



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Ανάπτυξη μοντέλου διερεύνησης των παραγόντων που επηρεάζουν την συμμόρφωση στις παρεμβάσεις για την διαχείριση της παιδικής παχυσαρκίας

Μελέτη και υλοποίηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΕΛΕΥΘΕΡΙΟΥ Χ. ΣΟΦΡΑ



Επιβλέπουσα: Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Νοέμβριος 2021



**Ανάπτυξη μοντέλου διερεύνησης των παραγόντων
που επηρεάζουν την συμμόρφωση στις παρεμβάσεις
για την διαχείριση της παιδικής παχυσαρκίας**

Μελέτη και υλοποίηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΕΛΕΥΘΕΡΙΟΥ Χ. ΣΟΦΡΑ

Επιβλέπουσα: Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 4η Νοεμβρίου 2021.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

.....
Ανδρέας Γεώργιος Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

.....
Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Copyright © - All rights reserved. Με την επιφύλαξη παντός δικαιώματος.

Ελευθέριος Σοφράς, 2021.

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

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

ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ ΚΑΙ ΑΝΑΛΗΨΗΣ ΠΡΟΣΩΠΙΚΗΣ ΕΥΘΥΝΗΣ

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ευνοπογράφως ότι είμαι αποκλειστικός συγγραφέας της παρούσας Πτυχιακής Εργασίας, για την ολοκλήρωση της οποίας κάθε βοήθεια είναι πλήρως αναγνωρισμένη και αναφέρεται λεπτομερώς στην εργασία αυτή. Έχω αναφέρει πλήρως και με σαφείς αναφορές, όλες τις πηγές χρήσης δεδομένων, απόψεων, θέσεων και προτάσεων, ιδεών και λεκτικών αναφορών, είτε κατά κυριολεξία είτε βάσει επιστημονικής παράφρασης. Αναλαμβάνω την προσωπική και ατομική ευθύνη ότι σε περίπτωση αποτυχίας στην υλοποίηση των ανωτέρω δηλωθέντων στοιχείων, είμαι υπόλογος έναντι λογοκλοπής, γεγονός που σημαίνει αποτυχία στην Πτυχιακή μου Εργασία και κατά συνέπεια αποτυχία απόκτησης του Τίτλου Σπουδών, πέραν των λοιπών συνεπειών του νόμου περί πνευματικών δικαιωμάτων. Δηλώνω, συνεπώς, ότι αυτή η Πτυχιακή Εργασία προετοιμάστηκε και ολοκληρώθηκε από εμένα προσωπικά και αποκλειστικά και ότι, αναλαμβάνω πλήρως όλες τις συνέπειες του νόμου στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής άλλης πνευματικής ιδιοκτησίας.

(Υπογραφή)

.....
Ελευθέριος Σοφράς

04/11/2021

Περίληψη

Η παιδική παχυσαρκία αποτελεί ένα παγκόσμιο πρόβλημα δημόσιας υγείας, εξαιτίας του οποίου εκατομμύρια παιδιά υφίστανται χαμηλής ποιότητας ζωή ενώ μακροπρόθεσμα βρίσκονται σε υψηλό κίνδυνο εμφάνισης σοβαρών νοσημάτων όπως ο Σακχαρώδης Διαβήτης, η αποφρακτική άπνοια και άλλα. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, η Ελλάδα και άλλες χώρες της Νότιας Ευρώπης παρουσιάζουν τα υψηλότερα ποσοστά παιδικής παχυσαρκίας στην Ευρώπη. Εκτός από τους γενετικούς παράγοντες που επηρεάζουν την εμφάνιση της παιδικής παχυσαρκίας σημαντικό ρόλο παίζουν και οι περιβαλλοντικοί παράγοντες, συμπεριλαμβανομένου του οικογενειακού περιβάλλοντος. Αντικείμενο της παρούσας διπλωματικής εργασίας είναι η διερεύνηση συσχετίσεων μεταξύ του οικογενειακού περιβάλλοντος και της παιδικής παχυσαρκίας με απώτερο στόχο την διαστρωμάτωση των παιδιών που θα επιτρέψει την εφαρμογή εξατομικευμένων και πιο αποτελεσματικών παρεμβάσεων. Προς αυτήν την κατεύθυνση, εφαρμόστηκαν τεχνικές μηχανικές μάθησης και πιο συγκεκριμένα Μπείσιανή Ανάλυση Δικτύου σε δεδομένα που παραχωρήθηκαν από την Μονάδα Αναπτυξιακής και Συμπεριφορικής Παιδιατρικής του Νοσοκομείου Παιδών «Αγία Σοφία». Τα δεδομένα περιλαμβάνουν πληροφορίες σχετικά με το βάρος, την γονεϊκή τυπολογία, τις γονεϊκές πρακτικές σίτισης και τις ειδικές συμπεριφορές υγείας τόσο των γονέων όσο και των παιδιών. Από αυτά τα δεδομένα, εξήχθησαν τα χαρακτηριστικά των γονέων και παιδιών τα οποία αποτέλεσαν τους κόμβους του Μπείσιανού Δικτύου. Η ισχύς της επιρροής μεταξύ των κόμβων υπολογίστηκε χρησιμοποιώντας την μέγιστη και μέση παρατηρούμενη τιμή της J-Απόκλισης, ενώ μέσω της ανάλυσης ευαισθησίας διερευνήθηκαν οι πιο σημαντικές μεταβολές των χαρακτηριστικών που επηρεάζουν το βάρος των παιδιών.

Λέξεις Κλειδιά

Μπείσιανή Ανάλυση Δικτύου, Παιδική Παχυσαρκία, Γονεϊκή Τυπολογία, Γονεϊκές Πρακτικές Σίτισης, Μπείσιανός Συμπερασμός, Συσχέτιση, Μη Ισορροπημένα Δεδομένα, J-Απόκλιση

Abstract

Childhood obesity is a major public health problem across the globe, due to which millions of children suffer from low quality of life and in the long run are at high risk of developing serious diseases such as diabetes mellitus, obstructive sleep apnea and others. According to the World Health Organization, Greece and other southern European countries have the highest rates of childhood obesity in Europe. The most common causes of childhood obesity are genetic and environmental factors, including family environment. The objective of the present thesis is to identify correlations between family environment and childhood obesity with the ultimate goal to stratify children enabling the development of personalized and effective interventions. Towards this direction, a Bayesian Network Analysis approach has been applied on data granted from the Division of Developmental and Behavioral Pediatrics of the Agia Sophia Children's Hospital. Data include information related to the child's weight, parenting styles, feeding practices, and health behaviors of parents and children. The features extracted by these data were modelled as the nodes of the Bayesian Network. The strength of influence between the nodes was computed using J-Divergence, while through applying sensitivity analysis the most important feature changes on the child's weight were revealed.

Keywords

Bayesian Network Analysis, Childhood Obesity, Parenting Styles, Parental Food Practices, Bayesian Inference, Correlation, Unbalanced Data, J-Divergence

στους γονείς μου

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά την επιβλέπουσα καθηγήτρια αυτής της διπλωματικής, κα. Κωνσταντίνα Νικήτα, για την ευκαιρία που μου έδωσε να ασχοληθώ με το συγκεκριμένο θέμα, το ερέθισμα που μου προσέφερε να γνωρίσω τον τομέα της υγείας, καθώς και για το ενδιαφέρον που μου καλλιέργησε κατά τη διάρκεια των σπουδών μου.

Ιδιαίτερες ευχαριστίες θα ήθελα να δώσω στην ερευνήτρια κα. Κωνσταντία Ζαρκογιάννη και την υποψήφια διδάκτορα Μαρία Αθανασίου για την καθοδήγησή τους και τη διαρκή και άμεση στήριξή τους. Τόσο οι τεχνικές όσο και οι θεωρητικές συζητήσεις μας πάνω στο θέμα της διπλωματικής αυτής αλλά και άλλων εργασιών μου δίδαξαν πολλά και ενίσχυσαν το ενδιαφέρον μου να συνεχίσω συνεχώς να μαθαίνω.

Ευχαριστώ, ακόμα, τα μέλη της επιτροπής, κ. Ανδρέα - Γεώργιο Σταφυλοπάτη και κ. Γεώργιο Στάμου, όχι μόνο ως εξεταστές, αλλά και για όσα πολύτιμα μου προσέφεραν ως καθηγητές μου.

Έπειτα, θέλω να ευχαριστήσω τους φίλους και κοντινούς μου ανθρώπους, με τους οποίους είμαστε συνοδοιπόροι όλα αυτά τα αξέχαστα χρόνια. Ευχαριστώ ιδιαίτερα την Ελένη, που υπήρξε το μεγαλύτερο στήριγμά μου.

Τέλος, το μεγαλύτερο ευχαριστώ οφείλω στην οικογένεια μου, στους γονείς μου, τον αδελφό μου, την αδελφή μου και την γιαγιά μου, για την ανιδιοτελή στήριξη και αγάπη τους όλα αυτά τα χρόνια.

Η παρούσα διπλωματική εργασία πραγματοποιήθηκε στο πλαίσιο του έργου ENDORSE (Τ1ΕΔΚ-03695).

Αθήνα, Νοέμβριος 2021

Ελευθέριος Σοφράς

Περιεχόμενα

Περίληψη	1
Abstract	3
Ευχαριστίες	7
1 Εισαγωγή	19
1.1 Αντικείμενο	20
1.2 Οργάνωση της Εργασίας	21
2 Παιδική Παχυσαρκία	23
2.1 Ο Ρόλος του Παιδιού	24
2.1.1 Κληρονομικότητα	24
2.1.2 Περιγεννητικοί Παράγοντες	24
2.1.3 Τρόπος Ζωής των Παιδιών	25
2.2 Ο Ρόλος της Οικογένειας	27
2.2.1 Γονεϊκός Έλεγχος	27
2.2.2 Γονεϊκός Τύπος	28
2.2.3 Γονεϊκές Πρακτικές Σίτισης	30
2.3 Ερευνητικές Προκλήσεις	31
3 Μοντελοποίηση της Παιδικής Παχυσαρκίας	33
3.1 Προηγηθείσες Παρατηρήσεις	33
4 Μπεϊσιανή Ανάλυση	37
4.1 Γνωσιακό Υπόβαθρο	37
4.1.1 Νόμος του Μπέις	37
4.1.2 Μπεϊσιανά Δίκτυα	38
4.1.3 Μπεϊσιανός Συμπερασμός	41
4.1.4 Διαγράμματα Επιρροής	41
4.1.5 Τύποι Ερμηνευσιμότητας - Επεξηγήσεων στα Μπεϊσιανά Δίκτυα	43
4.1.6 Αποδίδοντας Ερμηνεύσεις - Επεξηγήσεις	44
4.2 Αρχές Γενικού Σχεδιασμού	44
4.2.1 Πρόγονοι και απόγονοι	44
4.2.2 Μανδύας Markov	45
4.2.3 Διατεταγμένοι Κόμβοι	45

4.3	Δυναμική Στρατηγική	46
4.4	Σύγκριση των Κατανομών	49
4.5	Μετρικές Αποστάσεων	51
4.5.1	Ευκλείδεια Απόσταση	51
4.5.2	Απόσταση Hellinger	53
4.5.3	Απόσταση Kullback - Leibler	54
4.5.4	J - Απόκλιση	55
4.5.5	Απόσταση Συναρτήσεων Αθροιστικής Κατανομής	56
4.5.6	Συμπεράσματα	61
5	Μεθοδολογία και Υλοποίηση	63
5.1	Λογισμικό SMILE	63
5.2	Σύνολο Δεδομένων	64
5.2.1	Γονεϊκή Τυπολογία Μητέρων	64
5.2.2	Γονεϊκές Πρακτικές Σίτισης	65
5.2.3	Άλλες Κατηγορίες Δεδομένων	65
5.2.4	Αποτελέσματα Στατιστικής Ανάλυσης	66
5.3	Προεπεξεργασία Δεδομένων	66
5.3.1	Ελλιπείς Τιμές	66
5.3.2	Διακριτοποίηση	66
5.4	Κατασκευή Μπεϊσιανού Δικτύου	67
5.4.1	Αλγόριθμοι Μπεϊσιανής Μάθησης	68
5.4.2	Παραμετροποίηση	69
5.5	Μπεϊσιανός Συμπερασμός	70
5.5.1	Ισχύς Επιρροής	70
5.5.2	Ανάλυση Ευαισθησίας Δικτύου	71
5.6	Αξιολόγηση Μοντέλου	71
5.6.1	Συνάρτηση Log Likelihood	71
5.6.2	Διασταυρώμενη Επικύρωση	72
6	Αποτελέσματα και Συζήτηση	73
6.1	Παραγόμενα Μπεϊσιανά Δίκτυα	74
6.2	Αποτελέσματα Ισχύος Επιρροής στα Μπεϊσιανά Δίκτυα	76
6.3	Αποτελέσματα Ανάλυσης Ευαισθησίας στα Μπεϊσιανά Δίκτυα	80
7	Επίλογος	87
7.1	Συμπεράσματα	87
7.2	Μελλοντικές Επεκτάσεις	89
	Παράρτημα	91
	Α' Χρήσιμες Έννοιες	93
A'.1	P-Value	93
A'.2	Z-Score	93

Α'.2.1 BMI z-score	94
Βιβλιογραφία	105

Κατάλογος Σχημάτων

4.1	Τύποι Σύνδεσης σε ΚΑΓ	40
4.2	Παράδειγμα Διαγράμματος Επιρροής	42
4.3	Μανδύας Markov κόμβου A	45
4.4	Παράδειγμα Διατεταγμένων Κόμβων	46
4.5	Τέσσερις διαφορετικές καταστάσεις	47
4.6	Ένα μικρό δίκτυο	48
4.7	Απλό δίκτυο τριών κόμβων	51
4.8	Ευκλείδειες αποστάσεις της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$	52
4.9	Ευκλείδεια απόσταση της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	52
4.10	Ευκλείδεια απόσταση της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	53
4.11	Αποστάσεις Hellinger της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$	53
4.12	Απόσταση Hellinger της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	54
4.13	Απόσταση Hellinger της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	54
4.14	Αποστάσεις Kullback-Leibler της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$	55
4.15	Απόσταση Kullback-Leibler της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	55
4.16	Απόσταση Kullback-Leibler της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	56
4.17	Κανονικοποιημένη J-Απόκλιση με $\alpha = 10$ της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$	57

4.18	Κανονικοποιημένη J-Απόκλιση με $a = 10$ της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	57
4.19	Κανονικοποιημένη J-Απόκλιση με $a = 10$ της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	57
4.20	ΣΑΚ των κατανομών P (αριστερά) και Q (δεξιά)	58
4.21	ΣΑΚ της κατανομής R	58
4.22	Συνδυασμένες κατανομές P και Q (αριστερά) και P και R (δεξιά)	59
4.23	Αποστάσεις ΣΑΚ της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$	60
4.24	Αποστάσεις ΣΑΚ της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	60
4.25	Αποστάσεις ΣΑΚ της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0	60
5.1	Κατευθυνόμενο Ακυκλικό Γράφημα (ΚΑΓ)	67

Κατάλογος Εικόνων

4.1	Παράδειγμα Μπεϊσιανού Δικτύου	39
4.2	Κατανομή Πιθανοτήτων του κόμβου "Επιτυχία του εγχειρήματος"	42
4.3	Αναμενόμενη Ωφέλεια του κόμβου αξίας "Οικονομικό Όφελος"	43
5.1	Εκ των Υστέρων Πιθανότητα	71
6.1	Μπεϊσιανό Δίκτυο Γονεϊκών Πρακτικών Σίτισης στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)	74
6.2	Μπεϊσιανά Δίκτυα Γονεϊκών Τύπων στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)	74
6.3	Μπεϊσιανά Δίκτυα Κοινωνικοδημογραφικών Χαρακτηριστικών των Μητέρων στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)	74
6.4	Μπεϊσιανά Δίκτυα Περιγεννητικών Παραγόντων Κινδύνου στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)	75
6.5	Μπεϊσιανά Δίκτυα Σωματομετρικών Χαρακτηριστικών των Γονέων στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)	75
6.6	Μπεϊσιανά Δίκτυα του Τρόπου Ζωής των Γονέων στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)	75
6.7	Μπεϊσιανά Δίκτυα του Τρόπου Ζωής των Παιδιών στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)	75
6.8	Διαγράμματα Ευαισθησίας των Γονεϊκών Πρακτικών Σίτισης στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)	82
6.9	Διαγράμματα Ευαισθησίας των Γονεϊκών Συμπεριφορών στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)	82
6.10	Διαγράμματα Ευαισθησίας των Κοινωνικοδημογραφικών Χαρακτηριστικών των Γονέων στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)	83
6.11	Διαγράμματα Ευαισθησίας των Περιγεννητικών Παραγόντων Κινδύνου στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)	83
6.12	Διαγράμματα Ευαισθησίας των Σωματομετρικών Χαρακτηριστικών των Γονέων στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)	84

- 6.13 Διαγράμματα Ευαισθησίας του Τρόπου Ζωής των Γονέων στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω) 85
- 6.14 Διαγράμματα Ευαισθησίας του Τρόπου Ζωής των Παιδιών στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω) 85

Κατάλογος Πινάκων

4.1	Πίνακας δεσμευμένης πιθανότητας για το δίκτυο του σχήματος 4.4. Βρίσκοντας τον A επιπροσθέτως με τον B, ή το ανάποδο, θα έχει μικρή επίδραση στον C.	47
4.2	Άλλος πίνακας δεσμευμένων πιθανοτήτων για το δίκτυο 4.4. Η συνδυαστική επίδραση των γονεϊκών κόμβων είναι πολύ μεγαλύτερη από το άθροισμα των μεμονωμένων επιδράσεων.	49
4.3	Σύνοψη μετρικών αποστάσεων.	61
6.1	Καταγραφή της J-Απόκλισης (με $\alpha = 1$) που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία για το σύνολο των Κοριτσιών. Οι συσχετίσεις έχουν υπολογιστεί για την μέγιστη και μέση τιμή.	77
6.2	Καταγραφή της J-Απόκλισης (με $\alpha = 1$) που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία για το σύνολο των Αγοριών. Οι συσχετίσεις έχουν υπολογιστεί για την μέγιστη και μέση τιμή.	78
6.3	Καταγραφή των Συσχετίσεων που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία για το σύνολο των Κοριτσιών. Οι συσχετίσεις έχουν υπολογιστεί για την μέγιστη τιμή.	79
6.4	Καταγραφή των Συσχετίσεων που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία για το σύνολο των Αγοριών. Οι συσχετίσεις έχουν υπολογιστεί για την μέγιστη τιμή.	80
6.5	Καταγραφή της Ανάλυσης Ευαισθησίας, της Ακρίβειας και της συνάρτησης EM Log Likelihood για το σύνολο των Κοριτσιών. Η ευαισθησία των ΔΜΣ και ΔΜΣ ζ-βαθμολογίας έχει υπολογιστεί για την μέγιστη τιμή.	81
6.6	Καταγραφή της Ανάλυσης Ευαισθησίας, της Ακρίβειας και της συνάρτησης EM Log Likelihood για το σύνολο των Αγοριών. Η ευαισθησία των ΔΜΣ και ΔΜΣ ζ-βαθμολογίας έχει υπολογιστεί για την μέγιστη τιμή.	81

Κεφάλαιο **1**

Εισαγωγή

Εχοντας άμεση σχέση με τον διαβήτη τύπου ΙΙ, τα ισχαιμικά εγκεφαλικά, τις καρδιακές παθήσεις και πληθώρα καρκίνων, η παιδική παχυσαρκία αντιπροσωπεύει μία από τις πιο σοβαρές σύγχρονες απειλές για την υγεία και την ποιότητα ζωής των παιδιών[1, 2, 3, 4, 5, 6], αλλά και την μετέπειτα ζωή τους ως έφηβοι και ενήλικες. Η παιδική παχυσαρκία συνοδεύεται με μια ζωή με συνεχείς ανησυχίες για την υγεία, που κυμαίνονται από το αυξημένο κόστος υγειονομικής περίθαλψης[4, 7, 8], έως την επιβάρυνση με επιπρόσθετο στρες στη δυναμική της οικογένειας[9, 10].

