος των ανδρών προς τις γυναίκες να καθορισθεί και για κάθε τμήμα του πανεπιστημίου. Τα δεδομένα αυτά παρουσιάζονται στον πίνακα 2.

Από τον πίνακα 2 προκύπτει ότι σε κάθε τμήμα το ποσοστό επιτυχίας των γυναικών είναι μεγαλύτερο από αυτό των ανδρών. Οι γυναίκες έχουν ποσοστό επιτυχίας 24% στο τμήμα Α1 και 83% στο τμήμα Α2 ενώ οι άνδρες 19% και 78% αντίστοιχα. Παρατηρείται η εμφάνιση του παράδοξου Simpson, δηλαδή η αντιστροφή της σχέσης που ισχύει στο σύνολο των υποψηφίων. Ο λόγος αυτής της αντιστροφής είναι η τάση που εμφανίζεται στις γυναίκες να προτιμούν το τμήμα στο οποίο η εισαγωγή έχει το μεγαλύτερο βαθμό δυσκολίας, στην προκειμένη περίπτωση το Α1, όπου εισάγεται μόλις το 1/5 των

Ένας παράγων Γ αποτελεί κοινό αίτιο των Α και Β, μόνο όταν αποκρύπτει τον Α από τον Β και αντίστροφα. Με άλλα λόγια, το κοινό αίτιο διαλύει την εμφανιζόμενη συσχέτιση ανάμεσα στα Α και Β, καθιστώντας τα στατιστικά ανεξάρτητα.

Πρέπει να σημειωθεί ότι οι συνθήκες απόκρυψης είναι απλώς αναγκαίες, όχι ικανές για το χαρακτηρισμό ενός παράγοντα ως κοινό αίτιο δύο άλλων. Είναι δυνατόν ένας παράγων να ικανοποιεί τις συνθήκες απόκρυψης ως προς δύο συσχετισμένους παράγοντες, χωρίς ωστόσο να μπορεί να χαρακτηριστεί ως κοινό αίτιο αυτών. Σύμφωνα με τον Salmon ένας παράγων που ικανοποιεί τις συνθήκες απόκρυψης θα αποτελεί κοινό αίτιο των συσχετισμένων παραγόντων αν υπάρχουν δύο ανεξάρτητες αιτιακές διαδικασίες που να συνδέουν αντίστοιχα το κοινό αίτιο με καθέναν από τους συσχετισμένους παράγοντες.³

Συνοψίζοντας ο ορισμός της αιτίας μπορεί να αναδιατυπωθεί ως εξής:

(1*) Ο Α είναι αίτια του Ε αν

i) $P(E/A) > P(E/\sim A)$ και

ii) Δεν υπάρχει άλλος παράγων ο οποίος να πραγματοποιείται νωρίτερα ή ταυτόχρονα με τον Α και να αποκρύπτει τον Α από τον Ε.

Η συνθήκη (i) εξασφαλίζει την απαιτούμενη αύξηση της πιθανότητας πραγματοποίησης του αποτελέσματος παρουσία της αιτίας, ενώ η συνθήκη (ii) εξασφαλίζει ότι τα Α και Ε δεν αποτελούν αποτελέσματα κάποιας κοινής αιτίας. Ως συνέπεια η διατύπωση αυτή συλλαμβάνει τις ύποπτες συσχέτισεις. Στο παράδειγμα, που ήδη αναφέρθηκε, της σχέσης ως ανάμεσα στην ένδειξη του βαρομέτρου και στην πιθανότητα να δημιουργηθεί καταιγίδα, η έλευση του βαρομετρικού χαμηλού αποκρύπτει την ένδειξη του βαρομέτρου από την επερχόμενη καταιγίδα. Αυτό γίνεται, διότι η πιθανότητα καταιγίδας δοθέντος ότι πλησιάζει βαρομετρικό χαμηλό είναι η ίδια είτε η ένδειξη του βαρομέτρου μειωθεί είτε όχι.

Ο χαρακτηρισμός ενός παράγοντα ως αιτίας κάποιου άλλου γίνεται με τη βοήθεια μαθηματικών σχέσεων ανάμεσα σε πιθανότητες ελεύθερες ή υπό συνθήκη. Όμως εκφράσεις όπως είναι «πιθανότητα πραγματοποίησης του Ε» ή «πιθανότητα πραγματοποίησης του Ε δοθέντος ότι πραγματοποιείται το Α» αποκτούν νόημα μόνο σε σχέση με το στατιστικό δείγμα από το οποίο εξάγονται αυτές οι πιθανότητες. Αν, όπως παρατηρείται Cartwright (Cartwright 1983, 22-33) «υπάρχουν πολλές αιτίες σε δράση», ή ο Α είναι συσχετισμένος με άλλους παράγοντες, τότε ο τρόπος με τον οποίο καθορίζεται το δείγμα επιδρά στην αριθμητική τιμή της πιθανότητας εμφάνισης του αποτελέσματος.

³ Ο Salmon (1993) προσεγγίζει την αιτιότητα μέσω αιτιακών διαδικασιών και μεταφοράς «στίγματος» (mark transmission). Σύμφωνα με τον Salmon οι βασικές οντότητες που εμπλέκονται σε μια αιτιακή σχέση δεν είναι τα συμβάντα αλλά οι διαδικασίες. Η βασική διαφορά είναι ότι τα συμβάντα είναι χωροχρονικά οριοθετημένα ενώ οι διαδικασίες έχουν και χρονική και χωρική διάρκεια. Διακρίνει ανάμεσα σε αιτιακές και μη αιτιακές διαδικασίες. Αιτιακή διαδικασία είναι εκείνη η οποία είναι σε θέση να μεταφέρει ένα «στίγμα», ενώ μη αιτιακή εκείνη η οποία δεν μπορεί να μεταφέρει.

επιτυχόντων. Συγκεκριμένα υπάρχουν 33 υποψήφιοι στο τμήμα Α1 και 24 στο Α2. Το αντίστροφο ισχύει για τους άνδρες υποψηφίους. Με άλλα λόγια, ο λόγος εμφάνισης του παραδόξου Simpson είναι η ύπαρξη του παράγοντα "τάση επιλογής δύσκολου τμήματος", ο οποίος είναι διαφοροποιημένος στα δύο τμήματα του στατιστικού δείγματος (άνδρες-γυναίκες) και σχετίζεται με την πιθανότητα επιλογής του υποψηφίου.