Παρά την παγκόσμια επίγνωση επί του ζητήματος, η επικράτηση των υπέρβαρων και παχύσαρκων παιδιών (ηλικίας 2-19 ετών) αυξήθηκε 47,1% παγκοσμίως σε διάστημα 33 ετών, μεταξύ 1980 και 2013[11]. Το 2010, ο Παγκόσμιος Οργανισμός Υγείας υπολόγισε ότι αριθμός των υπέρβαρων παιδιών ηλικίας κάτω των 5 ετών είναι πάνω από 42 εκατομμύρια, με σχεδόν τα 35 εκατομμύρια από αυτά να ζουν σε αναπτυσσόμενες χώρες [12]. Στην πραγματικότητα, ο επιπολασμός της παχυσαρκίας αυξάνεται τόσο στις ανεπτυγμένες όσο και στις αναπτυσσόμενες χώρες[13]. Υπέρβαρα και παχύσαρκα παιδιά είναι πιθανό να παραμείνουν παχύσαρκα στην ενήλικη ζωή, γεγονός που ενισχύει τον κίνδυνο για ένα πλήθος ασθενειών όπως οι καρδιαγγειακές κ.α.[14]. Η συνεχώς αυξανόμενη επιβάρυνση της παχυσαρκίας είναι ανησυχητική λόγω του ευρέος φάσματος οξείων και μακροπρόθεσμων επιπλοκών που επιβάλλουν σημαντικό ανθρώπινο αλλά και οικονομικό κόστος για τα ίδια τα άτομα, τις οικογένειές τους, το εγχώριο σύστημα υγειονομικής περίθαλψης και εν τέλη την κοινωνία γενικότερα[15]. Μια μελέτη δείχνει ότι τα ποσοστά αυξανόμενης παιδικής παχυσαρκίας θα μειώσουν το μέσο προσδόκιμο ζωής στις ΗΠΑ κατά 2-5 χρόνια στις επόμενες δεκαετίες, ούτως ώστε η τρέχουσα γενιά να μπορεί δυνητικά να γίνει η πρώτη από τον Β΄ Παγκόσμιο Πόλεμο και έπειτα που θα έχει μικρότερο προσδόκιμο ζωής από τους γονείς του[16]. Από άλλους η παχυσαρκία θα μπορούσε να συγκριθεί σε σημαντικότητα με την αρτηριοσκλήρωση, αφού η πρώτη είναι μία από τους πολλούς παράγοντες κινδύνου για την δεύτερη. Για όλους τους παραπάνω λόγους, η εύρεση παραγόντων που συμβάλλουν σε αποτελεσματικές στρατηγικές πρόληψης και προγράμματα για την αντιμετώπιση αυτής της μάστιγας στην δημόσια υγεία χρήζει άμεσης ανάγκης[17].

Κάποιοι παιδίατροι αντιμετωπίζουν την παιδική παχυσαρκία περισσότερο ως ψυχολογική, παρά σωματική ασθένεια, αισθητική διαταραχή ή ανεπάρκεια του σώματος[18]. Αφετηρία για αυτόν τον ισχυρισμό αποτελεί η πολυαιτιότητα και η χρονική της πορεία. Συχνά η παιδική παχυσαρκία αποτελεί γενεσιουργό ασθένεια. Το παχύσαρκο παιδί έχει από την

πολύ αρχή της ζωής του διαφορετική ψυχοκινητική ανάπτυξη και εσφαλμένη αυτο-εικόνα για το ίδιο, συναισθηματικές διαταραχές και διαταραχές στη διαδικασία κοινωνικοποίησης. Απόρροια αυτού είναι ότι, τα αίτια της παιδικής παχυσαρκίας δεν μπορούν να οριστούν εύκολα[19]. Στην εργασία μας η παχυσαρκία μπορεί να οριστεί ως πρωτογενής νόσος, καθώς λαμβάνονται υπόψη διάφορες φυσιολογικές αλλαγές εξωγενούς προέλευσης. Η δευτερογενής παχυσαρκία για παράδειγμα συνδέεται με γενετικές ασθένειες και άλλα κλινικά σύνδρομα. Προέλευση της πρωτογενούς παχυσαρκίας δεν μπορεί να οριστεί ακόμη, οπότε η καταπολέμηση της είναι παρόμοια με την πρόληψη των πυρκαγιών, όπου παρά τους ανιχνευτές πυρκαγιάς, εξακολουθούν να υπάρχουν υλικά ανθεκτικά στη φωτιά, εκπαίδευση των ανθρώπων και πυροσβεστικές δυνάμεις. Στον αντίποδα, πρόληψη νόσων όπως η ευλογία ή η ελονοσία είναι δράση που κατευθύνεται απευθείας κατά του γνωστού βακίλου. Αν αυτό εξαλειφθεί, η ασθένεια δεν θα ξεσπάσει. Ως εκ τούτου, η πρόσβαση στην παιδική παχυσαρκία είναι πολυ-επιστημονική. Όσες μονόπλευρες μέθοδοι χρησιμοποιήθηκαν μέχρι τώρα ήταν ανεπιτυχείς για τους ακριβώς παραπάνω λόγους. Πολλές μελέτες εξετάζουν περισσότερο (κι έτσι ενδείκνυται), τη φαινομενολογία της παχυσαρκίας από την πηγή της[20, 21, 22].

1.1 Αντικείμενο

Από τα παραπάνω καθίσταται σαφής και επιτακτική η ανάγκη για έγκαιρη διάγνωση της υπερβαρότητας και της παχυσαρκίας στα παιδιά με την επιτυχή πρόβλεψη ανοδικής ή καθοδικής πορείας του σωματικού βάρους του παιδιού, ώστε να μειωθούν στο ελάχιστο δυνατό και αποτελεσματικά οι ποικίλες συνέπειες στην ψυχοσωματική ζωή του ατόμου και του περιγυρού του, μέχρι και την αποτροπή της θνησιμότητας. Στο παρελθόν, οι ερευνητές έχουν υιοθετήσει διάφορες στρατηγικές εξαγωγής πολυτροπικών χαρακτηριστικών για να πετύχουν τον συμπερασμό πολυποίκιλων και περιεκτικών αποτελεσμάτων ποσοτικής απεικόνισης και χρησιμοποιούν διάφορες μεθόδους μηχανικής μάθησης για τον έγκαιρο εντοπισμό του προβλήματος. Ωστόσο, οι υπάρχουσες τεχνικές στερούνται λεπτομερούς ανάλυσης του μεγέθους της συσχέτισης μεταξύ διαφορετικών μεμονωμένων διακριτών χαρακτηριστικών, τα οποία είναι πολύ σημαντικά για την κατανόηση της δυναμικής της νόσου[23].

Σε αυτή τη μελέτη, ενθαρρυνόμαστε από την προσεγγίσεις άλλων ερευνητών και χρησιμοποιούμε δεδομένα 48 υπέρβαρων και παχύσαρκων παιδιών που παραχωρήθηκαν από το Νοσοκομείο Παιδών "Η Αγία Σοφία" προκειμένου να εξάγουμε πληροφορία ούτως ώστε να κατηγοριοποιήσουμε το προφίλ των ασθενών, υλοποιώντας προσέγγιση "Bayesian Network Analysis", (εφεξής: "Μπεϊσιανή Ανάλυση Δικτύου"), με απώτερο και τελικό σκοπό την ποσοτικοποίηση της συσχέτισης μεταξύ διαφορετικών χαρακτηριστικών και την εύρεση της ισχύς της συσχέτισης αυτής (Strength of Influence). Τα χαρακτηριστικά και η συσχέτιση μεταξύ των χαρακτηριστικών, διαμορφώθηκαν ως κόμβοι και ακμές του δικτύου αντίστοιχα[23].

Μείζονα μετρική αποτελεί ο Δείκτης Μάζας Σώματος (Body Mass Index), εφεξής: ΔΜΣ (BMI), που αποτελεί μια απλή μέτρηση που χρησιμοποιεί την αναλογία βάρους προς ύψος και χρησιμοποιείται για την ταξινόμηση παιδιών και ενηλίκων σε λιπόβαρους, υπέρβαρους ή παχύσαρκους. Ο ΔΜΣ υπολογίζεται ως το βάρος του ατόμου σε κιλά (kg) διαιρούμενο με το τετράγωνο του ύψους του σε μέτρα (m)[24].

$$BMI = kg/m^2 \quad (1.1)$$

Καθώς ο ΔΜΣ αυξάνεται, ιδιαίτερα σε περιπτώσεις όπου ο ΔΜΣ ξεπερνά το 30 (σε ενήλικες), οι κίνδυνοι για την υγεία αυξάνονται επίσης ανάλογα[25]. Ένας υψηλότερος ΔΜΣ φαίνεται ότι σχετίζεται άμεσα με τον αυξημένο κίνδυνο πλήθους ασθενειών όπως η καρδιαγγειακή νόσος (cardiovascular heart disease, CHD), διαβήτη τύπου 2 και ορισμένων τύπων καρκίνου[26]. Έκθεση του Εθνικού Προγράμματος Εκπαίδευσης για τη Χοληστερόλη στο Τμήμα Θεραπείας Ενηλίκων III (National Cholesterol Education Program's Adult Treatment Panel III, NCEP ATP III)[27], αναγνώρισε το μεταβολικό σύνδρομο ως πολλαπλό παράγοντα κινδύνου για την καρδιαγγειακή νόσο καθώς και για τον διαβήτη. Ως εκ τούτου, είναι σημαντικό να δοθεί περισσότερη κλινική προσοχή σε αυτούς τους παράγοντες, καθώς περισσότεροι ασθενείς με μεταβολικό σύνδρομο είναι παχύσαρκοι ή υπέρβαροι[28]. Για να εφαρμόσουμε περισσότερη αποτελεσματική καθοδήγηση για τις δημόσιες αρχές υγείας, είναι πολύ σημαντικό να κατανοήσουμε τα πολύπλοκα συστήματα αλληλένδετων επιρροών στον ΔΜΣ[29].

1.2 Οργάνωση της Εργασίας

Η παρούσα εργασία είναι οργανωμένη σε συνολικά επτά κεφάλαια. Στο πρώτο κεφάλαιο γίνεται μια πρώτη ανάγνωση σχετικά με την παχυσαρκία. Στα κεφάλαια 2, 3 και 4 παρουσιάζεται το θεωρητικό υπόβαθρο των βασικών γνωστικών αντικειμένων που σχετίζονται με τη διπλωματική αυτή και οργανώνουμε τον στόχο μας. Αρχικά, στο κεφάλαιο 2 περιγράφεται ο ρόλος των παιδιών και της οικογένειας στην εμφάνιση και διατήρηση της παιδικής παχυσαρκίας. Στο κεφάλαιο 3 παρουσιάζονται σχετικές με το θέμα μελέτες και παρατηρήσεις που αφορούν τις γονεϊκές πρακτικές σίτισης και τους γονεϊκούς τύπους και την σχέση τους με την παιδική παχυσαρκία. Εν συνεχεία στο κεφάλαιο 4 γράφουμε τις αρχικές και απαραίτητες έννοιες της μπεϊσιανής ανάλυσης που θα ακολουθήσει αργότερα. Στα κεφάλαια 5 και 6 καταγράφουμε το πρακτικό μέρος της παρούσας μελέτης. Πρώτα περιγράφονται, στο κεφάλαιο 5, η μεθοδολογία που ακολουθούμε, ενώ στο κεφάλαιο 6 καταγράφουμε τα αποτελέσματα. Τέλος, στο κεφάλαιο 7 δίνεται η συνεισφορά της εργασίας αυτής, καθώς και οι μελλοντικές επεκτάσεις.

Κεφάλαιο 2

Παιδική Παχυσαρκία

Στο κεφάλαιο αυτό παρουσιάζονται οι παράγοντες κινδύνου που καλείται η ερευνητική κοινότητα να λάβει υπόψιν προκειμένου να αντιμετωπίσει την υπερβαρότητα και την παχυσαρκία, καθώς και αναλύεται η βαρύτητα και σημαντικότητα των παραγόντων αυτών προκειμένου να επικεντρωθεί η έρευνά μας, με γνώμονα πρόσφατες διεθνείς και εγχώριες μελέτες.

Η τεράστια ανάγκη αντιμετώπισης της παχυσαρκίας, και πόσο μάλλον της παιδικής, οδήγησε τόσο επαγγελματίες όσο και ερευνητές να καταβάλουν προσπάθειες για να αναπτύξουν μια ευρεία συλλογή από επιχειρήσεις που στοχεύουν στον αυξανόμενο πληθυσμό των παιδιών σε κίνδυνο[30]. Παραδοσιακά, αυτές οι παρεμβάσεις επικεντρώθηκαν αποκλειστικά στις διατροφικές πτυχές της ζωής ενός παιδιού. Ωστόσο, η πρόσφατη βιβλιογραφία έχει δείξει ότι οι ρίζες της παχυσαρκίας είναι πολύ πιο γενικές και περίπλοκες. Η πολυπλοκότητα αυτή ίσως αποτυπώνεται καλύτερα στο έργο των Puder και Munsch το 2010, στο οποίο σημειώνονται τα εξής: «Αντίθετα με μια σταθερή και στατική κατάσταση, η παιδική παχυσαρκία αντιπροσωπεύει μια κατεξοχήν δυναμική διαδικασία, κατά την οποία η συμπεριφορά, η γνωστική λειτουργία και η συναισθηματική ρύθμιση αλληλεπιδρούν αμοιβαία μεταξύ τους, τόσο με βιολογικές παραμέτρους, όσο και με παράγοντες συμφραζομένων, όπως οι στάσεις των γονέων και η οικογενειακή διατροφή, η φυσική δραστηριότητα και τα διατροφικά πρότυπα»[31]. Με βάση αυτά τα ευρήματα, το πεδίο των παραγόντων που καλούνται να αντιμετωπίσουν οι παρεμβάσεις διευρύνεται γρήγορα. Μεταξύ των περισσότερων παραγόντων, εξέχουσα ήταν η συμπερίληψη των ψυχολογικών χαρακτηριστικών, τόσο από το παιδί όσο και από την οικογένειά του[1, 31, 32, 33, 34] και ιδιαίτερα οι:

1. γονεϊκές πρακτικές σίτισης και (Parental Feeding Practises)
2. ο γονεϊκός τύπος (Parental Style), γνωστό και με τον όρο γονική μέριμνα

Ως αποτέλεσμα, οι αναδυόμενες ερευνητικές μελέτες έδωσαν, έκτοτε, στοιχεία σχετικά με την ύπαρξη συσχετίσεων μεταξύ συγκεκριμένων ψυχολογικών χαρακτηριστικών και της γονεϊκής επιρροής με την παρουσία παχυσαρκίας στα παιδιά[1, 33, 35, 36]. Ωστόσο, οι περισσότερες από αυτές έχουν επικεντρωθεί στην μονόπλευρη ανάλυση των χαρακτηριστικών σε σχέση με την παιδική παχυσαρκία[32, 34, 35, 37].

2.1 Ο Ρόλος του Παιδιού

2.1.1 Κληρονομικότητα

Η παιδική παχυσαρκία έχει εδραιωθεί από την παγκόσμια επιστημονική κοινότητα ως ένα πολύπλοκο και πολυαιτιακό ζήτημα, με την κληρονομικότητα και τους γενετικούς παράγοντες να κατέχουν εξέχουσα θέση[38]. Το 99% των περιπτώσεων παχυσαρκίας είναι αποτέλεσμα διαταραχής στο ενεργειακό ισοζύγιο, το οποίο ισοδυναμεί με μειωμένη φυσική δραστηριότητα σε συνδυασμό με υπερβολική λήψη τροφής, σε γενετικά προδιατεθειμένα άτομα, σύμφωνα με την πλειοψηφία ερευνητών που ασχολούνται τα τελευταία χρόνια με την παχυσαρκία. Σε αυτό το μήκος κύματος, έχουν αναγνωριστεί πάνω από 200 γονίδια που σχετίζονται με αυξημένο βάρος μέχρι τις μέρες μας[39]. Για παράδειγμα, πολύ πρόσφατη μετα-ανάλυση γενετικών μελετών για τον ΔΜΣ με περισσότερους από 339.000 συμμετέχοντες οδήγησε στον εντοπισμό 97 γενετικών περιοχών που επηρεάζουν τον κίνδυνο εμφάνισης παχυσαρκίας. Σύμφωνα με την μελέτη αυτή, τα εμπλεκόμενα γονίδια σχετίζονται σε ένα βαθμό με το νευρικό σύστημα και κυρίως με τα σήματα του εγκεφάλου που ελέγχουν την όρεξη και την ενεργειακή κατανάλωση του σώματος[40]. Άλλες μελέτες που εφαρμόστηκαν σε δίδυμα συμπεραίνουν ότι η συνεισφορά των γενετικών παραγόντων στην εμφάνιση της παχυσαρκίας κυμαίνεται σε ποσοστό από 40 μέχρι και 75%[41]. Η μελέτη αυτή υπέδειξε ότι η διακύμανση του ΔΜΣ που προέρχεται από γενετικούς παράγοντες είχε τη μικρότερη τιμή στην ηλικία των 4 ετών στα αγόρια (0.42) και στα κορίτσια (0.41), ενώ αργότερα είχε αυξητική τάση με μέγιστο το 0.75 και στα δύο φύλα στην ηλικία των 19 ετών. Το αποτέλεσμα αυτό υπογραμμίζει την μεγάλη σημασία της έγκαιρης παρέμβασης για την πρόληψη και καταπολέμηση της υπερβαρότητας[41].

2.1.2 Περιγεννητικοί Παράγοντες

Επίσης σημαντικός είναι ο ρόλος των περιβαλλοντικών παραγόντων που εμπλέκονται στην εμφάνιση της παιδικής παχυσαρκίας. Ποικίλουν σε μέγεθος επιρροής, σε χρονική εμφάνιση και διάρκεια, και ορισμένοι από αυτούς εμφανίζονται ακόμα και στην ενδομήτρια ζωή[38]. Έπειτα από συστηματική ανασκόπηση παραγόντων κινδύνου για την παιδική παχυσαρκία διάρκειας 2 ετών, δηλαδή περίπου στις πρώτες 1000 μέρες ζωής των παιδιών, στην οποία συμπεριελήφθησαν 282 μελέτες, αναδείχθηκαν οι εξής παράγοντες ως σημαντικότεροι για τον κίνδυνο εκδήλωσης της παιδικής παχυσαρκίας: υψηλός ΔΜΣ της μητέρας πριν την εγκυμοσύνη (maternal pre-pregnancy BMI), βάρος γέννησης μεγαλύτερο του φυσιολογικού (infant birth weight), ενδομήτρια έκθεση σε καπνό, υπερβολική αύξηση βάρους στη βρεφική ηλικία (accelerated infant weight gain) και υπερβολική πρόσληψη βάρους της μητέρας κατά την εγκυμοσύνη (excessive gestational weight gain)[42].

Άλλες μελέτες, λιγότερες στον αριθμό, υποστηρίζουν την σχέση κι άλλων παραγόντων με την εμφάνιση παιδικής παχυσαρκίας, για παράδειγμα το χαμηλό κοινωνικό και οικονομικό επίπεδο, η πρώιμη έναρξη στερεών τροφών (λιγότερο των 4 μηνών), η υπερβολική λήψη αντιβιοτικών στη βρεφική ηλικία, ο διαβήτης κύησης, η παρακολούθηση βρεφονηπιακού σταθμού, η διαταραγμένη σχέση μητέρας και βρέφους, ο διαταραγμένος ύπνος και η ακατάλληλη χρήση υποκατάστατων γάλακτος να είναι κάποια από αυτά[42]. Όσων αφορά

την σχέση μεταξύ παιδικής παχυσαρκίας και καπνίσματος της εγκύου, μια σύγχρονη μετα-ανάλυση αποτελούμενη από 39 μελέτες στις οποίες πήραν μέρος περισσότερα από 235.000 παιδιά, ανέδειξε αυξημένο κίνδυνο για υπερβαρότητα και παχυσαρκία σε εκείνα τα παιδιά των οποίων οι μητέρες κάπνιζαν κατά τη διάρκεια της εγκυμοσύνης[43]. Πιο συγκεκριμένα στην Ελλάδα, μια μελέτη αντιπροσωπευτικού δείγματός περίπου 5.125 ζευγαριών παιδιών και μητέρων, που διερεύνησε τη σχέση μεταξύ παραγόντων κινδύνου για παιδική παχυσαρκία κατά τη διάρκεια της εγκυμοσύνης και του ΔΜΣ των παιδιών στην ηλικία των 8 ετών, κατέληξε στο συμπέρασμα ότι τα παιδιά, των οποίων οι μητέρες αύξησαν το βάρος κατά την εγκυμοσύνη περισσότερο από τις συστάσεις του Institute of Medicine (IOM), είχαν μεγαλύτερο κίνδυνο εκδήλωσης παχυσαρκίας στα 8 έτη σε σχέση με τα παιδιά των οποίων οι μητέρες πήραν βάρος σύμφωνα με τις οδηγίες του IOM. Επιπλέον, η μέτρια φυσική δραστηριότητα έναντι της καθιστικής ζωής των εγκύων λειτουργούσε προστατευτικά στην αύξηση του βάρους, εν αντιθέσει με το κάπνισμα στην εγκυμοσύνη, που αύξησε τον κίνδυνο παχυσαρκίας στα 8 έτη[44].

Από τις μελέτες δεν μπορούσε να λείπει η διερεύνηση του μητρικού θηλασμού σχετικά με την παιδική παχυσαρκία. Επιδημιολογικές μελέτες αναδεικνύουν την προστατευτική δράση του θηλασμού έναντι της έξαρσης της παιδικής παχυσαρκίας[38]. Ακόμα μια μετα-ανάλυση 25 επιδημιολογικών μελετών με λίγο πάνω από 226.500 συμμετέχοντες υποστηρίζει ότι ο μητρικός θηλασμός μειώνει τον κίνδυνο παιδικής παχυσαρκίας[45]. Στον αντίποδα, κλινική δοκιμή για την προώθηση του μητρικού θηλασμού με περίπου 17.045 τυχαία ζεύγη μητέρων και παιδιών, εκ των οποίων περισσότερα από 13.550 παιδιά παρακολούθηθηκαν για 16 έτη, έδειξε ότι, παρόλο που η παρέμβαση της έρευνας αύξησε τη διάρκεια και την αποκλειστικότητα του μητρικού θηλασμού, όταν τα παιδιά έφτασαν το 16ο έτος της ηλικίας τους, δεν σχετίστηκε με χαμηλότερα ποσοστά παχυσαρκίας[46]. Στις τελευταίες κατευθυντήριες οδηγίες, για την παιδική παχυσαρκία, το 2017, λαμβάνοντας υπόψη τα αποτελέσματα αυτής της μεγάλης κλινικής μελέτης, η Endocrine Society καταλήγει ότι, συγκεκριμένα για τον μητρικό θηλασμό, οι έρευνες δεν αποδεικνύουν ξεκάθαρη προστατευτική δράση έναντι της παιδικής παχυσαρκίας (inconsistent evidence), αλλά συνεχίζεται να συστήνεται η προώθηση του μητρικού θηλασμού λόγω των πολλαπλών πλεονεκτημάτων που έχει για την υγεία των παιδιών[38].

2.1.3 Τρόπος Ζωής των Παιδιών

Η "Endocrine Society" αποτελεί αν όχι τον μεγαλύτερο, έναν από τους μεγαλύτερους διεθνείς οργανισμούς με στόχο την καταπολέμηση της παχυσαρκίας σε ενήλικες και παιδιά. Εκείνος, κι άλλοι οργανισμοί, τονίζουν την ιδιαίτερη σημασία της υγιεινής διατροφής, της άσκησης, της μείωσης της καθιστικής ζωής και του επαρκούς ύπνου για την πρόληψη αλλά και για την αντιμετώπιση της παιδικής παχυσαρκίας[38, 47]. Πυροδότηση κατευθυντήριων οδηγιών από τους οργανισμούς αποτέλεσε η διεθνής μελέτη "International Study of Childhood Obesity, Lifestyle and the Environment" με 5.844 παιδιά ηλικίας 9 έως 11 ετών από 12 χώρες που υπέδειξε ότι ο μέσος χρόνος που δαπανούν τα παιδιά σε καθιστικές δραστηριότητες ήταν 8,6 ώρες ανά ημέρα και το 54,2% των παιδιών ήταν μπροστά σε οθόνες (τηλεόραση, υπολογιστής, tablet) περισσότερο από 2 ώρες την ημέρα, με το τελευταίο να αποτελεί το όριο

που συστήνεται από τους περισσότερους διεθνείς οργανισμούς. Σε όλες τις χώρες που συμμετείχαν στην έρευνα, τα αγόρια δαπανούσαν περισσότερες ώρες σε καθιστικές δραστηριότητες και είχαν μεγαλύτερη ΔΜΣ ζ-βαθμολογία[48]. Άλλη πρόσφατη συστηματική ανασκόπηση 26 μελετών με περισσότερους από 14.730 συμμετέχοντες, εκ των οποίων περίπου 3.520 ήταν παχύσαρκοι, έκανε σύγκριση της φυσικής δραστηριότητας και του χρόνου που δαπανάται σε καθιστικές δραστηριότητες μεταξύ παχύσαρκων και μη-παχύσαρκων παιδιών και εφήβων ηλικίας από 1 έως 19 ετών. Τα αποτελέσματα έδειξαν ότι τα παχύσαρκα παιδιά και οι έφηβοι σε σχέση με την ομάδα ελέγχου είχαν στατιστικά σημαντικά μειωμένη διάρκεια μέτριας προς υψηλής φυσικής δραστηριότητας (αν και οι διαφορές ήταν σχετικά μικρές), ενώ δεν παρατηρήθηκε στατιστικά σημαντική διαφορά ως προς το χρόνο που δαπανούν σε καθιστικές δραστηριότητες. Η πλειονότητα των παχύσαρκων παιδιών είχαν φυσική δραστηριότητα μέτριας προς υψηλής έντασης μικρότερη από τις συστάσεις, δηλαδή λιγότερο από 60 λεπτά ανά ημέρα, και δαπανούσαν αρκετές ώρες σε καθιστικές δραστηριότητες ανεξάρτητα από το φύλο και την ηλικία τους[49].

Η Endocrine Society καλεί επίσης στην αποφυγή των πλούσιων σε θερμίδες, αλλά χαμηλών σε θρεπτικά συστατικά τροφίμων (calorie-dense, nutrient poor foods) στα παιδιά και τους νέους, προκειμένου να χτίσουν ισχυρές υγιεινές διατροφικές συνήθειες. Στην ζώνη αυτή κατατάσσονται οι χυμοί εμπορίου, τα φαγητά ταχυφαγείου (fast foods), τα αναψυκτικά με απλά σάκχαρα (sugar sweetened beverages - SSB), τα επεξεργασμένα προϊόντα πλούσια με απλά σάκχαρα, το αλάτι, τα πρόσθετα λίπη και τα υπερθερμιδικά σνακ. Παράλληλα, εμμένει στην ενθάρρυνση κατανάλωσης ολόκληρων φρούτων αντί για χυμούς[38]. Πρόσφατη συστηματική ανασκόπηση 32 πειραματικών μελετών για τη σχέση των σακχαρούχων αναψυκτικών (SSB) και του κινδύνου εκδήλωσης παιδικής ή εφηβικής παχυσαρκίας, κατέληξε στο συμπέρασμα ότι η πλειονότητα των μελετών δείχνουν θετική συσχέτιση μεταξύ της κατανάλωσης SSB και κινδύνου παχυσαρκίας ιδίως σε παιδιά που είναι ήδη υπέρβαρα[50]. Άλλη συστηματική ανασκόπηση 26 μελετών που διερεύνησαν τη σχέση κατανάλωσης υψηλής επεξεργασίας τροφίμων (ultra-processed foods) όπως υπερθερμιδικά σνακ, προϊόντα ταχυφαγείου, SSB και γλυκά με το σωματικό λίπος (μετρημένο με διπλά έμμεσες μεθόδους όπως η βιοηλεκτρική εμπέδηση και οι δερματικές πτυχές) σε παιδιά και εφήβους, κατέληξε στο συμπέρασμα ότι οι περισσότερες μελέτες βρήκαν θετική συσχέτιση ανάμεσα στην κατανάλωση αυτών των τροφίμων και την παρουσία αυξημένου σωματικού λίπους[51].

Πολύ σημαντικός παράγοντας είναι ο επαρκής σε διάρκεια και ποιοτικά πλούσιος ύπνος, ιδίως στις ηλικίες που χαρακτηρίζονται από έντονη σωματική αλλά και ψυχολογική ανάπτυξη του ατόμου. Ελληνική σύγχρονη μελέτη που διεξήχθη το 2015 σε πάνω από 177.090 παιδιά ηλικίας 8 έως 17 ετών έδειξε ότι η ανεπαρκής, για την ηλικία, διάρκεια ύπνου, δηλαδή λιγότερες από 9 ώρες για παιδιά ηλικίας 8 έως 11 ετών και λιγότερες από 8 ώρες για τους εφήβους, ήταν αρκετά συχνή. Η ανεπαρκής διάρκεια ύπνου σχετίστηκε επιπλέον με τον αυξημένο χρόνο μπροστά σε οθόνες, την υπερβαρότητα και την παχυσαρκία, καθώς και ανθυγιεινές διατροφικές συνήθειες όπως για παράδειγμα παράληψη πρωινού, συχνή κατανάλωση γλυκών και κατανάλωση φαγητού ταχυφαγείου[52]. Άλλη συστηματική ανασκόπηση 13 συγχρονικών και μίας προοπτικής μελέτης σε συνολικά 36.560 παιδιά ηλικίας 5 έως 17 ετών που διερεύνησε την συνδυαστική επίδραση φυσικής δραστηριότητας, ύπνου και καθιστικής ζωής σε δείκτες υγείας όπως η παχυσαρκία και οι καρδιομεταβολικοί βιοδείκτες,

κατέληξε στο συμπέρασμα ότι ο συνδυασμός μέτριας προς υψηλής φυσικής δραστηριότητας με επαρκή ύπνο και μειωμένη διάρκεια καθιστικών δραστηριοτήτων σχετίστηκε με μειωμένα ποσοστά παχυσαρκίας και ευνοϊκότερους καρδιομεταβολικούς δείκτες σε σχέση με παιδιά που δεν ήταν σωματικά δραστήρια, είχαν ανεπαρκή ύπνο και δαπανούσαν αρκετές ώρες σε καθιστικές δραστηριότητες[53]. Τέλος, μετα-ανάλυση 12 προοπτικών ερευνών που αναζήτησε τη σχέση διάρκειας ύπνου και της παιδικής παχυσαρκίας, κατέληξε στο συμπέρασμα ότι η μικρή διάρκεια ύπνου σχετίζεται με την μετέπειτα εμφάνιση παιδικής παχυσαρκίας[54].

2.2 Ο Ρόλος της Οικογένειας

2.2.1 Γονεϊκός Έλεγχος

Οι γονείς, ως πάροχοι της τροφής, ενθαρρυντές και πρότυπα, σίγουρα έχουν ισχυρή επιρροή στη διαμόρφωση της διατροφικής συμπεριφοράς των παιδιών τους και την κατάσταση του βάρους τους[55]. Ωστόσο, μελέτες έχουν τεκμηριώσει ότι ο γονικός έλεγχος (Parental Control) στο φαγητό μπορεί να έχει ακούσιες συνέπειες σε σχέση με το βάρος των παιδιών. Οι Birch και Davison αναφέρουν ότι ο γονικός περιορισμός της σίτισης συχνά προκαλείται από αντιλήψεις και ανησυχίες σχετικά με τον κίνδυνο παχυσαρκίας του παιδιού[56]. Υπάρχει απόδειξη ότι ο ίδιος ο γονικός έλεγχος, παρόλο που αρχικά αποσκοπεί στην αποτροπή της παχυσαρκίας, προκαλεί σε αρκετές περιπτώσεις δυσμενείς επιπτώσεις στη διατροφική συμπεριφορά και το βάρος του παιδιού[56, 57]. Έτσι για παράδειγμα, σε παιδιά που θεωρούνται υπέρβαρα, οι γονείς τους μπορεί να προσπαθούν να διαχειριστούν το βάρος τους περιορίζοντας τη διατροφική τους πρόσληψη που, με τη σειρά της, μπορεί να κάνει τα πράγματα χειρότερα. Οι Musher-Eizenman και Holub μελέτησαν ένα δείγμα 98 γονέων και ανέφερε ότι, όταν οι γονείς περιόρισαν αυτό που έτρωγαν τα παιδιά τους για λόγους υπερβαρότητας, το παιδί τους έτρωγε περισσότερο όταν του παρέχόταν ελεύθερη πρόσβαση σε ανθυγιεινά γεύματα[58]. Όπως προτάθηκε χαρακτηριστικά από τον Rhee, «φαίνεται ότι οι περιοριστικές συμπεριφορές σίτισης (restrictive) μπορεί να έχουν διαφορετικό αποτέλεσμα για το βάρος του παιδιού ανάλογα με τον κίνδυνο που διατρέχει το παιδί ή της τρέχουσας σωματικής του κατάστασης, με αποτέλεσμα μεγαλύτερη ζημιά σε σχέση από όσους κινδυνεύουν γενετικά ή βιολογικά ή ακόμα και σε σχέση με άτομα που είναι ήδη υπέρβαρα»[59].