Πίνακας 2: Επιτυχόντες κατά φύλο και κατά τμήμα

Τμήμα	Άνδρες		Ε/Υ	Γυναίκες		Ε/Υ	Παρατηρήσεις
	Επιτυχόντες	Υποψήφιοι		Επιτυχόντες	Υποψήφιοι		
A1	4	21	0,19	8	33	0,24	Μεγαλύτερη πιθανότητα επιτυχίας των γυναικών
A2	28	36	0,78	20	24	0,83	Μεγαλύτερη πιθανότητα επιτυχίας των γυναικών
Σύνολο	32	57		28	57		

Το παράδοξο Simpson από μαθηματική άποψη δεν είναι και τόσο παράδοξο. Υπάρχουν πάντα αριθμοί οι οποίοι να ικανοποιούν και τις τρεις ανισότητες:

$$\frac{\alpha}{\beta} > \frac{A}{B} \quad (1) \quad \frac{\gamma}{\delta} > \frac{\Gamma}{\Delta} \quad (2) \quad \frac{\alpha+\gamma}{\beta+\delta} < \frac{A+\Gamma}{B+\Delta} \quad (3)$$

Στο παράδειγμα που αναφέραμε τα κλάσματα α/β και Α/Β παρουσιάζουν τα ποσοστά επιτυχίας στο τμήμα Α1 γυναικών και ανδρών αντίστοιχα, ενώ τα γ/δ και Γ/Δ τα αντίστοιχα ποσοστά επιτυχίας για το τμήμα Α2. Λαμβάνοντας τιμές από τους πίνακες 1 και 2 προκύπτει ότι: α=8, β=33, γ=20, δ=24, Α=4, Β=21, Γ=28 και Δ=36 οπότε οι (1), (2) και (3) λαμβάνουν τη μορφή:

$$\frac{8}{33} > \frac{4}{21} \quad (1') \quad \frac{20}{24} > \frac{28}{36} \quad (2') \quad \frac{8+20}{33+24} = \frac{28}{57} < \frac{4+28}{21+36} = \frac{32}{57} \quad (3')$$

Το παράδοξο Simpson κωδικοποιήθηκε στο λογισμό πιθανοτήτων από τον Blyth (1972) ως η ταυτόχρονη επαλήθευση των ανιστήτων (i), (ii) και (iii)

$$\begin{aligned} (i) & P(E/A\&\Gamma) \geq P(E/\Gamma\&\sim A) \\ (ii) & P(E/A\&\sim\Gamma) \geq P(E/\sim\Gamma\&\sim A) \\ (iii) & P(E/A) < P(E/\sim A) \end{aligned}$$

Στη στατιστική οι πιθανότητες υπολογίζονται με τη βοήθεια των στοχαστικών μεταβλητών, δηλαδή συναρτήσεων που απεικονίζουν τα συμβάντα στην ευθεία των πραγματικών αριθμών. Σε μαθηματική γλώσσα η εμφάνιση του παραδόξου οφείλεται στο ότι δύο στοχαστικές μεταβλητές, οι οποίες είναι ανεξάρτητες μεταξύ τους, εμφανίζονται εξαρτημένες ως προς μια τρίτη στοχαστική μεταβλητή. Από μαθηματική άποψη, η κατασκευή μιας στοχαστικής μεταβλητής με τη συγκεκριμένη ιδιότητα είναι πάντοτε εφικτή. Όμως στον ενεργειακό κόσμο τα δεδομένα καθορίζουν τις στοχαστικές μεταβλητές που υπεισέρχονται στην ανάλυση του στατιστικού δείγματος, με αποτέλεσμα η εμφάνιση της αντιστροφής να μην είναι πάντα δυνατή.

Η αντιστροφή της σχέσης που εμφανίζεται ανάμεσα σε δύο παράγοντες Α και Ε λόγω της ύπαρξης ενός τρίτου συσχετισμένου παράγοντα Γ δεν είναι απαραίτητο να καθιστά τη σχέση αυτή ψευδή. Σύμφωνα με το κριτήριο του Cornfield (1959), αν ο Γ καθιστά τη σχέση ανάμεσα στους Α και Ε ψευδή, τότε προκαλεί αύξηση πιθανότητας του Α τουλάχιστον ίση με την αύξηση πιθανότητας του Ε που προκαλεί ο Α. Σε μαθηματική διατύπωση:

$$\text{Κριτήριο του Cornfield: } \frac{P(A/\Gamma)}{P(A/\sim\Gamma)} \geq \frac{P(E/A)}{P(E/\sim A)}$$

Το κριτήριο του Cornfield γίνεται κατανηρό από το εξής παράδειγμα που βασίζεται σε πραγματικά στοιχεία. Το 1958 ο Fisher συγκέντρωσε στοιχεία για την πιθανή σχέση καπνίσματος και γονότυπου. Η επεξεργασία των στοιχείων έδειξε ότι είχαν διαφορετική στάση απέναντι στο κάπνισμα το 51% ανάμεσα σε ετεροζυγωτούς διδύμους, ενώ μόλις το 24% ανάμεσα σε ομοζυγωτούς διδύμους. Αυτό επέτρεπε στον Fisher να συμπεράνει ότι υπάρχει κάποιος γενετικός παράγων υπεύθυνος για τη ροπή προς το κάπνισμα. Αυτό έχει ως συνέπεια η συχνότητα εμφάνισης καρκίνου των πνευμόνων στους καπνιστές να είναι πολύ μεγαλύτερη στους καπνιστές. Αλλά σύμφωνα με το κριτήριο του Cornfield, δοθέντος ότι η πιθανότητα εμφάνισης καρκίνου των πνευμόνων είναι 9 φορές μεγαλύτερη στους καπνιστές από ό,τι στους μη καπνιστές, ο γενετικός παράγων έπρεπε να αυξάνει την πιθανότητα καπνίσματος τουλάχιστον 9 φορές και όχι 2. Υπό αυτή την έννοια ο γενετικός παράγων μπορεί να θεωρηθεί συσχετισμένος τόσο με το κάπνισμα όσο και με τον καρκίνο, σε καμιά όμως περίπτωση δεν αίρει τη συσχέτιση καπνίσματος-καρκίνου πνευμόνων.

Το παράδοξο του Simpson αποτελεί ίσως την καλύτερη επιβεβαίωση της θέσης: κάθε συσχέτιση δεν συνιστά αιτιακή σχέση. Καθιστά την προσπάθεια εξαγωγής αιτιακών σχέσεων από στατιστικά δεδομένα μια διαδικασία με υψηλό βαθμό δυσκολίας. Αν η σχέση ανάμεσα σε δύο παράγοντες είναι δυνατόν να αναστραφεί λόγω ενός τρίτου παράγοντα, τότε η αναφορά σε αιτιακή σχέση γίνεται ιδιαίτερα επισφαλής. Αυτό το οποίο συνάγεται είναι ότι η εγκυρότητα της σχέσης P(E/A)>P(E/~A), η οποία αποτελεί την απαραίτητη προϋπόθεση για το χαρακτηρισμό του Α ως αιτία του Ε, εξαρτάται από το είδος του στατιστικού δείγματος από το οποίο εξήχθη. Μια πιθανοκρατική θεωρία αιτιότητας για να είναι εφαρμόσιμη πρέπει να αντιμετωπίζει τα παράδοξα του Simpson. Αυτό μπορεί να γίνει είτε καθορίζοντας συγκεκριμένους κανόνες με τους οποίους επιλέγεται το στατιστικό δείγμα είτε με το να δομηθεί με τέτοιο τρόπο ώστε να μην προβάλλεται από αυτό.

5. ΘΕΩΡΙΑ ΤΗΣ ΠΛΑΙΣΙΑΚΗΣ ΟΜΟΦΩΝΙΑΣ

Όπως αναφέρθηκε ήδη, έχει διατυπωθεί αριθμός θεωριών για την πιθανοκρατική αιτιότητα. Αυτές διαφέρουν στις προϋποθέσεις που θέτει καθεμία ώστε να διασφαλίζεται η εγκυρότητα της σχέσης $P(E/A) > P(E/\sim A)$, δηλαδή της σχέσης την οποία πρέπει να ικανοποιεί ο A , ώστε να χαρακτηριστεί αιτία του E .

Τα κύρια σημεία διαφοροποίησης είναι:

- Ο τρόπος με τον οποίο κατασκευάζεται ο πληθυσμός από τον οποίο θα ληφθούν τα δεδομένα, δηλαδή αν θα είναι τυχαίος, αν θα αποτελεί ένα τίμιο δείγμα ή αν θα διαμεριστεί σύμφωνα με κάποιον κανόνα.
- Ο τρόπος με τον οποίο αντιμετωπίζονται οι διαφορετικές αιτίες του E από τον υπό εξέταση παράγοντα.
- Ο τρόπος αντιμετώπισης των ύποπτων συσχετίσεων.
- Η ερμηνεία της πιθανότητας.

Η θεωρία της πλαισιακής ομοφωνίας (ΘΠΟ) εκφράστηκε κυρίως από τη Nancy Cartwright (1983, 22-33), τον Ellery Eells (1991, 80-126) και τον Bryan Skyrms (1980, 102-114), όχι όμως με ταυτόσημο τρόπο. Η κεντρική θέση σε όλες τις διατυπώσεις είναι:

Ο χαρακτηρισμός του A ως αιτία του E είναι έγκυρος μόνο αν εξασφαλίζεται ότι η πιθανότητα πραγματοποίησης του E επηρεάζεται αποκλειστικά από τον A . Δεν πρέπει να υπάρχουν αιτιακές συσχετίσεις που να επηρεάζουν την πιθανότητα πραγματοποίησης του E κατά τρόπο ανεξάρτητο από τον τρόπο με τον οποίο επηρεάζεται από τον A . Αυτό επιτυγχάνεται μόνο αν οι ανεξάρτητες από το A αιτίες του E διατηρηθούν σταθερές.

Έστω A_i , $i=1,2,..n$ όλοι οι παράγοντες, εκτός του A , οι οποίοι είναι τέτοιοι, ώστε κάθε A_i να είναι αιτία είτε του E είτε του $\sim E$. Κατασκευάζονται πληθυσμιακά δείγματα (αιτιακώς ομογενή δείγματα) K_j με $j=1,2,.. 2^n$ τέτοια ώστε τα μέλη καθενός από αυτά να χαρακτηρίζονται από ακριβώς την ίδια κατανομή των παραγόντων A_i . Κάθε τέτοιο δείγμα K_j θα χαρακτηρίζεται από μία μέγιστη σύζευξη των παραγόντων A_i ή των αρνήσεών τους και ο αριθμός τους θα είναι 2^n . Σύμφωνα με την ΘΠΟ ο A είναι θετική αιτία του E αν και μόνο αν αυξάνει την πιθανότητα πραγματοποίησης του E σε κάθε αιτιακώς ομογενές δείγμα. Δηλαδή:

ΘΠΟ: Ο παράγων A είναι αιτία του E αν και μόνο αν $P(E/A \& K_j) > P(E/K_j)$ για κάθε K_j .

Αντίστοιχα ορίζεται και η αρνητική αιτία, αν το σύμβολο του μεγαλύτερου αντικατασταθεί από το σύμβολο του μικρότερου.

Η διατύπωση αυτή της αιτιακής σχέσης συλλαμβάνει τις ψευδοσυσχετίσεις. Για παράδειγμα αν συμβολίσουμε την πτώση της ένδειξης του βαρομέτρου με PB , την έλευση του βαρομετρικού χαμηλού με XB και την πραγματοποίηση καταιγίδας με K , τότε σύμφωνα με την ΘΠΟ για να χαρακτηριστεί η μείωση της ένδειξης του βαρομέτρου αιτία για την καταιγίδα θα πρέπει η πιθανότητα καταιγίδας όταν μειώνεται η ένδειξη του βαρομέτρου να είναι μεγαλύτερη από την πιθανότητα καταιγίδας όταν δεν μειώνεται η ένδειξη, για όλα τα δείγματα ελέγχου. Ο μοναδικός παράγων που συσχετίζεται είναι η

έλευση του βαρομετρικού χαμηλού. Επομένως υπάρχουν δύο μόνο δείγματα ελέγχου, εκείνο όπου πλησιάζει βαρομετρικό χαμηλό και εκείνο όπου η ατμοσφαιρική πίεση μεταβάλλεται ομαλά. Θα πρέπει δηλαδή να ισχύει:

$$P(K/PB \& XB) > P(K/XB) \\ P(K/PB \& \sim XB) > P(K/\sim XB)$$

Επειδή όμως ο παράγων XB είναι κοινή αιτία των K και PB , οι ανισώσεις αυτές δεν μπορεί να ισχύουν. Άρα η πτώση της ένδειξης του βαρομέτρου δεν είναι αιτιακά συνδεδεμένη με την καταιγίδα.

Επίσης, η ΘΠΟ, λειτουργεί ικανοποιητικά ακόμα και σε περιπτώσεις όπου μια αιτία φαίνεται να μειώνει την πιθανότητα του αποτελέσματος. Για παράδειγμα, έστω ότι εξετάζεται αν υπάρχει αιτιακή σχέση ανάμεσα στο κάπνισμα και στην πιθανότητα εμφάνισης καρδιακής προσβολής, σε ένα δείγμα ανθρώπων όπου το μεγαλύτερο ποσοστό ασκείται. Αν η σχέση εξεταστεί στο σύνολο του δείγματος, τότε, επειδή η άθληση προστατεύει από την καρδιακή προσβολή, το κάπνισμα εμφανίζεται να μειώνει την πιθανότητα εμφάνισης καρδιακής προσβολής. Αν όμως η σχέση εξεταστεί ξεχωριστά στο τμήμα του δείγματος που ασκείται και στο τμήμα που δεν ασκείται, όπως προτείνει η ΘΠΟ, τότε η σχέση εμφανίζεται να είναι η αναμενόμενη, δηλαδή το κάπνισμα να αυξάνει την πιθανότητα εμφάνισης καρδιακής προσβολής και στα δύο τμήματα. Αυτές οι περιπτώσεις αποτελούν περιπτώσεις εμφάνισης του παραδόξου Simpson και εξετάζονται στη συνέχεια αναλυτικά.