Σε μια προοπτική ανάλυση των στάσεων γονικής σίτισης και του ΔΜΣ του παιδιού, διαπιστώθηκε ότι οι συμπεριφορές των γονέων ήταν σταθερές για περίοδο περίπου 2 έτη[60]. Μετά τον έλεγχο σε προηγούμενη κατάσταση βάρους του παιδιού, κατέληξαν στο συμπέρασμα ότι ορισμένοι γονικοί έλεγχοι της σίτισης των παιδιών αυξάνουν την κανονικοποιημένη ΔΜΣ ζ-βαθμολογία (BMI z-score) μεταξύ παιδιών υψηλού κινδύνου, προτείνοντας αμοιβαία επιρροή μεταξύ παιδιού και γονέα[60]. Αυτά τα ευρήματα υπογραμμίζουν την ανάγκη αξιολόγησης του ρόλου των γονεϊκών πρακτικών σίτισης (parents' child-feeding practices) των παιδιών στο πλαίσιο της οικογένειας ώστε να κατανοηθεί καλύτερα η αιτιότητα της παιδικής παχυσαρκίας.

2.2.2 Γονεϊκός Τύπος

Σύμφωνα με τον Baumind, από τους πρωτοπόρους στην έρευνα για τον γονεϊκό τύπο, υπάρχουν τρία κύρια στυλ γονεϊκού τύπου και μέριμνας:

1. η Εξουσιαστική (Authoritative),
2. η Αυταρχική (Authoritarian) και
3. η Επιτρεπτική (Permissive)[61].

Το εξουσιαστικό στυλ ανατροφής αναφέρεται σε έναν γονέα που απαιτεί από το παιδί να «συμμορφωθεί σε ένα εύλογο σύνολο κανόνων και ρυθμίσεων και ανταποκρίνεται ταυτόχρονα στις ανάγκες του παιδιού, ενώ επίσης υποστηρίζει την αυτονομία και την ατομικότητα του παιδιού». Από την άλλη, ένα αυταρχικό στυλ διατροφικής ρύθμισης αναφέρεται σε έναν γονέα που τονίζει τη σημασία της συμμόρφωσης, του γονικού ελέγχου, του σεβασμού στην εξουσία και τη διατήρηση της τάξης[62]. Οι Maccoby και Martin[63] αργότερα πρότειναν μια άλλη τυπολογία του στυλ ανατροφής που συνυπάρχει σε δύο διαστάσεις: αυτήν της «Απαιτητικότητας» (Demandingness) και της «Ανταπόκρισης» (Responsiveness) των γονέων. Εδώ, η «απαιτητικότητα» αναφέρεται στις προσδοκίες για την επίδειξη ωριμότητας από τα παιδιά, τον γονικό έλεγχο και την πειθαρχία, ενώ η «ανταπόκριση» αναφέρεται στη γονική εμφάνιση της ζεστασιάς, της ευαισθησίας, της στοργής και της ουσιαστικής αλληλεπίδρασης με τα παιδιά. Τέσσερις γονεϊκοί τύποι μπορούν να προκύψουν υπό αυτήν την προσέγγιση: η Εξουσιαστική, η Αυταρχική, η Επιτρεπτική και η Παραμελιτική (Neglectful) γονική μέριμνα. Ένας εξουσιαστικός γονέας σκοράρει ψηλά και στην «απαιτητικότητα» και στην «ανταπόκριση». Με την ίδια λογική, ένας αυταρχικός γονέας θα βαθμολογούταν υψηλά στην «απαιτητικότητα» αλλά χαμηλά στην «ανταπόκριση» και ένας επιτρεπτικός γονέας θα ήταν απλώς το αντίθετο από έναν αυταρχικό γονέα στις δύο διαστάσεις που προαναφέραμε. Τα στοιχεία δείχνουν ότι η εξουσιαστική γονική μέριμνα στην διατροφή συνδέεται με ποικίλες θετικές επιδράσεις στην παιδική ανάπτυξη, όπως υψηλότερα ακαδημαϊκά επιτεύγματα, αύξηση της ικανότητας αυτορύθμισης του ατόμου, συχνότερη χρήση προσαρμοστικών στρατηγικών και λιγότερες συμπεριφορές ανάληψης κινδύνου[64, 65].

Οι έρευνες που σχετίζονται με την διατροφή έχουν αρχίσει να μελετούν συστηματικά αυτή τη δομή σε μια προσπάθεια κατανόησης της επιρροής των γονεϊκών τύπων στις διατροφικές συμπεριφορές των παιδιών. Για παράδειγμα, μια μελέτη συμπεραίνει ότι η κατανάλωση φρούτων και λαχανικών ήταν πιο ευνοϊκή μεταξύ εφήβων που ήταν αναθρεμμένοι υπό ένα εξουσιαστικό τύπο γονέα[66]. Αποτελέσματα από άλλες μελέτες που εξετάζουν, ομοίως με παραπάνω, τις επιπτώσεις των γονεϊκών τύπων στην κατανάλωση τροφών πλούσιων σε βιταμίνες και ιχνοστοιχεία, ή στις συμπεριφορές του παιδιού σχετικά με το βάρος, φαίνεται να υποστηρίζουν ότι ένα στυλ γονικής μέριμνας που είναι υψηλό στην διάσταση της απαιτητικότητας και της ανταπόκρισης, συνδέεται γενικότερα με πιο υγιή αποτελέσματα.

Από μια μελέτη που πραγματοποιήθηκε στην Ολλανδία διαπιστώθηκε πώς ο τρόπος με τον οποίο οι έφηβοι αντιλαμβάνονται το γονεϊκό τύπο είχε καθοριστική επίδραση στην αποτελεσματικότητα των γονεϊκών πρακτικών σίτισης, και πιο συγκεκριμένα στην κατανάλωση ροφημάτων με ζάχαρη[67]. Μελέτες που έγιναν από τους Musher-Eizenman και Holub[58]

και επιπλέον από τους Joyce και Zimmer-Gembeck[68] κατέγραψαν επίσης σημαντικές επιδράσεις του γονεϊκού τύπου ανατροφής στη σχέση μεταξύ των περιοριστικών (Restrictive) πρακτικών σίτισης και των διατροφικών συνθηκών των παιδιών και του βάρους τους. Μια άλλη μελέτη διαπίστωσε ότι ένα εξουσιαστικό στυλ ανατροφής προέβλεπε καλύτερο γλυκαιμικό έλεγχο και την τήρηση του θεραπευτικού σχήματος μεταξύ παιδιών ασθενών με διαβήτη τύπου 1[69]. Τέλος, σε μια κριτική πολλών καίριων μελετών σχετικά με τον γονεϊκό τύπο και τα συμπτώματα συνδεδεμένα με το βάρος, αλλά και σχετικά με τις συμπεριφορές μεταξύ εφήβων, Enten και Golan[70] κατέληξαν στο συμπέρασμα ότι οι θετικές συμπεριφορές των εφήβων που σχετίζονται με τις διατροφικές συνήθειες, υποστηρίζονται από την εξουσιαστική πρακτική σίτισης των γονέων. Ωστόσο, η μεγάλη πλειοψηφία των μελετών αυτών σηκώνουν πλήθος αντιρρήσεων λόγω του βασικού τους σχεδιασμού.

Λίγες μελέτες υιοθέτησαν μια πιο "ευρεία" προσέγγιση ώστε να εξετάσουν, όσο πιο αντικειμενικά γίνεται, τη σχέση μεταξύ της γονικής μέριμνας και του βάρους του παιδιού. Για παράδειγμα, ο Rhee et al.[71] εξέτασε τις άμεσες επιδράσεις των γονεϊκών πρακτικών σίτισης στις μεταβολές του βάρους ενός συνόλου μαθητών της πρώτης Δημοτικού, που όμως το επέβλεπαν εξονυχιστικά. Σε αντίστοιχο μήκος κύματος και ο Berge et al.[72] διαπίστωσε ότι ένα εξουσιαστικό στυλ ανατροφής μπορεί να παίζει προστατευτικό και θετικό ρόλο έναντι της υπερβολικής πρόληψης βάρους στους εφήβους. Και οι δύο παραπάνω μελέτες απέδειξαν ότι τα παιδιά με αυταρχικούς γονείς ήταν πιο πιθανό να είναι υπέρβαρα, σε σύγκριση με τα παιδιά με εξουσιαστικούς γονείς. Αυτές οι μελέτες τονίζουν ότι η κατανόηση του τρόπου με τον οποίο οι γονεϊκοί τύποι συνδέονται με τον κίνδυνο πρόληψης βάρους θα μπορούσε να οδηγήσει στην ανάπτυξη ολοκληρωμένων και αποτελεσματικών παρεμβάσεων για την πρόληψη της παχυσαρκίας μεταξύ των μικρών παιδιών και των εφήβων. Μια κριτική από τους Ventura και Birch[73] συνοψίζει τόσο με αιτιολογικά όσο και με συσχετιστικά στοιχεία τις αμοιβαίες επιρροές μεταξύ της γονικής μέριμνας και της διατροφής και του βάρους του παιδιού, υποδεικνύοντας εύλογους τρόπους για να εφαρμοστούν αυτές οι επιρροές.

Άλλη ανασκόπηση[74] επεσήμανε με συγκεκριμένα αποτελέσματά τον επιδραστικό ρόλο του γονεϊκού τύπου στη σχέση μεταξύ των γονεϊκών πρακτικών και των ανθρωπομετρικών αποτελεσμάτων στα παιδιά. Στην πραγματικότητα, στη βιβλιογραφία σχετικά με την κοινωνικοποίηση, οι Darling και Steinberg[75] προτείνουν εδώ και καιρό ότι η γονική μέριμνα πρέπει να εκληφθεί άμεσα ως πλαίσιο που επηρεάζει την επιρροή συγκεκριμένων πρακτικών σίτισης στο παιδί. Υποστηρίζουν, ότι η γονική μέριμνα επηρεάζει την γενικότερη ανάπτυξη του παιδιού μέσω τριών χαρακτηριστικών των γονέων:

1. Τις αξίες και τους στόχους που έχουν οι γονείς στην κοινωνικοποίησή των παιδιών
2. Τις πρακτικές ανατροφής που χρησιμοποιούν οι γονείς
3. Τις συμπεριφορές που εξωτερικεύουν απέναντι στα παιδιά τους

Με αυτόν τον τρόπο διακρίνουν περαιτέρω την έννοια των γονεϊκών πρακτικών από τους γονεϊκούς τύπους και υποστηρίζουν ότι η γονεϊκή πρακτική είναι «η συμπεριφορά που καθορίζεται από συγκεκριμένο περιεχόμενο και συγκεκριμένους στόχους κοινωνικοποίησης» ενώ ο γονεϊκός τύπος, δηλαδή η γονική μέριμνα, αναφέρονται κατεξοχήν στις αλληλεπιδράσεις

γονέων-παιδιών σε διάφορες καταστάσεις της καθημερινότητας[75]. Με βάση αυτό το εννοιολογικό μοντέλο του γονεϊκού τύπου, εάν το αποτέλεσμα που ενδιαφέρει είναι η κατάσταση του βάρους του παιδιού, το γονεϊκό στυλ θα επηρεάσει άμεσα αυτήν την πτυχή, κυρίως λόγω των επιδραστικών επιπτώσεών του στη σχέση μεταξύ των γονεϊκών πρακτικών σίτισης των παιδιών και του βάρους του. Ο Rhee[59] προτείνει επίσης ότι οι γονικού τύποι μπορούν να επηρεάσουν τη σχέση μεταξύ των συμπεριφορών διατροφής των παιδιών και των τιμών του βάρους τους έτσι ώστε μια «συμπεριφορά που παρέχεται εντός του πλαισίου ενός πιο θετικού γονεϊκού τύπου, θα έχει σημαντικά διαφορετικό αντίκτυπο στο παιδί από αυτό που του παρέχεται περισσότερο έναν αρνητικό γονεϊκό τύπο».

2.2.3 Γονεϊκές Πρακτικές Σίτισης

Οι γονεϊκές πρακτικές σίτισης και η σχέση τους με τις διατροφικές συνήθειες των παιδιών θα μας απασχολήσει ιδιαίτερα στην παρούσα εργασία. Αφορμή, η δημοσίευση συστηματικής ανασκόπησης με μετα-ανάλυση στην οποία συμπεριλήφθηκαν 78 μελέτες, εκ των οποίων οι 37 κρίθηκαν κατάλληλες για τη μετα-ανάλυση[75]. Οι πρακτικές σίτισης που μελετήθηκαν ήταν οι παρακάτω:

1. ενεργή καθοδήγηση/ενθάρρυνση (active guidance/education),
2. περιοριστική καθοδήγηση/χρήση κανόνων (restrictive guidance/rule-making),
3. διαθεσιμότητα (availability),
4. προσβασιμότητα (accessibility),
5. γονείς-πρότυπα (modeling),
6. πίεση για φαγητό (pressure to eat),
7. χρήση του φαγητού ως επιβράβευση (food as reward),
8. επιβράβευση με λεκτικό τρόπο (rewarding with verbal praise) και
9. η επιβράβευση με διατροφική επίτρεψη (rewarding food consumption materially).

Οι διατροφικές συνήθειες που επικεντρώθηκαν οι μελέτες ήταν η κατανάλωση φρούτων και λαχανικών (κατηγορία: υγιεινές συνήθειες) και η κατανάλωση σακχαρούχων αναψυκτικών και σνακ (κατηγορία: μη υγιεινές συνήθειες). Τα αποτελέσματα έδειξαν ότι η διαθεσιμότητα και οι γονείς-πρότυπα είχαν τις ισχυρότερες συσχετίσεις τόσο για τις υγιεινές όσο και για τις μη υγιεινές συνήθειες. Οι ερευνητές τονίζουν ότι η αποτελεσματικότητα μερικών πρακτικών σίτισης ίσως εξαρτάται από το είδος του τροφίμου (υγιεινό/ανθυγιεινό) και την ηλικία του παιδιού. Πολύ ενδιαφέρον έχει το γεγονός ότι, όσων αφορά τα υγιεινά τρόφιμα, η ενεργή καθοδήγηση και ενθάρρυνση φαίνεται να είναι πιο αποτελεσματική σε αντίθεση με τα ανθυγιεινά τρόφιμα, όπου η περιοριστική καθοδήγηση και χρήση κανόνων μοιάζει να είναι πιο αποτελεσματική. Αντίστοιχα, για παιδιά ηλικίας επτά ετών και άνω, η περιοριστική καθοδήγηση και χρήση κανόνων δείχνει να είναι πιο αποτελεσματική για την πρόληψη ανθυγιεινής διατροφής, ενώ για παιδιά ηλικίας έξι ετών και κάτω, η επιβράβευση με λεκτικό

τρόπο ίσως είναι πιο αποτελεσματική στην προώθηση της υγιεινής διατροφής, αλλά και στην πρόληψη της ανθυγιεινής διατροφής[76].

2.3 Ερευνητικές Προκλήσεις

Κατά συνέπεια, όταν οι έρευνες αξιολόγησης των συσχετίσεων και των στρατηγικών πρόληψης της υπερβαρότητας αναζητούν πληροφορίες πέρα από τα επιφανειακά παραδοσιακά ποσοστά επικράτησης, οι προσπάθειές τους οδηγούν συχνά σε αμφιλεγόμενα και διφορούμενα αποτελέσματα. Παραμένει, επομένως, ασαφές πώς αυτές οι ψυχολογικές συσχετίσεις μπορούν να οδηγήσουν στην επιτυχία των στρατηγικών παρέμβασης και την αποτελεσματική αντιμετώπιση της νόσου[77].

Η εργασία μας καλείται να αντιμετωπίσει ακριβώς αυτή την αβεβαιότητα. Υποστηρίζουμε ότι υπάρχει ανάγκη να προχωρήσουμε πέρα από τις συσχετίσεις σε πολυμορφικά χαρακτηριστικά και να καθιερώσουμε τη σημασία των αλληλεπιδράσεων μεταξύ αυτών των χαρακτηριστικών αφού αναφέρονται στην αποτελεσματικότητα της εκάστοτε προσπάθειας νοσηλείας.

Κεφάλαιο **3**

Μοντελοποίηση της Παιδικής Παχυσαρκίας

Στο κεφάλαιο αυτό καταγράφουμε τις πιο ενδιαφέρουσες μελέτες που έχουν προηγηθεί, σχετικές με την μοντελοποίηση της παιδικής παχυσαρκίας, τις τεχνικές και μεθόδους που εφάρμοσαν για να προσεγγίσουν το συγκεκριμένο πρόβλημα, και τις προσεγγίσεις που φαίνονται πιο πειστικές και με τις οποίες μπορούμε να οραματιστούμε και ύστερα να στηρίξουμε την δική μας μελέτη και υλοποίηση.

3.1 Προηγηθείσες Παρατηρήσεις

Η μελέτη των Hughes και Power[78] είναι η πρώτη που διερεύνησε τον αντίκτυπο των γονεϊκών πρακτικών σίτισης, και συγκεκριμένα των μητέρων, σχετικά με το βάρος των παιδιών, στην πάροδο του χρόνου. Το πέτυχαν με ένα δείγμα ισπανόφωνων παιδιών που προέρχονται από οικογένειες χαμηλού εισοδήματος μέσα από το πρόγραμμα Head Start. Προς υποστήριξη της υπόθεσης, η "Επιεικής" πρακτική σίτισης αξιολογήθηκε όταν το παιδί ήταν κατά μέσο όρο 4,8 ετών και συσχετίστηκε με αυξημένο δείκτη ΔΜΣ 18 μήνες αργότερα. Εξουσιαστικές, αυταρχικές και αμέτοχες πρακτικές σίτισης δεν συσχετίστηκαν με αλλαγές στη ΔΜΣ ζ-βαθμολογία στο πέρασμα του χρόνου. Η σχέση ενός επιεικούς τρόπου διατροφής στις ΔΜΣ ζ-βαθμολογίες των παιδιών συνέβαλαν σημαντικά στην ερμηνεία των αυξήσεων του ΔΜΣ του παιδιού πολύ πάνω από το βασικό όριο, τα δημογραφικά στοιχεία, και τις διατροφικές συμπεριφορές του παιδιού στην ηλικία των 4 ετών. Άλλοι σημαντικοί προγνωστικοί δείκτες του χρόνου για την ΔΜΣ ζ-βαθμολογία στο τελικό μοντέλο ήταν η ΔΜΣ ζ-βαθμολογία του παιδιού την πρώτη χρονική περίοδο σύγκρισης (ισχυρότερος προγνωστικός δείκτης), το πολιτισμικό επίπεδο του γονέα και η χρήση των περιοριστικών πρακτικών και πρακτικών παρακολούθησης από τους γονείς. Ενώ οι δύο πρώτοι ήταν παράγοντες κινδύνου για μεγαλύτερη ΔΜΣ ζ-βαθμολογία στη δεύτερη χρονική περίοδο (θετικές συσχετίσεις), η παρακολούθηση κατά τον ένα χρόνο ήταν προστατευτική ή σχετίζεται αρνητικά με τον δείκτη ΔΜΣ του παιδιού (αρνητική συσχέτιση).