6. ΠΑΡΑΔΟΞΑ SIMPSON ΟΦΕΙΛΟΜΕΝΑ ΣΕ ΣΥΣΧΕΤΙΣΗ ΔΥΟ ΠΑΡΑΓΟΝΤΩΝ

Όπως αναφέρθηκε στην §4, για να εμφανισθεί η αντιστροφή της σχέσης ανάμεσα σε δύο παράγοντες A και E , θα πρέπει να υπάρχει ένας τρίτος συσχετισμένος με αυτούς παράγων. Ο παράγων αυτός Γ χωρίζει το πληθυσμιακό δείγμα στις δύο συμπληρωματικές κλάσεις:

$$A_1 = \{x \mid x \text{ είναι } \Gamma\} \\ A_2 = \{x \mid x \text{ είναι } \sim \Gamma\}$$

Αν για την εξαγωγή της σχέσης ανάμεσα σε A και E δεν συνυπολογιστεί ο διαμερισμός αυτός, τότε στο σύνολο του πληθυσμού η εμφανιζόμενη σχέση θα είναι αντίστροφη της σχέσης που θα εμφανιζόταν, αν λαμβανόταν υπόψη ο διαμερισμός που επιβάλλει ο Γ .

Η ΘΠΟ εξετάζοντας την αιτιακή δράση του A ως προς τον E σε δείγματα, όπου όλοι οι αιτιακώς ανεξάρτητοι παράγοντες από τον A διατηρούνται σταθεροί, αποφεύγει τα παράδοξα που απορρέουν από το συσχετισμό των A και E λόγω κάποιου τρίτου παράγοντα.

Αυτό γίνεται φανερό, αν εξεταστεί το προαναφερθέν παράδειγμα του πανεπιστημίου του Berkeley υπό το πρίσμα της ΘΠΟ.

Έστω ότι ακολουθείται ο συμβολισμός:

- A = άνδρας υποψήφιος
 $\sim A$ = γυναίκα υποψήφιος
 E = επιλογή υποψηφίου
 Γ = τάση για επιλογή τμήματος με μεγάλο βαθμό δυσκολίας
 X_i = διάφοροι παράγοντες όπως είναι ευφυΐα, γνώσεις κ.λπ.

Οι παράγοντες X_i θεωρούνται σταθεροί για όλους τους υποψηφίους, οπότε ο μοναδικός συναφής παράγων είναι ο Γ . Επομένως υπάρχουν δύο μόνο ομογενή αιτιακά δείγματα, το K_{Γ} το οποίο χαρακτηρίζεται από την παρουσία του Γ και το $K_{\sim\Gamma}$ που χαρακτηρίζεται από την απουσία του Γ .

Για να καθορισθεί ο ρόλος του A για τον E , θα πρέπει να εξεταστούν αφενός οι σχέσεις ανάμεσα στις πιθανότητες $P(E/A\&\Gamma)$ και $P(E/\sim A\&\Gamma)$, και αφετέρου $P(E/A\&\sim\Gamma)$, $P(E/\sim A\&\sim\Gamma)$. Στους πίνακες 3 και 4 παρουσιάζονται τα δεδομένα για τα δύο δείγματα.

Πίνακας 3. Δεδομένα για το ομογενές αιτιακό δείγμα K_{Γ}

Τμήμα	Άνδρες		Γυναίκες		$P(E/A\&\Gamma)$	$P(E/\sim A\&\Gamma)$
	Επιτυχόντες	Υποψήφιοι	Επιτυχόντες	Υποψήφιοι		
A1	4	21	5	21		
A2	28	36	30	36		
Σύνολο	32	57	35	57	0,56	0,61

Πίνακας 4. Δεδομένα για το ομογενές αιτιακό δείγμα $K_{\sim\Gamma}$

Τμήμα	Άνδρες		Γυναίκες		$P(E/A\&\sim\Gamma)$	$P(E/\sim A\&\sim\Gamma)$
	Επιτυχόντες	Υποψήφιοι	Επιτυχόντες	Υποψήφιοι		
A1	6	33	8	33		
A2	19	24	20	24		
Σύνολο	25	57	28	57	0,43	0,49

Από τους πίνακες 3 και 4 προκύπτει ότι:

$$\begin{aligned}
 P(E/A\&\Gamma) &= 0,56 & P(E/\sim A\&\Gamma) &= 0,61 \\
 P(E/A\&\sim\Gamma) &= 0,43 & P(E/\sim A\&\sim\Gamma) &= 0,49
 \end{aligned}$$

Οπότε οι ζητούμενες σχέσεις είναι:

$$\begin{aligned}
 P(E/A\&\Gamma) &< P(E/\sim A\&\Gamma) \text{ και} \\
 P(E/\sim A\&\sim\Gamma) &< P(E/A\&\sim\Gamma)
 \end{aligned}$$

Φαίνεται λοιπόν ότι ο παράγων Γ δεν αυξάνει την πιθανότητα επιτυχίας ανδρών υποψηφίων σε κανένα από τα δύο δείγματα. Αντιθέτως τη μειώνει και στα δύο και μπορεί να χαρακτηριστεί ως αρνητικός παράγων για την επιτυχία του υποψηφίου. Δηλαδή το πανεπιστήμιο όχι μόνο δεν μεροληπτεί υπέρ των ανδρών αλλά θα μπορούσε να κατατηγρηθεί για μεροληψία υπέρ των γυναικών.

Έτσι λοιπόν, η στρατηγική της Θ ΠΟ να διατηρεί σταθερούς τους ανεξάρτητους από τον A αιτιακούς παράγοντες για τον E , πέρα από τα όποια επιστημικά προβλήματα θέτει, είναι αποδοτική. Θωρακίζει τη Θ ΠΟ από τα παράδοξα Simpson που προέρχονται από ανεξάρτητους του A παράγοντες ενώ ταυτόχρονα μπορεί και καθορίζει τη σχέση ανάμεσα σε A και E .

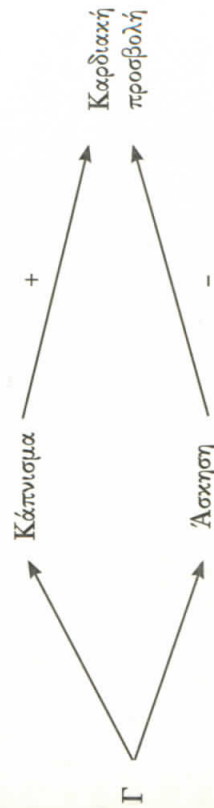
Η κατάσταση δεν είναι η ίδια αναφορικά με τους ενδιάμεσους παράγοντες, δηλαδή τους παράγοντες που αποτελούν τους κρίκους της αιτιακής αλυσίδας από τον A στον E .

7. ΠΑΡΑΔΟΞΑ SIMPSON ΟΦΕΙΛΟΜΕΝΑ ΣΕ ΣΥΣΧΕΤΙΣΗ ΕΝΔΙΑΜΕΣΩΝ ΠΑΡΑΓΟΝΤΩΝ

Η Θ ΠΟ αντιμετωπίζει εντελώς διαφορετικά τους παράγοντες υποβάθρου και τους ενδιάμεσους παράγοντες. Για τους πρώτους απαιτεί ομοφωνία, για τους δεύτερους λαμβάνεται υπόψη το μέσο στατιστικό αποτέλεσμα σε κάθε πλαισιακό υπόβαθρο.

Ο τρόπος αντιμετώπισης των ενδιάμεσων παραγόντων είναι όμοιος με τον τρόπο που αντιμετωπίζει η κλασική πιθανοκρατική θεωρία (αιτία = αύξηση πιθανότητας του αποτελέσματος) όλους τους παράγοντες. Λόγω αυτού του τρόπου αντιμετώπισης εμφανίζονται αντιστροφές σχέσεων στην κλασική θεώρηση. Άρα είναι αναμενόμενο ότι και στη Θ ΠΟ θα είναι δυνατή η εμφάνιση παραδόξων Simpson ως προς τους ενδιάμεσους παράγοντες. Το μέγεθος του προβλήματος αναδεικνύεται από το εξής παράδειγμα που αναφέρει ο Dupre (1984, 1990):

Έστω ότι α υπάρχει ένας γενετικός παράγων, ο οποίος προκαλεί αφενός την τάση για έντονο κάπνισμα και αφετέρου την τάση για άσκηση και β) το κάπνισμα αποτελεί θετική αιτία για να υποστεί κανείς καρδιακή προσβολή, ενώ η άσκηση αρνητική. Το σχήμα 2 αναπαριστά αυτή την αιτιακή δομή.



Σχήμα 2. Αιτιακή δομή του παραδείγματος Dupre

Σε αυτή την περίπτωση, το κάπνισμα και η άσκηση αποτελούν ενδιάμεσους αιτιακούς παράγοντες, για τον καθορισμό της αιτιακής σχέσης ανάμεσα στο γενετικό παράγοντα και στην καρδιακή προσβολή. Σύμφωνα με τη ΘΠΟ δεν πρέπει να κρατηθούν σταθεροί. Επιπλέον ας υποθεθεί ότι, λαμβάνοντας δείγματα από κάποιον πληθυσμό, η αιτιακή αυτή σχέση βρίσκεται θετική.

Το ζήτημα που τίθεται είναι η δυνατότητα εύρεσης κάποιου παράγοντα τέτοιου ώστε, αν ο πληθυσμός διαμεριστεί ως προς αυτόν, να είναι δυνατή η αντιστροφή της αιτιακής σχέσης. Ο παράγων αυτός είναι 'γνώση της χρήσης του καπνού' ο οποίος προφανώς είναι συσχετισμένος με το κάπνισμα, το οποίο αποτελεί ενδιάμεσο αιτιακό παράγοντα του Α και Ε. Αν λοιπόν ο πληθυσμός διαμεριστεί ως προς αυτό τον παράγοντα, τότε στον πληθυσμό που δεν γνωρίζει τη χρήση του καπνού ο γενετικός παράγων θα είναι αρνητική αιτία για την καρδιακή προσβολή. Αν λοιπόν ληφθούν δείγματα και από αυτό τον πληθυσμό, τότε η ευρεθείσα αιτιακή σχέση θα αντιστραφεί. Δηλαδή, ο γενετικός παράγων δεν θα αυξάνει την πιθανότητα εμφάνισης καρδιακής προσβολής σε όλα τα δείγματα, με συνέπεια να μην μπορεί να χαρακτηριστεί ως αιτία.

Παρατηρείται δηλαδή η δυνατότητα εμφάνισης του παραδόξου Simpson αν υπάρχει κάποιος παράγων που είναι συσχετισμένος με κάποιον ενδιάμεσο παράγοντα της αιτιακής αλυσίδας από τον Α στον Ε.

Η ΘΠΟ δεν μπορεί να αντιμετωπίσει αυτού του τύπου τις αντιστροφές. Ο Eells (1987, 1991, 23-34) προσαθεί να απαντήσει, διαμορφώνοντας μια διαφορετική εκδοχή της πλαισιακής ομοφωνίας. Αποδέχεται ότι η αιτιακή σχέση είναι μια σχέση στην οποία εμπλέκονται τέσσερις παράμετροι: ο παράγων αιτία, ο παράγων αποτέλεσμα, ο πληθυσμός δειγμάτων στον οποίο εξετάζεται η αιτιακή σχέση και τέλος το είδος του πληθυσμού δειγμάτων. Σύμφωνα με τον Eells το είδος του πληθυσμού καθορίζει σε σημαντικό βαθμό την τιμή της πιθανότητας πραγματοποίησης του αποτελέσματος. Συνεπώς το είδος του πληθυσμού θα πρέπει να πληροί κάποιες προϋποθέσεις. Συγκεκριμένα, το είδος του πληθυσμού θα πρέπει να επιτρέπει και την παρουσία ενός παράγοντα που παράγοντα ο οποίος αποτελεί την υπό εξέταση αιτία. Επίσης οι τιμές $P(E/\Delta\&K)$ και $P(E/\sim\Delta\&K)$ όπως και οι σχέσεις ανάμεσά τους δεν πρέπει να αποτελούν λογική συνέπεια του γεγονότος ότι ο πληθυσμός εκφράζει ένα συγκεκριμένο είδος. Η συσχέτιση της αιτιακής σχέσης με το είδος του πληθυσμού επιτρέπει το χαρακτηρισμό ενός παράγοντα ως μεικτή αιτία, δηλαδή ο Α μπορεί να είναι θετική αιτία για τον Ε ως προς ένα πληθυσμό είδους M1 και αρνητική αιτία ως προς έναν πληθυσμό είδους M2.

Με αυτή την οπτική, στο αναφερθέν παράδειγμα, ο γενετικός παράγων χαρακτηρίζεται ως μεικτή αιτία για την καρδιακή προσβολή. Θετική για το είδος πληθυσμού όπου το κάπνισμα είναι δυνατό και αρνητική για το είδος του πληθυσμού, όπου το κάπνισμα δεν είναι δυνατό.

Ο ελιγμός αυτός έχει κόστος. Στο εν λόγω παράδειγμα, ο συνολικός πληθυσμός διαχωρίζεται σε δύο συμπληρωματικές κλάσεις. Το κριτήριο διαχωρισμού είναι η δυνατότητα του καπνίσματος, το οποίο αποτελεί ενδιάμεσο παράγοντα. Δηλαδή, εμμέσως, η διερεύνηση της αιτιακής σχέσης του γενετικού παράγοντα και της καρδιακής προσβολής γίνεται σε δείγματα που διατηρούν έναν ενδιάμεσο παράγοντα σταθερό. Αν όμως αυτό είναι θεμιτό για έναν ενδιάμεσο παράγοντα, τότε είναι θεμιτό και για τους

υπόλοιτους. Αν επιτραπεί το είδος του πληθυσμού να χαρακτηρίζεται από την κατοχή ή όχι ενός ενδιάμεσου παράγοντα, η κατάσταση γίνεται αρκετά περίπλοκη. Στην πραγματικότητα επιβάλλεται ένας νέος λεπτότερος διαμερισμός του πληθυσμού. Συγκεκριμένα κάθε ομογενές αιτιακό δείγμα είναι δυνατόν να διαμεριστεί σε κλάσεις K_m , στις οποίες οι ενδιάμεσοι παράγοντες θα διατηρούνται σταθεροί. Έτσι αν ένας παράγων Α χαρακτηρίζεται αιτία του Ε, αυτό σημαίνει ότι έχει την ίδια επίδραση στην πιθανότητα πραγματοποίησης του Ε για κάθε δείγμα ελέγχου K_i και για κάθε διαμερισμό K_m , σε διαφορετική περίπτωση θα χαρακτηρίζεται ως μεικτή αιτία. Με αυτό τον τρόπο επιβάλλεται στη ΘΠΟ τόσο οι ενδιάμεσες αιτίες όσο και οι αιτίες του Ε οι ανεξάρτητες από τον Α να αντιμετωπίζονται με τον ίδιο τρόπο. Αν αυτό δεν επιτραπεί, το προηγούμενο παράδειγμα διατηρεί όλη του την ισχύ.