Αξιοσημείωτο είναι το γεγονός ότι το πιο επιεικές στυλ σίτισης ήταν αυτό που συσχετίστηκε με την ΔΜΣ ζ-βαθμολογία του παιδιού με την πάροδο του χρόνου, καθώς ο επιεικής τρόπος διατροφής συσχετίζεται σταθερά με υψηλότερη ΔΜΣ ζ-βαθμολογία του παιδιού και περαιτέρω ποσοστά σε άλλες διασταυρωμένες μελέτες που αξιολογούν την ίδια δομή μεταξύ δειγμάτων Αφροαμερικανών και Ισπανόφωνων [79, 80, 81], Αφροαμερικανών που ζουν σε αγροτικές περιοχές, Ισπανών, λοιπών Ευρωπαίων[82] και πρόσφατων μεταναστών, όλοι

με χαμηλά εισοδήματα[83]. Η επιεικής πρακτική διατροφής συμβαίνει όταν οι γονείς έχουν λίγες απαιτήσεις από τα παιδιά τους κατά τη διάρκεια της σίτισης και χρησιμοποιούν κυρίως οδηγίες διατροφής με επίκεντρο τα παιδιά. Αυτές οι σχετικά επιτρεπτικές πρακτικές δεν παρέχουν στα παιδιά το ίδιο επίπεδο κλιμάκωσης, ρύθμισης ορίων και κανόνων μεταξύ των γονέων που χρησιμοποιούν εξουσιαστικές ή αυταρχικές πρακτικές. Προηγούμενη μελέτη σε οικογένειες με χαμηλό εισόδημα παρατήρησε ότι επιεικές και αμέτοχες διατροφικές πρακτικές σχετίζονται με χαμηλότερη πρόσληψη φρούτων, λαχανικών και γαλακτοκομικών[84], παρέχοντας έναν μόνο πιθανό μηχανισμό μέσω του οποίου τα διάφορα στυλ διατροφής επηρεάζουν την κατάσταση του βάρους του παιδιού με την πάροδο του χρόνου. Ωστόσο, είναι εμφανές στην ανάλυση που παρουσιάζεται ότι το επιεικές στυλ σίτισης τοποθετεί τα παιδιά σε μεγαλύτερο κίνδυνο υπερβολικής αύξησης βάρους με την πάροδο του χρόνου. Αυτό υποδηλώνει ότι εμπλέκονται επιπλέον μηχανισμοί. Είναι εύλογη η υπόθεση ότι οι γονείς που είναι επιεικές μπορεί να χρησιμοποιούν το φαγητό για να δείξουν αγάπη και στοργή για το παιδί τους, το οποίο μπορεί να συμβάλει σε μεγαλύτερη πρόσληψη ενέργειας και να αποτελέσει έναν επιπλέον μηχανισμό με τον οποίο συμβάλλουν οι επιεικές τρόποι διατροφής στην υπερβολική αύξηση βάρους των παιδιών[81]. Επιπλέον, η επιεικής πρακτική διατροφής μπορεί να επηρεάσει την αυτορύθμιση των παιδιών σε θερμιδικές προσλήψεις εξαιτίας των χαμηλών επιπέδων γονικής ρύθμισης ορίων και η χρήση κανόνων μπορεί να οδηγήσει σε υπερκατανάλωση τροφής και αδιαφορία των παιδιών στα σημάδια του εσωτερικού κορεσμού τους[85]. Μία δύναμη στην δομή του στυλ σίτισης είναι ότι αξιολογεί τις επιρροές των γονέων στην παιδική διατροφή στο πλαίσιο τόσο του παιδιού όσο και του γονέα, δίνοντας μια πιο σφαιρική αξιολόγηση του τρόπου με τον οποίο οι γονείς αλληλεπιδρούν με το παιδί τους κατά τη διάρκεια των γευμάτων σε σύγκριση με το πρακτικές σίτισης των παιδιών.

Η θετική συσχέτιση των περιοριστικών πρακτικών γονικής μέριμνας και η αρνητική συσχέτιση της παρακολούθησης για τα τρόφιμα συνέβαλαν μόνο στην ΔΜΣ ζ-βαθμολογία του παιδιού στην δεύτερη περίοδο της μελέτης, όταν ελέγχθηκε επίσης το στυλ διατροφής των γονέων. Είναι πιθανό ότι οι πρακτικές σίτισης των γονέων εκφράζονται διαφορετικά μεταξύ διαφορετικών μορφών διατροφής, και επομένως και τα δύο πρέπει να λαμβάνονται υπόψη κατά την αξιολόγηση της αύξησης του βάρους των παιδιών με την πάροδο του χρόνου. Στην πραγματικότητα, ο περιορισμός και η παρακολούθηση είχε πιο μικτά αποτελέσματα στις μελέτες αξιολόγησης και στη συσχέτισή τους με το βάρος ή το ΔΜΣ των παιδιών, σε σύγκριση με τα υπόλοιπα γνωστά στυλ διατροφής[86]. Οι περισσότερες από τις διασταυρωμένες μελέτες έχουν υποστηρίξει μια θετική συσχέτιση με τον περιορισμό και τον ΔΜΣ των παιδιών, αλλά τα διαχρονικά ευρήματα ήταν αμφίβολα[86, 87]. Η παρακολούθηση, από την άλλη πλευρά, γενικά δεν έχει συνδεθεί με τον ΔΜΣ των παιδιών σε διασταυρωμένες ή διαχρονικές μελέτες[86] αλλά συνδέθηκε με καλύτερη ποιότητα διατροφής δύο χρόνια αργότερα[88]. Μια μικρή μελέτη συσχέτισε την παρακολούθηση στην ηλικία των 5 ετών με μειωμένη ΔΜΣ ζ-βαθμολογία των ίδιων παιδιών στην ηλικία των 7, μεταξύ ατόμων σε κίνδυνο παχυσαρκίας, αλλά όχι μεταξύ παιδιών υψηλού κινδύνου παχυσαρκίας, όπως αυτό καθορίζεται από την κατάσταση βάρους του γονέα[60]. Επομένως, οι ερευνητές και οι πάροχοι υγειονομικής περίθαλψης πρέπει να λαμβάνουν υπόψη και τις πρακτικές γονικής μέριμνας και τις πρακτικές διατροφής των παιδιών κατά την παρέμβαση και τη θεραπεία της παιδικής παχυσαρκίας.

Έχει υποστηριχθεί ότι το μεγάλο μέρος της συσχέτισης των τρόπων διατροφής και πρα-

κτικών σίτισης με το βάρος των παιδιών οφείλεται στην ανταπόκριση των γονεϊκών συμπεριφορών στις διατροφικές συμπεριφορές του παιδιού τους, οι οποίες μπορεί πραγματικά να οδηγήσουν στην συσχέτιση[89]. Στην ανάλυση που παρουσιάζουν οι Hughes και Power[78], αξιολογήθηκαν τρία χαρακτηριστικά διατροφικής συμπεριφοράς του παιδιού (ανταπόκριση στα τρόφιμα, συναισθηματική υπερφαγία και ανταπόκριση στον κορεσμό) που προηγουμένως φαίνεται να συνδέονται είτε με τις γονικές πρακτικές διατροφής είτε με την κατάσταση βάρους του παιδιού σε συγχρονικές μελέτες[90, 91]. Ενώ η ανταπόκριση στον κορεσμό ήταν σημαντικά αρνητικά συσχετισμένη με τη ΔΜΣ ζ-βαθμολογία του παιδιού στην πρώτη περίοδο, καμία από τα αυτές τις διατροφικές συμπεριφορές των παιδιών δεν συσχετίστηκαν με την ΔΜΣ ζ-βαθμολογία την δεύτερη χρονική περίοδο. Ούτε συνέβαλε κάποια συμπεριφορά σε μεταβολές της ΔΜΣ ζ-βαθμολογία του παιδιού με την πάροδο του χρόνου κατά τον έλεγχο των δημογραφικών στοιχείων και του πολιτιστικού επιπέδου ή όταν εξετάστηκαν οι πρακτικές σίτισης.

Ο πολιτισμός, ή το πόσο προσαρμόζεται ένας μετανάστης στη νέα κουλτούρα στην οποία κατοικεί φαίνεται να είναι θετικά συσχετισμένα με τον κίνδυνο παιδικής παχυσαρκίας[92, 93, 94, 95]. Έχει αποδειχθεί προηγουμένως ότι το επίπεδο προσαρμογής των ισπανόφωνων γονέων κατά την έναρξη της μελέτης συσχετίστηκε με χαμηλότερη χρήση περιοριστικών πρακτικών, ενώ οι γονείς που γεννήθηκαν στις ΗΠΑ ήταν πιο πιθανό να αναφέρουν έναν πιο επιεική τρόπο διατροφής και λιγότερο πιθανό να αναφέρουν έναν αυταρχικό τρόπο σίτισης[96]. Έτσι προκύπτει ότι το επίπεδο πολιτισμού του γονέα στην αγγλική/αμερικανική κουλτούρα όπως αξιολογείται από την χρήση και το επίπεδο της αγγλικής γλώσσας και την χρήση αγγλικών μέσων ενημέρωσης ήταν προγνωστικό για αυξημένη ΔΜΣ ζ-βαθμολογία του παιδιού με την πάροδο του χρόνου. Αυτά τα αποτελέσματα συμφωνούν με άλλες μελέτες και καταλήγουν στο ότι πιο επιμορφωμένοι Ισπανόφωνοι τείνουν να έχουν υψηλότερα επίπεδα παχυσαρκίας από τους πρόσφατους μετανάστες[92, 93, 97, 98].

Κεφάλαιο 4

Μπεϊσιανή Ανάλυση

Ο Συμπερασμός στα Μπεϊσιανά Δίκτυα (Bayesian Networks Inference) χρησιμοποιείται για τον υπολογισμό της εκ των υστέρων κατανομής πιθανότητας (Posterior Probability Distribution) μη παρατηρούμενων μεταβλητών σε ένα δίκτυο. Αυτές οι εκ των υστέρων κατανομές πιθανότητες χρησιμοποιούνται ευρέως για την εξαγωγή συμπερασμάτων και αποτελούν τη βάση για αποφάσεις, στον τομέα εφαρμογής ενός συγκεκριμένου μοντέλου. Ο συμπερασμός είναι μια πολύπλοκη διαδικασία και μπορεί να είναι δύσκολο να κατανοηθεί ακόμη και από τους πιο έμπειρους χρήστες μπεϊσιανών δικτύων[99]. Σε αυτή τη διατριβή, εκμεταλλευόμενοι το ακαδημαϊκό πρόγραμμα GeNIe Academic, χρησιμοποιούμε τεχνικές για να απεικονίσουμε τις σημαντικότερες πτυχές ενός μπεϊσιανού δικτύου και να καταστήσουμε τη διαδικασία συμπερασμάτων φιλική στον αναγνώστη. Η κυριότερη τεχνική συνίσταται στην βελτίωση της οπτικής αναπαράστασης του μπεϊσιανού δικτύου με επιπλέον πληροφορίες. Η λειτουργία των ακμών σε ένα μπεϊσιανό δίκτυο είναι να υποδείξει τις συσχετίσεις μεταξύ των μεταβλητών. Για αυτό το λόγο, παρουσιάζουμε τις ακμές στο δίκτυο ώστε να εμφανίζουν πρόσθετες πληροφορίες: (i) το πάχος ενός τόξου προσαρμόζεται αυτόματα ώστε να αντιπροσωπεύει το μέτρο της δύναμης της επιρροής (Strength of Influence) μεταξύ δύο άμεσα συνδεδεμένων κόμβων και (ii) το χρώμα ενός τόξου ρυθμίζεται για να υποδείξει την κατεύθυνση επιρροής μεταξύ δύο άμεσα συνδεδεμένων κόμβων. Πέρα από αυτήν την οπτική απεικόνιση, καταγράφουμε και έμμεσες συσχετίσεις μεταξύ μη άμεσα συνδεδεμένων κόμβων.

4.1 Γνωσιακό Υπόβαθρο

4.1.1 Νόμος του Μπέις

Στη θεωρία πιθανοτήτων, όσον αφορά τις δεσμευμένες πιθανότητες (Conditional Probabilities), ο νόμος του Μπέις[100] (Bayes Rule) παίζει εξέχων ρόλο. Ο κανόνας μπορεί μαθηματικά να διατυπωθεί ως εξής:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \quad (4.1)$$

Νοηματικά αυτό σημαίνει ότι η εκ των υστέρων πιθανότητα ενός ενδεχομένου Y μετά την παρατήρηση κάποιου στοιχείου X ισούται με την πιθανότητα παρατήρησης του X δεδομένου

του Y , πολλαπλασιασμένο με την προγενέστερη πιθανότητα του Y , διαιρούμενη με την προγενέστερη πιθανότητα του X . Τώρα, αν αναθέσουμε μια ασθένεια στο Y και ένα σύμπτωμα αυτής της νόσου στο X , μπορούμε να αντιληφθούμε τη σημασία αυτού του θεωρήματος. Συχνά είναι πολύ πιο εύκολο να προσδιοριστεί η πιθανότητα του συμπτώματος X δεδομένης της νόσου Y , και ύστερα να προσδιοριστεί η πιθανότητα της ασθένειας Y δεδομένου του συμπτώματος X , παρόλο που η δεύτερη πιθανότητα είναι πολύ ενδιαφέρουσα να γνωρίζουμε. Ο κανόνας του Μπέις καθιστά αυτή τη γνώση δυνατή.

4.1.2 Μπεϊσιανά Δίκτυα

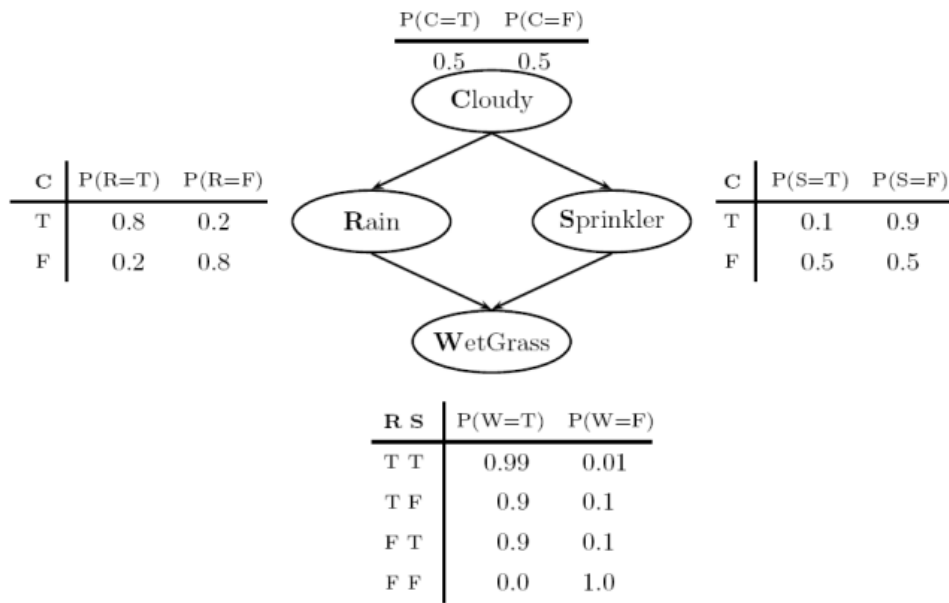
Το Μπεϊσιανό δίκτυο (Bayesian Network) είναι ένα πιθανοτικό γραφικό δίκτυο. Αντιπροσωπεύει μεταβλητές ενός συγκεκριμένου θεματικού τομέα και οπτικοποιεί τις πιθανοτικές σχέσεις μεταξύ τους. Το πιο σημαντικό χαρακτηριστικό του είναι ότι αυτές οι σχέσεις μπορούν να θεωρηθούν ως αιτιώδεις σχέσεις. Ο επίσημος μαθηματικός ορισμός του Μπεϊσιανού δικτύου έχει ως εξής[101]:

1. Ένα σύνολο τυχαίων μεταβλητών αποτελεί τους κόμβους του δικτύου. Οι μεταβλητές μπορεί να είναι διακριτικές ή συνεχές, αλλά στην πράξη προτιμούνται οι διακριτές.
2. Ένα σύνολο κατευθυνόμενων ακμών συνδέει ζεύγη κόμβων. Αν υπάρχει μία κατευθυνόμενη ακμή από τον κόμβο X στον κόμβο Y , τότε το X λέγεται γονέας (Parent) του Y .
3. Κάθε κόμβος X_i έχει κατανομή δεσμευμένης πιθανότητας $P(X_i|Parents(X_i))$ που ποσοτικοποιεί την επίδραση των γονέων στον κόμβο.
4. Το γράφημα δεν έχει κατευθυνόμενους κύκλους και ως εκ τούτου αποτελεί ένα κατευθυνόμενο ακυκλικό γράφημα, ΚΑΓ (Directed Acyclic Graph, DAG).