Ο Eells αναφέρει ρητά ότι δεν πρέπει να διατηρούνται σταθεροί οι ενδιάμεσοι παράγοντες. Το επιχείρημα το οποίο προβάλλει είναι το εξής (Eells 1991): Έστω ότι ο Α αποτελεί αιτία του Ε και ο μοναδικός τρόπος με τον οποίο επιτυγχάνεται αυτό είναι η παραγωγή από τον Α ενός ενδιάμεσου παράγοντα Ζ, ο οποίος με τη σειρά του παράγει το Ε (σχήμα 3). Κατ' αυτό τον τρόπο η αύξηση πιθανότητας του Ε επιτυγχάνεται μέσω της αύξησης πιθανότητας του Ζ. Αν ο Ζ κρατηθεί σταθερός σε ένα αιτιακά ομογενές δείγμα, η πιθανότητα πραγματοποίησης του Ζ στο δείγμα αυτό θα είναι 0 ή 1 και στην παρουσία και στην απουσία του Α. Σε αυτή τη περίπτωση ο Α δεν θα μπορεί να μεταβάλει την πιθανότητα πραγματοποίησης του Ζ και εξ αυτού την πιθανότητα πραγματοποίησης του Ε, δηλαδή θα χαρακτηριστεί αιτιακά ουδέτερος ως προς τον Ε, κάτι το οποίο δεν είναι αληθές δηλαδή, κρατώντας ενδιάμεσους παράγοντες σταθερούς αλλοιώνεται η αιτιακή δράση του Α προς το Ε ή ακόμα σε ορισμένες περιπτώσεις αναιρείται πλήρως.



Σχήμα 3. Αιτιακή δομή περίπτωσης όπου ο Α αυξάνει την πιθανότητα πραγματοποίησης του αποτελέσματος, αυξάνοντας την πιθανότητα πραγματοποίησης ενός ενδιάμεσου παράγοντα

Έτσι λοιπόν, ενώ ο Eells απαιτεί να μη διατηρούνται σταθεροί οι ενδιάμεσοι παράγοντες όμως επιτρέποντας το είδος ενός πληθυσμού να χαρακτηρίζεται από την έλλειψη ή την κατοχή ενός ενδιάμεσου παράγοντα, δεν κάνει τίποτε άλλο παρά να επιτρέπει τη διατήρηση ενός ενδιάμεσου παράγοντα σταθερού.

Ακολουθώντας αυτή την συλλογιστική, αν υπάρχουν n ενδιάμεσοι παράγοντες στην αιτιακή αλυσίδα η οποία συνδέει την αιτία και το αποτέλεσμα, τότε κάθε ομογενές αιτιακό δείγμα είναι δυνατόν να ληφθεί από 2m διαφορετικά πληθυσμιακά είδη, το κάθε ένα από τα οποία χαρακτηρίζεται από τη διατήρηση ενός ενδιάμεσου παράγοντα σταθερού. Ως συνέπεια ο Α είναι δυνατόν να συνιστά θετική αιτία για τον Ε ως προς ένα πληθυσμό, αρνητική ως προς άλλον, να είναι αιτιακά ουδέτερος ως προς τρίτο κ.λπ., δηλαδή ο αριθμός των μεικτών αιτιών αυξάνεται δραματικά.

Φαίνεται λοιπόν ότι ούτε η διατύπωση του Eells μπορεί να δώσει κάποια λύση στην περίπτωση εμφάνισης αντιστροφών λόγω κάποιου συσχετισμού με ενδιάμεσο παράγοντα. Είναι δυνατόν όμως να υπάρξουν και αντιστροφές, δηλαδή παράδοξα του Simpson,

που οφείλονται στην ίδια τη δομή της ΘΠΟ. Πριν όμως εξεταστεί αυτή η περίπτωση θα πρέπει, έστω συνοπτικά, να αναλυθεί η έννοια της αρνητικής αιτίας.

Με τον όρο «αρνητική αιτία» νοείται ένας παράγοντας παρουσία του οποίου η πιθανότητα εμφάνισης του αποτελέσματος μειώνεται. Πολλές φορές η αρνητική αιτία δηλώνει την απουσία της αιτίας. Για παράδειγμα: «η απροσεξία του οδηγού αποτελεί αρνητική αιτία για την ασφαλή οδήγηση». Εδώ η λέξη «απροσεξία» υποδηλώνει την απουσία προσοχής του οδηγού που αποτελεί προϋπόθεση για ασφαλή οδήγηση. Αυτή η θεωρητική αρνητικής αιτίας δεν είναι εφαρμόσιμη σε κάθε περίπτωση· η φράση «η άσκηση αποτελεί αρνητική αιτία για την καρδιακή προσβολή» σε καμιά περίπτωση δεν μπορεί να εκληφθεί ότι δηλώνει την απουσία των αιτιών της καρδιακής προσβολής. Λέγοντας ότι παρουσία της αρνητικής αιτίας το αποτέλεσμα δεν πραγματοποιείται, σημαίνει ότι υπό την προϋπόθεση να είναι παρούσες οι θετικές αιτίες για το αποτέλεσμα, απουσίας της αρνητικής αιτίας τα αποτελέσματα θα πραγματοποιούνταν. Αυτό με τη σειρά του φανερώνει ότι παράλληλα με την αρνητική αιτία βρίσκονται σε δράση και θετικές αιτίες. Συνέπεια είναι ότι για να δράσει μια αρνητική αιτία θα πρέπει παράλληλα να δρουν και θετικές αιτίες για το αποτέλεσμα, διαφορετικά θα είναι αιτιακά αδρανής. Ως παράδειγμα, επεξηγηματικό των ανωτέρω μπορεί να θεωρηθεί το εξής: έστω ότι κάποιος είναι αλλεργικός στην επαφή του με τριαντάφυλλα, με συνέπεια να παρουσιάζει κοκκινίλες στο πρόσωπό του. Για να αντιμετωπίσει την αλλεργία κάνει, ως υποθέσουμε, ένεση κορτιζόνης. Εάν κάνει την ένεση και έρθει σε επαφή με τριαντάφυλλα, τότε η κορτιζόνη θα δράσει και δεν θα εμφανίσει κοκκινίλες· αν όμως δεν έρθει σε επαφή με τριαντάφυλλα, η κορτιζόνη δεν θα έχει κανένα αποτέλεσμα ως προς τις κοκκινίλες, θα είναι αιτιακά αδρανής.