Ένα Μπεϊσιανό δίκτυο καθορίζει μια πλήρη από κοινού κατανομή πιθανότητας (complete joint probability distribution) πάνω στα X από τον τύπο:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i|Parents(X_i)) \quad (4.2)$$

Για να επεξηγήσουμε περαιτέρω τις παραπάνω έννοιες θα εισαγάγουμε ένα παράδειγμα Μπεϊσιανού Δικτύου. Στην Εικόνα 3.1[101] εμφανίζεται ένα μπεϊσιανό δίκτυο με τέσσερις κόμβους και έναν πίνακα δεσμευμένων πιθανοτήτων για κάθε έναν από τους κόμβους αυτούς. Μοντελοποιείται η ακόλουθη κατάσταση: Ο καιρός είτε είναι νεφελώδης (Cloudy) είτε όχι, επηρεάζει την πιθανότητα να βρέξει (Rain) και την πιθανότητα να είναι ο ψεκαστήρας (Sprinkler) ανοιχτός. Εάν είναι νεφελώδης, ο ψεκαστήρας κατά πάσα πιθανότητα δεν θα είναι ενεργοποιημένος. Η υγρασία στο γρασίδι (WetGrass) επηρεάζεται τόσο από τη βροχή όσο και από τον ψεκαστήρα. Αν υπάρχει βροχή και ο ψεκαστήρας είναι ενεργοποιημένος, η πιθανότητα να είναι υγρό το γρασίδι είναι η υψηλότερη, δηλαδή 0,99. Εάν δεν υπάρχει βροχή και ο ψεκαστήρας είναι απενεργοποιημένος, είναι σίγουρο ότι το γρασίδι δεν είναι υγρό και η πιθανότητα είναι 1. Μια κατευθυνόμενη ακμή μεταξύ δύο κόμβων δείχνει ότι οι δύο αυτοί κόμβοι είναι εξαρτημένοι, πράγμα που σημαίνει ότι επηρεάζουν ο ένας τον άλλον.

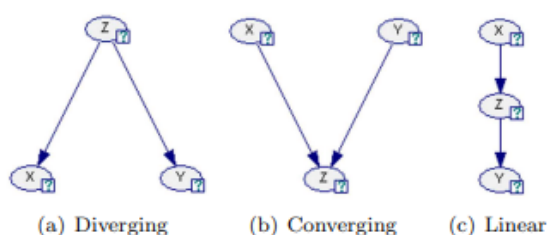


Εικόνα 4.1: Παράδειγμα Μπεϊσιανού Δικτύου

Εάν δεν υπάρχει ακμή μεταξύ δύο κόμβων, τότε δεν έχουν επιρροή ο ένας στον άλλο, ή τουλάχιστον όχι άμεσα. Επίσης, εάν βλέπουμε ότι, για παράδειγμα, το γρασίδι είναι υγρό, τότε έχουμε παρατηρήσει τη μεταβλητή (ή κόμβο) WetGrass, και άρα η μεταβλητή έχει γίνει στοιχείο, παρατήρηση ή εύρημα. Αυτοί οι τρεις όροι μπορούν να χρησιμοποιηθούν εναλλακτικά. Τελικά, αυτό το δίκτυο μπορεί να χρησιμοποιηθεί για να συμπεράνει πιθανότητες όπως αυτή ότι ο ουρανός είναι νεφελώδης όταν γνωρίζουμε ότι το γρασίδι είναι υγρό αλλά ο ψεκαστήρας είναι απενεργοποιημένος, ή την πιθανότητα του ενεργοποιημένου ψεκαστήρα όταν γνωρίζουμε ότι το γρασίδι είναι υγρό και δεν έβρεξε. Θα εξηγήσουμε παρακάτω πώς εκμαιεύουμε αυτήν την πληροφορία. Κάθε ερώτημα σχετικά με τον τομέα, συμπεριλαμβανομένων αυτών που μόλις τέθηκαν, καθορίζονται από την πλήρη από κοινού κατανομή πιθανότητας $P(\text{Cloudy}, \text{Rain}, \text{Sprinkler}, \text{WetGrass})$.

Η πλήρης από κοινού κατανομή πιθανότητας αποτελείται, σε αυτή την περίπτωση, από $2^4 = 16$ καταχωρήσεις, την πιθανότητα κάθε δυνατού συνδυασμού μεταβλητών. Το δίκτυο του σχήματος 3.1 αντιπροσωπεύει την ίδια ακριβώς κατανομή, αλλά έχει μόνο εννέα καθορισμένες πιθανότητες στους πίνακες δεσμευμένων πιθανοτήτων. Υπάρχουν συνολικά δεκαοκτώ αριθμοί στην εικόνα, αλλά όλες οι μεταβλητές είναι δυαδικές επομένως η πιθανότητα μιας κατάστασης είναι 1 μείον την πιθανότητα της άλλης κατάστασης και ως εκ τούτου χρειάζονται μόνο οι εννέα αριθμοί. Αυτή η μείωση των απαραίτητων στοιχείων είναι πολύ σημαντικό πλεονέκτημα των μπεϊσιανών δικτύων και προκαλείται από τις δεσμευμένες ανεξάρτητες παραδοχές που δημιουργούνται από το ίδιο το δίκτυο. Όσο μεγαλύτερο είναι το δίκτυο ή ο τομέας εφαρμογής, τόσο μεγαλύτερη είναι η εξοικονόμηση στοιχείων.

Έστω τώρα μεταβλητές X , Y και Z . Εάν οι X και Y είναι ανεξάρτητες, τότε ισχύει η ακόλουθη μαθηματική έκφραση: $P(X, Y) = P(X)P(Y)$. Τώρα αν το X είναι δεσμευμένο ανεξάρτητο από το Y δεδομένου Z , μπορούμε να γράψουμε: $P(X, Y|Z) = P(X|Z)P(Y|Z)$. Αυτό είναι γνωστό ως Αποκλίνουσα Σύνδεση (Diverging Connection), όπως φαίνεται στο



Σχήμα 4.1: Τύποι Σύνδεσης σε ΚΑΓ

σχήμα 3.2 (α). Μπορούμε τώρα να αποσυνθέσουμε την από κοινού κατανομή πιθανότητας: $P(X, Y, Z) = P(X, Y|Z)P(Z) = P(X|Z)P(Y|Z)P(Z)$. Ένας συγκλίνων κόμβος (Converging), όπως φαίνεται στο σχήμα 3.2 (β), εκφράζει το γεγονός ότι οι X και Y είναι οριακά ανεξάρτητες (Marginal Independent), αλλά δεσμευμένα εξαρτημένες δεδομένου του Z . Αυτό μας επιτρέπει να παραγοντοποιήσουμε ξανά την από κοινού κατανομή πιθανότητας: $P(X, Y, Z) = P(Z|X, Y)P(X)P(Y)$. Ένας τρίτος και τελευταίος τύπος σύνδεσης είναι ο γραμμικός, βλ. Σχήμα 2.2 (γ). Αυτός μοντελοποιεί την κατάσταση ότι, για παράδειγμα, το X είναι το γεγονός ότι κάποιος/κάποια άφησε τους προβολείς του αυτοκινήτου του/της ανοιχτούς για μια νύχτα. Αυτό είναι γνωστό ότι προκαλεί το άδειασμα της μπαταρίας το ακόλουθο πρωί, ενδεχόμενο Z , το οποίο με τη σειρά του προκαλεί το αυτοκινήτου να μην ξεκινά, ενδεχόμενο Y . Τώρα, εάν μέσω κάποιας διαδικασίας διαγνωστεί ότι η μπαταρία είναι άδεια, τότε το να μάθουμε ότι το αυτοκίνητο δεν ξεκινά δεν θα επηρεάσει περαιτέρω την πιθανότητα ότι όλα αυτά προκαλούνται με το να αφήνουμε τους προβολείς αναμμένους όλη τη νύχτα. Η από κοινού κατανομή πιθανότητας παραγοντοποιείται τότε ως εξής: $P(X, Y, Z) = P(Y|Z)P(Z|X)P(X)$. Αυτό μας επιτρέπει να εισαγάγουμε την έννοια του δ-διαχωρισμού (d-seperation) σε ένα μπεϊσιανό δίκτυο. Λέμε ότι δύο κόμβοι είναι δ-διαχωρισμένοι με συγκεκριμένα στοιχεία ή παρατηρήσεις, εάν είναι ανεξάρτητοι με βάση αυτές τις συγκεκριμένες παρατηρήσεις. Με άλλα λόγια, αν δύο οι κόμβοι είναι δ-διαχωρισμένοι, δεν επηρεάζουν ο ένας τον άλλον. Αν συμβαίνει αυτό, δεν υπάρχει ενεργή διαδρομή μεταξύ των δύο κόμβων. Μια διαδρομή είναι ενεργή εάν, κοιτάζοντας το σχήμα 3.2:

1. Η σύνδεση αποκλίνει και το Z δεν έχει παρατηρηθεί.
2. Η σύνδεση συγκλίνει και το Z , ή ένας απόγονός του, έχει παρατηρηθεί.
3. Η σύνδεση είναι γραμμική και το Z δεν έχει παρατηρηθεί.

Έτσι, αν επιστρέψουμε στο παράδειγμά μας, γίνεται εμφανές ότι η αρχική από κοινού κατανομή πιθανότητας $P(Cloudy, Rain, Sprinkler, WetGrass)$ μπορεί να αποσυντεθεί σε:

$$P(Cloudy)P(Rain|Cloudy)P(Sprinkler|Cloydy)P(WetGrass|Rain, Sprinkler)$$

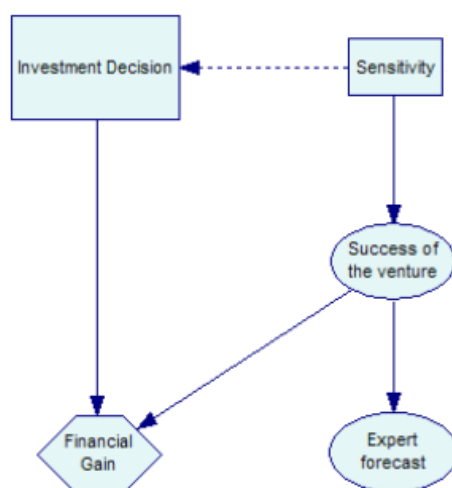
όπου όλες οι παραπάνω ζητούμενες πιθανότητες δίνονται στους πίνακες δεσμευμένων πιθανοτήτων του σχήματος 3.1. Χρησιμοποιώντας αυτήν την νοητική διαδικασία μπορεί κανείς να συναγάγει οποιαδήποτε πιθανότητα στον τομέα εφαρμογής που τον ενδιαφέρει. Αυτή ακριβώς η διαδικασία ονομάζεται Συμπερασμός (Inference).

4.1.3 Μπεϊσιανός Συμπερασμός

Ο βασικός και κύριος σκοπός ενός μπεϊσιανού δικτύου είναι να υπολογίσει τις εκ των υστέρων κατανομές πιθανότητας για ένα σύνολο ερωτηθέντων μεταβλητών, δεδομένης της παρατήρησης ενός συνόλου παρατηρούμενων μεταβλητών. Αυτή η διαδικασία είναι γνωστή ως Μπεϊσιανός Συμπερασμός (Bayesian Inference)[102], και στην βιβλιογραφία ονομάζεται επίσης Μπεϊσιανή Ενημέρωση (Bayesian Updating), Ενημέρωση Πειποιθήσεων (Belief Updating) ή Συλλογισμός (Reasoning). Υπάρχουν δύο τρόποι προσέγγισης του συμπερασμού, είτε με ακρίβεια είτε προσεγγιστικά. Και οι δύο αυτές προσεγγίσεις είναι στην χειρότερη περίπτωση NP-Hard επιλύσιμες[103], πράγμα που σημαίνει ότι η υπολογιστική πολυπλοκότητά τους είναι μη-ντετερμινιστικά πολυωνυμική, δηλαδή δεν υπάρχει ντετερμινιστικός αλγόριθμος που να επιλύει το πρόβλημα σε πολυωνυμικό (και λιγότερο) χρόνο. Προφανώς, η ακριβής μέθοδος δίνει το ακριβές αποτέλεσμα, ενώ μια προσεγγιστική μέθοδος προσπαθεί να πλησιάσει το σωστό αποτέλεσμα όσο το δυνατόν πιο πολύ. Λόγω της υπολογιστικής πολυπλοκότητας του προβλήματος του συμπερασμού, η ακριβής μέθοδος είναι δυνατή μόνο σε περιορισμένες κατηγορίες δικτύων, και πιο συγκεκριμένα στα δίκτυα που ανήκουν στην κατηγορία των μεμονωμένα-συνδεδεμένων δικτύων (Singly Connected Networks), γνωστών και ως πολυδένδρων (polytrees). Ένα δίκτυο ανήκει σε αυτήν την κατηγορία εάν το υποκείμενο μη κατευθυνόμενο ακυκλικό γράφημα έχει είτε καμία είτε μόλις μία διαδρομή μεταξύ δύο κόμβων. Αυτό το υποκείμενο γράφημα είναι αυτό που λαμβάνει κανείς όταν αγνοήσει την κατεύθυνση των ακμών. Όταν το δίκτυο, από την άλλη, είναι πολλαπλασιαστικά-συνδεδεμένο, είναι δυνατό να χρησιμοποιηθούν τεχνικές ομαδοποίησης (clustering) για να μετατραπεί σε μεμονωμένα-συνδεδεμένο, από το οποίο μπορεί να γίνει ακριβής συμπερασμός. Στην πράξη, όμως, τα δίκτυα είναι τέτοιου μεγέθους ώστε η ακριβής εξαγωγή συμπερασμάτων ή/και η ομαδοποίηση καθίστανται ανέφικτα. Αυτός είναι και ο λόγος για τον οποίο τελικά ακολουθούμε τις προσεγγιστικές μεθόδους. Υπάρχουν πολλές διαφορετικές προσεγγιστικές λύσεις, και ποια είναι η καλύτερη εξαρτάται στο δίκτυο που μας ενδιαφέρει να μελετήσουμε. Οι πιο γνωστοί προσεγγιστικοί αλγόριθμοι είναι η Πιθανολογική Λογική Δειγματοληψία (Probabilistic Logic Sampling)[104], η Πιθανοτική Δειγματοληψία (Likelihood Sampling)[105, 106], η Οπισθοδρομική Δειγματοληψία (Backward Sampling)[107], η Προσαρμοστικής Σπουδαιότητας Δειγματοληψία (Adaptive Importance Sampling)[108], και η Προσεγγιστική εκ των Υστέρων Σπουδαιότητας Δειγματοληψία (Approximate Posterior Importance Sampling)[109].

4.1.4 Διαγράμματα Επιρροής

Ένα Διάγραμμα Επιρροής[110] είναι ένα μπεϊσιανό δίκτυο που συμπεριλαμβάνει κόμβους απόφασης και κόμβους τιμών. Μοντελοποιεί ένα συγκεκριμένο πρόβλημα απόφασης και ο στόχος είναι να επιλεχθεί η εναλλακτική απόφαση για την οποία το αναμενόμενο κέρδος (ή ωφέλεια) έχει την υψηλότερη τιμή. Ένα παράδειγμα διαγράμματος επιρροής φαίνεται στο σχήμα 3.3. Εκτός από τους δύο οβάλ κόμβους, οι οποίοι είναι κανονικοί κόμβοι πιθανότητας όπως σε έναν μπεϊσιανό δίκτυο, υπάρχουν ακόμα τρεις κόμβοι διαφορετικού σχήματος: δύο με ορθογώνιο σχήμα και ένας σε σχήμα ρόμβου. Οι κόμβοι σε σχήμα ορθογωνίου είναι κόμβοι απόφασης (Decision Nodes) και αντιπροσωπεύουν μεταβλητές που βρίσκονται υπό τον έλεγχο του υπεύθυνου λήψης αποφάσεων και μοντελοποιούν τις εναλλα-



Σχήμα 4.2: Παράδειγμα Διαγράμματος Επιρροής

κτικές αποφάσεις που διαθέτει ο λαμβάνων της απόφασης. Κάθε κόμβος απόφασης συνήθως έχει πολλές εναλλακτικές αποφάσεις. Στην πιο απλή μορφή, όπως ο κόμβος «Απόφαση Επένδυσης» (Investment Decision) του σχήματος 3.3 υπάρχουν δύο αποφάσεις: επένδυση (invest) και μη επένδυση (donotinvest). Εάν παρθεί μια συγκεκριμένη απόφαση, δηλαδή, παρατηρηθεί ένας κόμβος απόφασης, θα επηρεαστεί η αξία όλων των απογόνων κόμβων του. Κάθε παιδί έχει μια κατανομή πιθανότητας που καθορίζεται για κάθε πιθανή απόφαση ή για κάθε αναμενόμενο κέρδος, εάν ο κόμβος είναι κόμβος αξίας (Value Node) (βλπ. επόμενη παράγραφο). Για παράδειγμα, η εικόνα 3.4 δείχνει τον πίνακα δεσμευμένων πιθανοτήτων του κόμβου "Επιτυχία του εγχειρήματος" (Success of the venture), ο οποία εξαρτάται από τον κόμβο απόφασης «Ευαισθησία» (Sensitivity).

Probability distributions for different policies:

Sensitivity	Low	Nominal	High
Success	0.1	0.2	0.35
Failure	0.9	0.8	0.65

Εικόνα 4.2: Κατανομή Πιθανοτήτων του κόμβου "Επιτυχία του εγχειρήματος"

Ο κόμβος σε σχήμα ρόμβου, "Οικονομικό Όφελος" (Financial Gain), είναι ένας κόμβος αξίας. Ένας κόμβος αξίας αντιπροσωπεύει την ωφέλεια, δηλαδή αποτελεί ένα μέτρο της επιθυμίας των αποτελεσμάτων της διαδικασίας της απόφασης. Ποσοτικοποιείται από την ωφέλεια καθενός από τους πιθανούς συνδυασμούς αποτελεσμάτων των γονικών κόμβων. Αυτές οι ωφέλειες είναι εκ φύσεως υποκειμενικές και μπορεί να είναι οποιοσδήποτε αριθμός. Ο στόχος ολόκληρης της διαδικασίας λήψης αποφάσεων είναι να μεγιστοποιηθεί αυτός ο αριθμός. Ο ορισμός του κόμβου "Οικονομικό Όφελος" φαίνεται στο Σχήμα 3.5.

Σε ένα διάγραμμα επιρροής, κάθε κατευθυνόμενη ακμή μεταξύ δύο κόμβων απόφασης έχει μια ειδική σημασία. Υποδεικνύει τη σειρά με την οποία λαμβάνονται οι αποφάσεις. Η

Expected utilities for different policies:						
Sensitivity	Low		Nominal		High	
Investment De...	Invest	DonotInvest	Invest	DonotInvest	Invest	DonotInvest
Exp. utility	-3500	500	-2000	500	250	500

Εικόνα 4.3: Αναμενόμενη Ωφέλεια του κόμβου αξίας "Οικονομικό Όφελος"

απόφαση στην ουρά της ακμής αναμένεται να ληφθεί πριν από την απόφαση στην άκρη της. Στο Σχήμα 3.3 υπάρχει μια τέτοιο ακμή μεταξύ του κόμβου "Ευαισθησία" και του κόμβου "Απόφαση Επένδυσης", είναι ένα διακεκομμένο βέλος.

Το διάγραμμα επιρροής του σχήματος 3.3 μοντελοποιεί την ακόλουθη κατάσταση: Είμαστε σε θέση να επενδύσουμε σε ένα συγκεκριμένο εγχείρημα. Εάν επενδύσουμε και το εγχείρημα είναι επιτυχές, τα έσοδά μας (ή το οικονομικό όφελος) θα είναι τα υψηλότερα. Αν επενδύσουμε και το εγχείρημα αποτύχει, χάνουμε όλα μας τα χρήματα. Η τελική επενδυτική απόφαση διαμορφώνεται από τον κόμβο απόφασης «Επενδυτική Απόφαση». Αλλά πριν πάρουμε την απόφαση, πρέπει να πάρουμε μια απόφαση για τον κόμβο "Ευαισθησία". Αυτός ο κόμβος μοντελοποιεί την αβεβαιότητά μας σχετικά με την πιθανότητα ότι το εγχείρημα θα είναι επιτυχές, που με την σειρά του διαμορφώνεται από τον κόμβο "Επιτυχία του Εγχειρήματος". Ο κόμβος απόφασης "Ευαισθησία" έχει τρεις εναλλακτικές αποφάσεις, κάθε μία από τις οποίες καταλήγει σε διαφορετική κατανομή πιθανότητας στον κόμβο «Επιτυχία του εγχειρήματος». Επίσης, πριν λάβουμε την απόφαση για το αν θα επενδύσουμε ή όχι, μπορούμε να συμβουλευτούμε έναν ειδικό που προβλέπει την επιτυχία του εγχειρήματος. Τέλος, ο κόμβος αξίας "Οικονομικό κέρδος" μοντελοποιεί το τελικό αποτέλεσμα της επένδυσής μας, δηλαδή πόσα χρήματα κερδίζουμε ή χάνουμε. Σε αυτή την περίπτωση η ωφέλεια μπορεί να ερμηνευτεί ως χρήμα, αλλά γενικότερα η ωφέλεια δεν έχει μονάδες.

4.1.5 Τύποι Ερμηνευσιμότητας - Επεξηγήσεων στα Μπεϊσιανά Δίκτυα

Στη βιβλιογραφία σχετικά με το θέμα της ερμηνευσιμότητας στα μπεϊσιανά δίκτυα, διακρίνονται τρία είδη επεξηγήσεων[111]. Το πρώτο ονομάζεται Απαγωγή (Abduction). Η απαγωγή είναι η διαδικασία προσδιορισμού των πιο πιθανών τιμών των μη παρατηρούμενων μεταβλητών στο δίκτυο. Μια τέτοια διαμόρφωση αναφέρεται συνήθως ως MPE (Most Probable Explanation), και μπορεί να περιέχει κάθε μη παρατηρούμενη μεταβλητή, που σε αυτή την περίπτωση ονομάζεται ολική απαγωγή, ή μπορεί να περιέχει μόνο ένα υποσύνολο των μη παρατηρούμενων μεταβλητών, όπου και ονομάζεται μερική απαγωγή. Γενικά η απαγωγή περιλαμβάνει τη μεγιστοποίηση της πιθανότητας ενός συνόλου μη παρατηρούμενων μεταβλητών όταν δίνονται μία ή περισσότερες παρατηρήσεις. Είναι επίσης δυνατό να δημιουργηθεί ένα σύνολο MPE, για παράδειγμα πέντε διαμορφώσεις με τις πέντε υψηλότερες πιθανότητες.

Τα άλλα δύο είδη επεξηγήσεων είναι οι στατικές και οι δυναμικές επεξηγήσεις. Μια στατική επεξήγηση λαμβάνει υπόψιν μόνο τις πληροφορίες που περιέχονται στο μοντέλο του μπεϊσιανού δικτύου, δηλαδή, χωρίς να πραγματοποιείται Συλλογισμός. Ένας άλλος τρόπος να περιγράψουμε την στατική ερμηνευσιμότητα είναι ότι προσφέρει επεξηγήσεις από

τις παραδοχές του μοντέλου. Μια στατική επεξήγηση θα μπορούσε, για παράδειγμα, να κάνει σαφείς τις δηλώσεις ανεξαρτησίας που περιέχονται στο μοντέλο, ή θα μπορούσε να περιγράψει την εκ των προτέρων (prior) πιθανότητα των μεταβλητών. Μια δυναμική επεξήγηση, από την άλλη πλευρά, είναι μια επεξήγηση της συλλογιστικής διαδικασίας σε ένα μπεϊσιανό δίκτυο. Έτσι, δεδομένου ενός ή περισσότερων παρατηρήσεων και μιας μεταβλητής ενδιαφέροντος, μια δυναμική επεξήγηση προσπαθεί να δώσει στον χρήστη μια εικόνα για τη διαδικασία που προκάλεσε τη μεταβλητή ενδιαφέροντος να επηρεαστεί με τον τρόπο που εκείνη επηρεάστηκε. Πιο συγκεκριμένα, προσπαθεί να εξηγήσει τις αλλαγές στην εκ των υστέρων πιθανότητα της μεταβλητής ενδιαφέροντος σε σχέση με τις παρατηρήσεις. Αυτό το είδος επεξήγησης, μπορεί κανείς να υποστηρίξει, ότι προσπαθεί να απαντήσει στις ερωτήσεις: «Ποιες ήταν οι πιο σημαντικές παρατηρήσεις;» και «Γιατί έχει επιρροή μια συγκεκριμένη παρατήρηση;». Μια παρατήρηση έχει επιρροή όταν επηρεάζει την εκ των υστέρων πιθανότητα της μεταβλητής ενδιαφέροντος είτε με θετικό είτε με αρνητικό τρόπο. Μέσα στις δυναμικές επεξηγήσεις μπορεί να γίνει κι άλλη διάκριση. Υπάρχει επουσιώδης διαφορά μεταξύ μικροσκοπικών και μακροσκοπικών δυναμικών επεξηγήσεων[112]. Η μικροσκοπική δυναμική επεξήγηση προσπαθεί να δικαιολογήσει τις παραλλαγές της κατανομής πιθανότητας ενός συγκεκριμένου κόμβου, ενώ μια μακροσκοπική δυναμική επεξήγηση προσπαθεί να κάνει σαφείς στον χρήστη τις βασικές γραμμές συλλογισμού από τις παρατηρήσεις στη μεταβλητή ενδιαφέροντος και ως εκ τούτου λαμβάνει υπόψιν ένα μεγαλύτερο μέρος του μοντέλου.

4.1.6 Αποδίδοντας Ερμηνεύσεις - Επεξηγήσεις

Μια επεξήγηση θα πρέπει να παρουσιάζεται με τρόπο αποτελεσματικό, βολικό, καθώς και εύκολα προσβάσιμο. Η διάκριση της απόδοσης που μπορεί να πραγματοποιηθεί από αυτή την άποψη είναι κυρίως μεταξύ των λεκτικών και γραφικών επεξηγήσεων. Μια λεκτική επεξήγηση θα μπορούσε να είναι για παράδειγμα: «Η μεταβλητή A εξαρτάται από μεταβλητή B, αλλά δεδομένης της μεταβλητής Γ είναι ανεξάρτητες» ή «Η κατάσταση 0 είναι περισσότερο πιθανή από την κατάσταση 1». Μια γραφική επεξήγηση χρησιμοποιεί γραφικά μέσα για να επικοινωνήσει μια ερμηνεία. Η πιο προφανής και βασική επεξήγηση αυτού του τύπου είναι η οπτικοποίηση της δομής του δικτύου. Εάν ο χρήστης έχει αρκετές γνώσεις σχετικά με τα μπεϊσιανά δίκτυα, μπορεί να συναγάγει τις εξαρτήσεις και μη μεταξύ των μεταβλητών στον τομέα εφαρμογής του μοντέλου. Ένα άλλο παράδειγμα είναι η εμφάνιση των πιθανοτήτων των διαφόρων καταστάσεων μιας μεταβλητής χρησιμοποιώντας γραφικές ράβδους που κυμαίνονται από το μηδέν έως εκατό τοις εκατό[99].

4.2 Αρχές Γενικού Σχεδιασμού

4.2.1 Πρόγονοι και απόγονοι

Ένα απλό αλλά δυνητικά πολύ χρήσιμο χαρακτηριστικό είναι η δυνατότητα να βλέπεις εύκολα, για έναν συγκεκριμένο κόμβο, ποιοι είναι οι γονείς και τα παιδιά του. Σε ένα περίπλοκο δίκτυο με πολλά δεδομένα δεν είναι πάντα ξεκάθαρο ποιοι κόμβοι συνδέονται με έναν συγκεκριμένο κόμβο. Γενικότερα, πιστεύουμε ότι θα ήταν χρήσιμο να επισημαίνονται γραφικά οι πρόγονοι ή/και οι απόγονοι ενός κόμβου. Ο χρήστης μπορεί έτσι να είναι σε

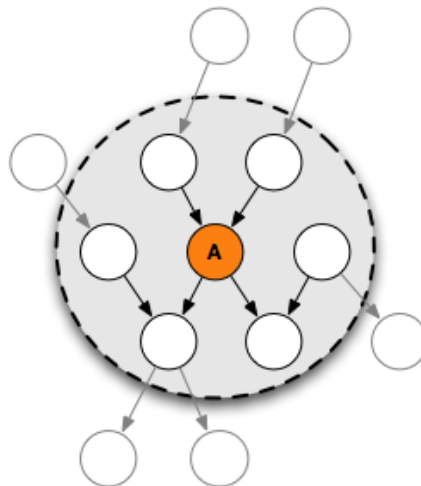
θέση να πει πόσες γενιές προγόνων ή απογόνων επιθυμεί να δει. Για παράδειγμα, οι γονείς μπορούν να είναι οι πρόγονοι της πρώτης γενιάς και τα παιδιά να είναι οι απόγονοι επίσης της πρώτης γενιάς.

4.2.2 Μανδύας Markov

Ο μανδύας Markov ενός κόμβου σε ένα Μπεϊσιανό δίκτυο είναι το σύνολο των κόμβων που αποτελούνται από τους γονείς του κόμβου, τα παιδιά του και τα παιδιά των γονέων του. Ο μανδύας Markov είναι μια ενδιαφέρουσα έννοια, επειδή "προστατεύει" τον κόμβο από το υπόλοιπο δίκτυο. Με δεδομένο τον κόμβο A και τον μανδύα Markov του $Markov(A)$, ο κόμβος είναι ανεξάρτητος από το υπόλοιπο δίκτυο:

$$P(A|Markov(A), B) = P(A|Markov(A)) \quad (4.3)$$

όπου B είναι ένα σύνολο άλλων κόμβων στο δίκτυο. Αυτό σημαίνει ότι οι κόμβοι που αποτελούν μέρος του μανδύα Markov ενός συγκεκριμένου κόμβου, είναι οι μόνοι κόμβοι που χρειάζονται για να προβλέψει κανείς την συμπεριφορά αυτού του κόμβου (Σχήμα 4.1).

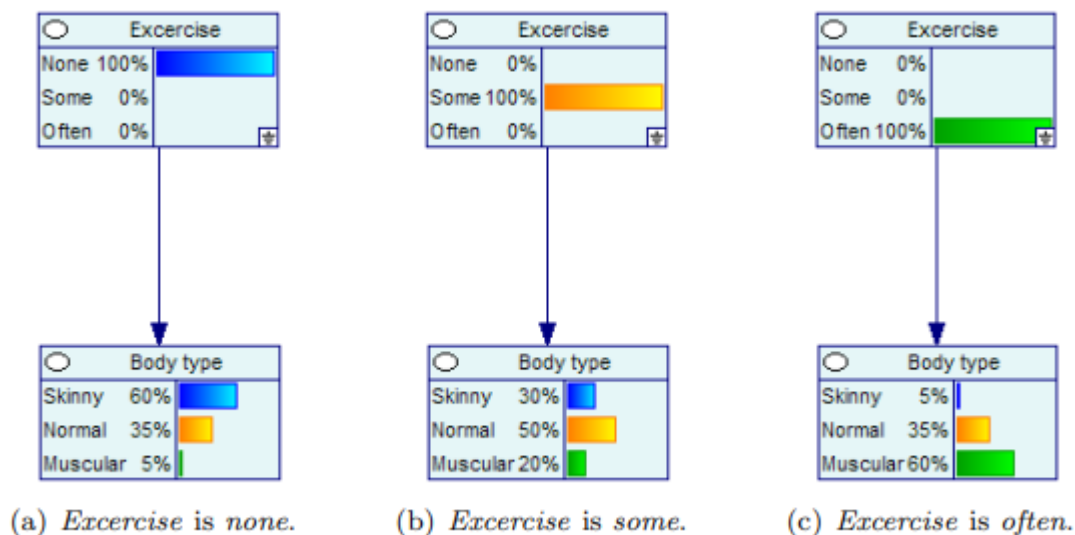


Σχήμα 4.3: Μανδύας Markov κόμβου A

4.2.3 Διατεταγμένοι Κόμβοι

Για να υλοποιήσουμε τους στόχους μας πρέπει να εισαγάγουμε την έννοια των διατεταγμένων κόμβων (Ordinal Nodes). Σε ένα μπεϊσιανό δίκτυο, οι διάφορες καταστάσεις που έχει ένας κόμβος δεν απαιτείται να είναι ορισμένες με οποιαδήποτε σειρά. Αν όμως έχουμε κόμβους με διατεταγμένες καταστάσεις, πράγμα που συμβαίνει συχνά σε ένα διαγνωστικό δίκτυο, μπορούμε να λάβουμε υπόψιν αυτές τις επιπλέον πληροφορίες κατά τον προσδιορισμό του βαθμού της επιρροής μεταξύ των δύο συνδεδεμένων κόμβων, ώστε να συμπεράνουμε μια εξήγηση που να δείχνει εάν η επιρροή είναι θετική ή αρνητική. Για παράδειγμα στο Σχήμα 4.2. απεικονίζεται ένα δίκτυο που αποτελείται από δύο κόμβους, "Άσκηση" (Exercise) και "Τύπος Σώματος" (Body Type) (σχεδιάστηκαν ως ορθογώνιοι, αλλά είναι κανονικοί τυχαίοι

κόμβοι κι όχι κόμβοι απόφασης). Και οι δύο έχουν τρεις καταστάσεις και είναι διατεταγμένες. Σε αυτή την περίπτωση οι καταστάσεις και των δύο κόμβων είναι σε αυξουσα σειρά. Σε κάθε μία από τις εικόνες παρατηρείται μια κατάσταση. Στο σχήμα 4.2 (a) παρατηρείται η τιμή "Καμία" (none), στο Σχήμα 4.2 (b) παρατηρείται "Κάποια" (some) και το "Συχνά" (often) παρατηρείται το σχήμα 4.2 (c). Όταν παρατηρείται υψηλότερη τιμή "Άσκησης" αυξάνονται επίσης οι πιθανότητες των υψηλότερων τιμών του "Τύπου Σώματος". Σε αυτή την περίπτωση λοιπόν η "Άσκηση" επηρεάζει θετικά τον "Τύπο Σώματος".

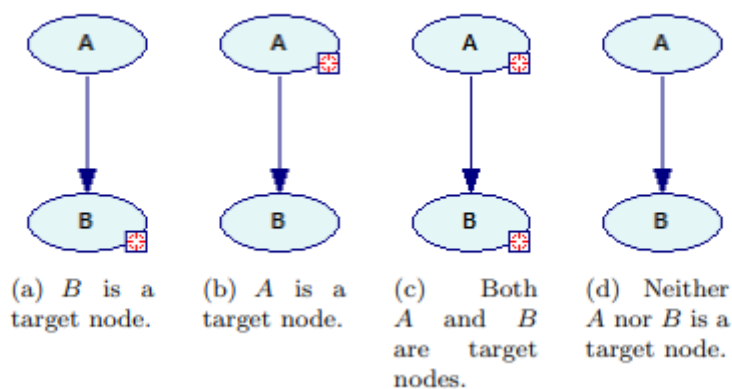


Σχήμα 4.4: Παράδειγμα Διατεταγμένων Κόμβων

4.3 Δυναμική Στρατηγική

Οι πληροφορίες που θέλουμε να παρέχουμε στον αναγνώστη είναι το μέγεθος της επιρροής που έχει ένας κόμβος σε έναν άλλο. Οι πλειοψηφία των μέχρι πρότινος προσεγγίσεων όπως περιγράφονται στην ενότητα 2 χρησιμοποιούν την κοινή κατανομή πιθανότητας, ενώ άλλες καθορίζουν την επιρροή κοιτάζοντας τους πίνακες δεσμευμένων πιθανοτήτων και τον προσδιορισμό της επιρροής που έχει ένας γονικός κόμβος σε έναν παιδικό κόμβο. Αυτή η προσέγγιση για τον προσδιορισμό της επιρροής ενός γονικού κόμβου σε έναν κόμβο παιδί είναι στατικός. Αυτό σημαίνει ότι οι υπολογισμοί δεν λαμβάνουν υπόψη τυχόν τρέχουσες παρατηρήσεις. Θα μπορούσε όμως να είναι, με ορισμένες παρατηρήσεις, η επίδραση ενός συγκεκριμένου γονικού κόμβου σε έναν κόμβο παιδί σημαντικά διαφορετική από την περίπτωση χωρίς παρατήρηση, κι έτσι θα κατέληγαν οι στατικές πληροφορίες ανακριθής. Οι στατικές πληροφορίες μπορούν να χρησιμοποιηθούν για να αποκτήσουν μια γενική εικόνα του αλληλεπιδράσεις μεταξύ των κόμβων, αλλά δεν είναι προσαρμοσμένες σε μια συγκεκριμένη κατάσταση.