Μετά από αυτά ας εξετασθεί το εξής παράδειγμα:

Έστω ότι:

- (i) Ένα συγκεκριμένος τύπος καρκίνου των πνευμόνων X προκαλείται είτε από την έκθεση σε αμιάντο λόγω της εισπνοής κόκκων αμιάντου (παράγον Α) είτε ως αποτέλεσμα του καπνίσματος. Επίσης δεν υπάρχει άλλη αιτία ανάπτυξης του συγκεκριμένου τύπου καρκίνου, ενώ η άσκηση αποτελεί αρνητική αιτία για την ανάπτυξη του συγκεκριμένου τύπου καρκίνου.
- (ii) Υπάρχει ένας γενετικός παράγον Γ, ο οποίος προκαλεί στο 50% των φορέων του παράγοντα την τάση για κάπνισμα και στο υπόλοιπο 50% την τάση για άθληση. Ο παράγον αυτός δεν προκαλεί σε καμιά περίπτωση την τάση και για κάπνισμα και για άθληση.
- (iii) Η αύξηση της πιθανότητας εμφάνισης καρκίνου στην περίπτωση καπνιστών που παράλληλα εκτίθενται σε αμιάντο είναι μικρότερη από τη μείωση της πιθανότητας εμφάνισης καρκίνου στην περίπτωση ασκούμενων που εκτίθενται σε αμιάντο.
- (iv) Η παρουσία ή απουσία του παράγοντα Α «εισπνοή κόκκων αμιάντου» δεν επηρεάζει την πιθανότητα να είναι κάποιος καπνιστής (Κ) ή αθλητής (Ε) λόγω του γενετικού παράγοντα. Με αυτό γίνεται προφανές ότι ισχύει η ακόλουθη σχέση ανάμεσα στις πιθανότητες:

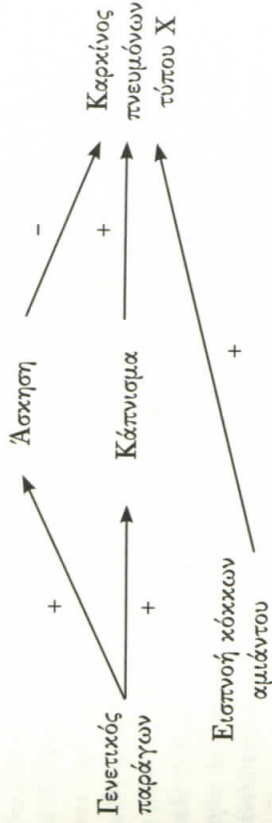
$$P(K/\Gamma) = P(K/\Gamma \& A) = P(K/\Gamma \& \sim A) = 0.5$$
και αντίστοιχα για τον παράγοντα Ε.

(v) Ο γενετικός παράγον (Γ) είναι η μοναδική αιτία για το κάπνισμα και την άσκηση. Δηλαδή ισχύει:

$$P(K/\sim\Gamma) = P(K/\sim\Gamma \& A) = P(K/\sim\Gamma \& \sim A) = 0$$

και αντίστοιχα για τον παράγοντα Ε.

Η αιτιακή δομή του παραδείγματος παρίσταται στο σχήμα 4.



Σχήμα 4. Αιτιακή δομή περίπτωσης αλληλεπίδρασης με ενδιάμεσο αιτιακό παράγοντα.

Σύμφωνα με τη ΘΠΟ ο παράγον «έκθεση σε αμιάντο» είναι ανεξάρτητος από το γενετικό παράγοντα, άρα διατηρείται σταθερός. Επειδή δεν υπάρχουν άλλοι αιτιακώς συχετισμένοι παράγοντες, θα υπάρχουν μόνο δύο αιτιακά δείγματα, το δείγμα του οποίου τα μέλη εκτίθενται σε αμιάντο και το δείγμα του οποίου τα μέλη δεν εκτίθενται σε αμιάντο. Για να χαρακτηριστεί ο γενετικός παράγον ως αιτία του καρκίνου των πνευμόνων τύπου X, θα πρέπει να αυξάνει την πιθανότητα εμφάνισής του και στα δύο δείγματα.

Στο δείγμα του οποίου τα μέλη εκτίθενται σε αμιάντο:

Ο γενετικός παράγον δημιουργεί τους ενδιάμεσους παράγοντες «κάπνισμα» και «άθληση», επομένως θα υπάρχουν οι συζεύξεις {«κάπνισμα» & «έκθεση σε αμιάντο»} και {«άθληση» & «έκθεση σε αμιάντο»}. Λόγω της υπόθεσης (iii) η σύζευξη {«άθληση» & «έκθεση σε αμιάντο»} θα μειώσει την πιθανότητα ανάπτυξης καρκίνου περισσότερο από όσο την αυξήσει η σύζευξη {«κάπνισμα» & «έκθεση σε αμιάντο»} άρα:

$$P(X/A \& \Gamma) < P(X/A \& \sim \Gamma)$$

Στο δείγμα του οποίου τα μέλη δεν εκτίθενται σε αμιάντο:

Όπως αναφέρθηκε, ο γενετικός παράγον δημιουργεί τους ενδιάμεσους παράγοντες «κάπνισμα» και «άθληση». Επειδή δεν υπάρχει έκθεση σε αμιάντο, δεν εντοπίζεται στο δείγμα άλλη θετική αιτία για την εμφάνιση καρκίνου πνευμόνων πέραν του καπνίσματος. Η άθληση δεν μπορεί να επιδράσει στην πιθανότητα εμφάνισης καρκίνου, γιατί εμφανίζεται σε άτομα τα οποία ούτε καπνίζουν ούτε εκτίθενται σε αμιάντο, άρα σε μέλη του πληθυσμού στα οποία δεν είναι δυνατόν να αναπτυχθεί ο καρκίνος του συγκεκριμένου τύπου. Οπότε σε δράση είναι μόνο ο παράγον «κάπνισμα». Προφανώς $P(X/\sim A \& \sim \Gamma) = 0$ επειδή ελλείπει κάθε αιτία του X, $P(X/\sim A \& \Gamma) > 0$ επειδή είναι σε δράση μόνο ο παράγον «κάπνισμα». Άρα:

$$P(X/\sim A\&G) > P(X/\sim A\&\sim G)$$

Σύμφωνα λοιπόν με τη ΘΠΟ, ο Γ δεν μπορεί να χαρακτηριστεί ως αιτία του X, επειδή δεν αυξάνει την πιθανότητα πραγματοποίησης του X σε κάθε αιτιακό δείγμα. Αυτό διαισθητικά δεν φαίνεται σωστό, αφού ο γενετικός παράγων από μόνος του είναι ικανός για την ανάπτυξη του συγκεκριμένου τύπου καρκίνου των πνευμόνων.

Στα παραδείγματα τα οποία εξετάστηκαν, δηλαδή γενετικού παράγοντα καρδιακής προσβολής (σχήμα 2) και γενετικού παράγοντα καρκίνου πνευμόνων τύπου X (σχήμα 3), η αιτιακή αλυσίδα που συνδέει αιτία και αποτέλεσμα έχει δύο κλάδους, στον ένα εκ των οποίων υπάρχει αρνητική αιτία για το αποτέλεσμα. Στο πρώτο η εμφάνιση του παραδόξου του Simpson οφείλεται στο διαμερισμό του πληθυσμού με τέτοιο τρόπο, ώστε να καθίσταται αδρανής ο ένας από τους δύο κλάδους της αιτιακής αλυσίδας. Στο δεύτερο παράδειγμα η εμφάνιση του παραδόξου οφείλεται σε διαφορετικούς λόγους. Ο τρόπος με τον οποίο συσχετίζεται ο ανεξάρτητος από την αιτία παράγων (έκθεση σε αμίαντο) με τους ενδιάμεσους παράγοντες των δύο κλάδων της αιτιακής αλυσίδας (κάπνισμα, άσκηση) είναι τέτοιος ώστε στο δείγμα που χαρακτηρίζεται από την παρουσία του, ο υπό εξέταση παράγων (γενετικός παράγων) να προξενεί μείωση της πιθανότητας πραγματοποίησης του αποτελέσματος. Στο δείγμα που χαρακτηρίζεται από την απουσία του ο κλάδος που περιέχει την αρνητική αιτία καθίσταται ανενεργός, με αποτέλεσμα ο υπό εξέταση παράγων να προξενεί αύξηση της πιθανότητας πραγματοποίησης του αποτελέσματος.

8. ΣΥΜΠΕΡΑΣΜΑΤΑ

Εξετάστηκε η αντιμετώπιση των ύποπτων συσχετίσεων και των παραδόξων του Simpson από τη θεωρία της πλαισιακής ομοφωνίας.

Οι ύποπτες συσχετίσεις συλλαμβάνονται από τη ΘΠΟ κατά τρόπο ικανοποιητικό. Αυτό βέβαια ήταν εφικτό και από την κλασική πιθανοκρατική προσέγγιση, η διαφορά συνίσταται στον κομψότερο τρόπο με τον οποίο το επιτυγχάνει η ΘΠΟ. Ο έλεγχος των ύποπτων συσχετίσεων γίνεται με συστηματικότερο τρόπο μέσω της απαίτησης της αύξησης της πιθανότητας πραγματοποίησης του αποτελέσματος σε κάθε δείγμα ελέγχου. Επίσης δεν είναι αναγκαία η αναφορά σε μια ανεξάρτητη από τον ορισμό αρχή, όπως η αρχή του κοινού αιτίου.

Τα παράδοξα του Simpson αντιμετωπίζονται με επιτυχία, όταν οφείλονται σε συσχέτιση του αποτελέσματος με έναν παράγοντα ανεξάρτητο της υπό εξέταση αιτίας.

Δεν συμβαίνει όμως το ίδιο, όταν τα παράδοξα του Simpson οφείλονται σε συσχετίσεις όπου εμπλέκονται οι ενδιάμεσοι αιτιακοί παράγοντες. Ειδικότερα, στις περιπτώσεις όπου η αιτιακή αλυσίδα παρουσιάζει διακλαδώσεις σε μια εκ των οποίων υπάρχει αρνητική αιτία για το αποτέλεσμα, η αντιμετώπιση των παραδόξων δεν είναι επαρκής. Σε αυτές τις περιπτώσεις η ΘΠΟ χαρακτηρίζει τους υπό εξέταση παράγοντες είτε ως μεικτές αιτίες σύμφωνα με τον Eells είτε αδυνατεί πλήρως να τους χαρακτηρίσει.

Φαίνεται λοιπόν ότι η ΘΠΟ μπορεί να αντιμετωπίσει ικανοποιητικώς τα παράδοξα

Simpson, λόγω συσχετισμού ενός παράγοντα ανεξάρτητου της υπό εξέταση αιτίας με κάποιον ενδιάμεσο παράγοντα της αιτιακής αλυσίδας η οποία συνδέει την αιτία και το αποτέλεσμα, μόνο στην περίπτωση όπου στην αιτιακή αλυσίδα δεν υπάρχουν διακλαδώσεις.

BIBΛΙΟΓΡΑΦΙΑ

- BLYTH, C. R., (1972): «On Simpson's Paradox and the Sure Thing Principle», *Journal of the American Statistical Association* 67 (Theory and Methods Section).
- CARTWRIGHT, N. (1983): *How the Lows of Physics Lie*, Oxford: Clarendon Press.
- CORNFIELD, J., HAENSZEL, W., HAMMOND, E., LILIENTHAL, A., SHIMKIN, M., and WYNDER, E. (1959): «Smoking and lung cancer: Recent evidence and a discussion of some questions», *Journal of the National Cancer Institute*, 22, 173-203.
- DUPRE, J. (1984): «Probabilistic Causality Emancipated», *Midwest Studies in Philosophy* IX, 169-175.
- DUPRE, J. (1990): «Probabilistic Causality: A Rejoinder to Ellery Eells», *Philosophy of Science* 57, 690-698.
- EELLS, E. (1991): *Probabilistic Causality*, Cambridge University Press, 80-126, 23-34.
- EELLS, E. (1987): «Probabilistic Causality: Reply to John Dupre», *Philosophy of Science* 54, 105-114.
- HUME, D. (1748): «An Enquiry Concerning Human Understanding», στο *Enquiries Concerning Human Understanding and Concerning, the Principles of Morals*, L.A. Selby-Biggl & P.H. Nidditch (eds), Oxford: Clarendon Press 1974.
- KIM, J. (1971): «Causes and Events: Mackie on Causation», *Journal of Philosophy* 68: 429-41. Reprinted in E. SOSA & M. TOOLEY (eds): *Causation*, Oxford: Oxford University Press, 1993, 205-7.
- MACKIE, J. L. (1974): *The Cement of the Universe: A Study of Causation*, Oxford: Clarendon Press, 87-92.
- PSILLOS, S. (2002): *Causation & Explanation*, Acumen Publishing Limited.
- PSILLOS, S. (2005): «Αιτιότητα», *Αρχαία Ελληνική Ιατρική*, 22, 594-605.
- REICHENBACH, H. (1956): *The Direction of Time*, University of California Press.
- SALMON, W.C. (1984): *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press, 158-168.
- SALMON, W.C. (1993): «Causality: Production and Propagation» στο *Causation* edited by E. SOSA and M. TOOLEY, Oxford University Press.
- SIMPSON, E. H., (1951): «The interpretation of interaction in contingency tables», *Journal of the Royal Statistical Society, Series B*, 13, 238-241.
- SKYRMS, B. (1980): *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*, New Heaven: Yale University Press.
- SUPPES, P. (1970): *A probabilistic theory of causality*, North-Holland Publishing Company Amsterdam.