Εκτός από αυτό, ενώ είναι αλήθεια ότι ένας γονικός κόμβος επηρεάζει το παιδί ή τα παιδιά του εάν εκείνος παρατηρηθεί, ένα παιδί, όταν παρατηρείται, μπορεί επίσης να επηρεάσει την κατανομή πιθανότητας του γονέα. Αυτές οι δύο επιρροές μπορεί να είναι αρκετά διαφορετικές



Σχήμα 4.5: Τέσσερις διαφορετικές καταστάσεις

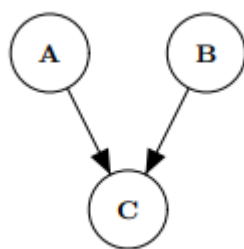
Πίνακας 4.1: Πίνακας δεσμευμένης πιθανότητας για το δίκτυο του σχήματος 4.4. Βρίσκοντας τον *A* επιπροσθέτως με τον *B*, ή το ανάποδο, θα έχει μικρή επίδραση στον *C*.

	a		-a	
	b	-b	b	-b
c	0.99	0.98	0.98	0.01
-c	0.01	0.02	0.02	0.99

μεταξύ τους. Για παράδειγμα, εάν πέσει ένας φορητός υπολογιστής από ένα υψηλό κτίριο, το ξέρουμε σχεδόν σίγουρα ότι θα καταλήξει να σπάσει σε πολλά κομμάτια. Αν όμως βρούμε ένα φορητό υπολογιστή που έχει σπάσει σε πολλά κομμάτια, δεν μπορούμε να είμαστε το ίδιο σίγουροι για το τι το προκάλεσε αυτό. Θα μπορούσε να είχε πέσει από ένα ψηλό κτίριο, αλλά θα μπορούσε επίσης να έχει περάσει από πάνω του ένα αυτοκίνητο, ή ίσως κάποιος νευρίασε και το πάτησε. Έτσι, ενώ η πιθανότητα "έχει πέσει από ένα υψηλό κτίριο" θα αυξηθεί όταν βρεθεί ένας κατεστραμμένος φορητός υπολογιστής, η πιθανότητα δεν θα αυξηθεί τόσο πολύ όσο αντίστροφα, δηλαδή, αυτή του "ο φορητός υπολογιστής θα καταστραφεί" όταν το αφήσουμε από ένα ψηλό κτίριο.

Επομένως, θα προσεγγίσουμε την εργασία διαφορετικά. Πολλά δίκτυα περιέχουν ένα ή περισσότερους κόμβους στόχους ή κόμβους υπόθεσης. Αυτό που θα μας ενδιαφέρει περισσότερο είναι η επιρροή των άλλων κόμβων που έχουν σε αυτούς τους κόμβους. Όταν ένα βέλος συνδέει έναν κόμβο στόχο με έναν κόμβο μη-στόχο, θα προσδιορίζει την επίδραση που έχει ο κόμβος μη στόχος στον κόμβο στόχο, ανεξάρτητα της κατεύθυνσης του βέλους. Όταν υπάρχουν δύο μη στόχοι ή δύο κόμβοι στόχοι που συνδέονται με ένα βέλος, θα χρησιμοποιούμε, από προεπιλογή, τον μέσο όρο των επιρροών και προς τις δύο κατευθύνσεις. Συνολικά υπάρχουν τέσσερις πιθανές καταστάσεις, που φαίνονται στο σχήμα 5.3. Στην πρώτη κατάσταση (Σχήμα 5.3.α), θα απεικονίσουμε την επίδραση που έχει ο *A* στον *B*. Στη δεύτερη περίπτωση, αυτή του Σχήματος 5.3.β, θα εξετάσουμε η επίδραση που έχει ο *B* στον *A*. Στις δύο τελευταίες καταστάσεις, που απεικονίζονται στα σχήματα 5.3.γ και 5.3.δ, θα εξετάσουμε την επιρροή και προς τις δύο κατευθύνσεις και ύστερα τον μέσο όρο τους.

Επιπλέον, η προσέγγιση που προτείνουμε είναι δυναμική: εξετάζουμε το δίκτυο στην τρέχουσα κατάσταση, συμπεριλαμβανομένων τυχόν παρατηρήσεων. Ουσιαστικά υποδεικνύει



Σχήμα 4.6: Ένα μικρό δίκτυο

πόση πιθανή επίδραση έχει ένας κόμβος σε έναν άμεσο διάδοχο ή προκάτοχο, και επομένως και την επιρροή που θα μπορούσε να έχει ένας κόμβος αν παρατηρηθεί στη συνέχεια. Αυτό μπορεί να δώσει πολύ μεγαλύτερη εικόνα σε μια κατάσταση. Για παράδειγμα, λάβετε υπόψη το Σχήμα 5.4, που δείχνει ένα απλό δίκτυο που αποτελείται από δύο γονικούς κόμβους, τον A και τον B, με κοινό θυγατρικό κόμβο C, και όπου ο C είναι ένας κόμβος στόχος. Χωρίς παρατηρήσεις, θεωρούμε ότι η πιθανή επιρροή που έχει ο A στο Σ είναι μεγαλύτερη από την πιθανή επιρροή που έχει ο B στο C. Αν παρατηρήσουμε τον B, αυτό θα μπορούσε να έχει τόσο μεγάλο αντίκτυπο στην κατανομή της πιθανότητας του Σ , ώστε η παρατήρηση του A επιπλέον στο B να σημαίνει μικρή ή καθόλου διαφορά. Αυτό σημαίνει ότι η πιθανή επιρροή του A στον Σ έχει μειωθεί σημαντικά. Ο πίνακας δεσμευμένων πιθανοτήτων του Πίνακα 5.1 διαφωτίζει αυτή την κατάσταση. Όλοι οι κόμβοι έχουν δύο καταστάσεις: αληθής και ψευδής. Αν ο A είναι αληθής και ο B είναι ψευδής, τότε η διαπίστωση ότι ο B είναι αληθής δεν έχει μεγάλο αντίκτυπο στην πιθανότητα ότι ο Σ είναι αληθής. Στην προσέγγισή μας, όταν παρατηρείται είτε ο A είτε ο B, το πάχος του τόξου θα υπολογίζεται εκ νέου για να αντικατοπτρίζει τη νέα κατάσταση.

Ένα άλλο πλεονέκτημα της προσέγγισής μας σε αντίθεση με τη στατική είναι το γεγονός ότι σε ορισμένες περιπτώσεις δεν χρειάζεται να ληφθεί υπόψη η συνέργεια μεταξύ των διαφορετικών γονέων ενός κόμβου με περισσότερους από έναν γονείς, απλώς και μόνο επειδή δεν είναι εκεί πια σε αυτές τις καταστάσεις. Ο ορισμός της συνέργειας μπορεί να δοθεί ως εξής: *«Η αλληλεπίδραση δύο ή περισσότερων παραγόντων ή δυνάμεων έτσι ώστε το συνδυαστικό τους αποτέλεσμα είναι μεγαλύτερο από το άθροισμα των επιμέρους αποτελεσμάτων τους»*. Σε περιπτώσεις μπεϊσιανών δικτύων, αυτό ισχύει για τη συνδυασμένη επίδραση της παρατήρησης πολλαπλών γονέων ενός κόμβου σε αυτόν τον κόμβο. Το συνδυασμένο αποτέλεσμα μπορεί να είναι μεγαλύτερο από τις μεμονωμένες επιπτώσεις. Ο πίνακας 4.2 δίνει ένα παράδειγμα αυτού. Κοιτάζοντάς τον παρατηρούμε ότι όταν και ο A και ο B είναι ψευδείς, η πιθανότητα του C είναι 0,01. Όταν είτε ο A είτε ο B είναι αληθείς, η πιθανότητα του C αυξάνεται στο 0,1. Όταν όμως και οι δύο A και B είναι αληθείς, η πιθανότητα του C αυξάνεται πολύ περισσότερο, στο 0,99. Άλλος τρόπος για να γίνει αντιληπτό αυτό είναι να από το Σχήμα 5.5, το οποίο δείχνει οπτικά μια συνέργεια μεταξύ δύο μεταβλητών x και y , που και οι δύο κυμαίνονται από 0 έως 1, και ορίζονται ως $(x + y)^2$. Όταν μόνο μία από τις δύο μεταβλητές προσεγγίζει το 1, η τιμή δεν είναι τόσο υψηλή όσο όταν και το x και το y προσεγγίζουν το 1.

Αυτό το φαινόμενο δεν μπορεί να συλληφθεί με ακρίβεια με το πάχος των ακμών, το οποίο

Πίνακας 4.2: Άλλος πίνακας δεσμευμένων πιθανοτήτων για το δίκτυο 4.4. Η συνδυαστική επίδραση των γονεϊκών κόμβων είναι πολύ μεγαλύτερη από το άθροισμα των μεμονωμένων επιδράσεων.

	a		-a	
	b	-b	b	-b
c	0.99	0.1	0.1	0.01
-c	0.01	0.9	0.9	0.99

είναι μονοδιάστατο. Στη δυναμική προσέγγιση, όταν όλοι πλην ενός από τους γονείς ή τα παιδιά ενός κόμβου παρατηρούνται, πράγμα που σημαίνει, ότι δεν υπάρχει συνέργεια πια, είμαστε σε θέση να εμφανίσουμε με ακρίβεια την πραγματική κατάσταση, γιατί εξετάζουμε πιθανές επιδράσεις στην τρέχουσα κατάσταση του δικτύου. Όταν και μόλις υπάρξει αλλαγή στο δίκτυο, για παράδειγμα μια άλλη παρατήρηση, το πάχος ενός τόξου επαναυπολογίζεται εάν είναι απαραίτητο. Τότε, θα καθορίσουμε τη δύναμη της επιρροής εξετάζοντας την εκ των προτέρων κατανομή πιθανότητας ενός κόμβου, για κάθε πιθανή κατάσταση του γονικού ή παιδικού κόμβου, ανάλογα με τον τύπο σύνδεσης όπως συζητήθηκε νωρίτερα (Εικόνα 4.3). Για έναν κόμβο με n καταστάσεις, αυτό θα οδηγήσει σε n δυνητικά διαφορετικές εκ των προτέρων κατανομές πιθανότητας του συνδεδεμένου κόμβου. Θα υπολογίσουμε το μέγεθος της διαφοράς μεταξύ αυτών των κατανομών και θα βασίσουμε τον τελικό προσδιορισμό της επιρροής είτε στο μέσο όρο όλων των διαφορών, είτε στο μέγιστο όλων των διαφορών είτε στον σταθμισμένο μέσο όρο. Ο σταθμισμένος μέσος όρος ορίζεται ως

$$\sum_{i=0}^n a_i D(P(A), P(B|A = a_i)) \quad (4.4)$$

όπου οι A και B αποτελούν απευθείας συνδεδεμένους κόμβους, ο A έχει n καταστάσεις και D είναι η συνάρτηση μετρικής των αποστάσεων μεταξύ των δύο κατανομών (Κεφάλαιο 5).

Αυτή η μέθοδος προϋποθέτει ότι για κάθε μία από τις καταστάσεις κάθε κόμβου που είναι γονέας ή παιδί, οι πιθανότητες όλων των άμεσων προκατόχων και διαδόχων του πρέπει να έχουν ενημερωθεί. Εάν ένα δίκτυο έχει N κόμβους και κάθε κόμβος έχει n καταστάσεις αυτό θα απαιτούσε, στη χειρότερη περίπτωση, $N * n$ ενημερώσεις του δικτύου. Αυτό οδηγεί σε πολυπλοκότητα τάξης $O(N)$. Η ίδια η εκτέλεση συμπερασμάτων έχει αποδειχθεί ότι είναι NP-hard [103], αλλά με την χρήση της τρέχουσας κατάστασης και την ταχύτητα των αλγορίθμων μπεϊσιανής ενημέρωσης (Κεφάλαιο 5), αυτή η διαδικασία, για τα περισσότερα δίκτυα, θα ολοκληρωθεί μέσα σε κάποια λεπτά.

4.4 Σύγκριση των Κατανομών

Δεδομένου δύο κατανομών πιθανοτήτων διακριτών τιμών, θέλουμε να γνωρίζουμε πόσο αυτές διαφέρουν ώστε να μπορούμε να κάνουμε μια συστηματική σύγκριση μεταξύ των διαφόρων διαφορών και να εξάγουμε στην συνέχεια έγκυρα συμπεράσματα για το ποια αλλαγή είναι πιο σημαντική ή πιο επιδραστική από την άλλη. Στην περίπτωσή μας, υπάρχουν δύο διαφορετικά πιθανά σενάρια, ένα όπου οι καταστάσεις των κατανομών είναι διατεταγμένες και ένα στο οποίο δεν είναι. Αν οι καταστάσεις των κατανομών είναι διατεταγμένες, ίσως

Θελήσουμε να προσαρμόσουμε την ερμηνεία μας για μια αλλαγή μιας κατανομής. Η κύρια σκέψη είναι ότι, εάν οι καταστάσεις διατάσσονται από αριστερά προς τα δεξιά, από το λιγότερο σημαντικό στο πιο σημαντικό, τόσο περισσότερο μετατοπίζεται η πιθανότητα από τη μία πλευρά της κατανομής στην άλλη πλευρά, δηλαδή τόσο πιο σημαντική είναι η αλλαγή. Αν έχουμε τρεις κατανομές, $A = [1, 0, 0, 0]$, $B = [0, 1, 0, 0]$ και $C = [0, 0, 0, 1]$, τότε, στην περίπτωση που οι τέσσερις καταστάσεις δεν διατάσσονται με κανέναν τρόπο, η αλλαγή από την A στη B θα θεωρούνταν εξίσου μεγάλη με την αλλαγή από την A στην C . Αν από την άλλη, γνωρίζουμε ότι οι καταστάσεις είναι σε αύξουσα σειρά διατεταγμένες, θα θεωρούσαμε ότι η αλλαγή από την A στην B είναι λιγότερο σημαντική από την αλλαγή από την A προς την C .

Πρέπει επίσης να δώσουμε προσοχή στον τρόπο με τον οποίο θα αντιμετωπίσουμε τις διαφορές όταν μία από τις πιθανότητες πλησιάζει τις ακραίες τιμές, δηλαδή είτε το 0 είτε το 1. Πρόκειται, για παράδειγμα, να αντιμετωπίσουμε μια διαφορά πιθανότητας από το 0.0001 έως το 0.01 με τον ίδιο τρόπο όπως από το 0.71 στο 0.72. Ίσως θελήσουμε να θεωρήσουμε το πρώτο πιο σημαντικό αύξηση από το τελευταίο, ενώ και στις δύο περιπτώσεις η διαφορά είναι στο 0.01. Αν φανταζόμαστε έναν μετεωρολόγο που προβλέπει 71% πιθανότητα βροχής, ενώ το η πραγματική πιθανότητα είναι 72%, θα λέγαμε ότι έκανε καλή δουλειά. Αν όμως ένας ειδικός προβλέπει πιθανότητα 1% εμφάνισης σοβαρής νόσου, όταν η πραγματική πιθανότητα είναι μόλις 0.01%, πιθανότατα θα λέγαμε ότι δεν ήταν πολύ ακριβής. Αυτό οδηγεί στη διάκριση μεταξύ απόλυτων και σχετικών διαφορών. Η απόλυτη διαφορά εκφράζεται στις ίδιες μονάδες με τις δύο συγκριθείσες τιμές. Εκεί δεν υπάρχουν μονάδες με σχετική διαφορά, εκφράζονται σε ποσοστά. Στο δικό μας παράδειγμα οι απόλυτες διαφορές είναι 0.0099 και 0.01, αντίστοιχα, ενώ οι σχετικές διαφορές είναι:

$$\frac{|0.71 - 0.72|}{0.71} * 100 = 1.41\% \quad (4.5)$$

και

$$\frac{|0.01 - 0.0001|}{0.01} * 100 = 99\% \quad (4.6)$$

Έτσι, οι απόλυτες διαφορές είναι σχεδόν ίσες, αλλά οι σχετικές διαφορές δεν είναι καθόλου ίσες.

Επίσης, πρέπει να συγκρίνουμε ποικίλες διαφορές μεταξύ τους, κάτι που απαιτεί ότι η μετρική μας είναι συμμετρική. Συμμετρική σημαίνει ότι, για μια ορισμένη μετρική απόστασης D και για δύο κατανομές πιθανότητας P και Q , ισχύει το ακόλουθο[99]:

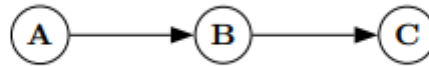
$$D(P, Q) = D(Q, P) \quad (4.7)$$

Αν κάναμε συγκρίσεις μόνο από μία πλευρά, εννοώντας ότι η μία από αυτές τις δύο κατανομές είναι σταθερή κατά τη διάρκεια των συγκρίσεων, θα μπορούσαμε απλώς να φροντίσουμε να χρησιμοποιούμε την ίδια σειρά ορισμάτων στη συνάρτηση απόστασης κάθε φορά. Αυτό όμως δεν ισχύει στα μπεϊσιανά δίκτυα. Πιο συγκεκριμένα, για παράδειγμα, στο δίκτυο του σχήματος 4.5, αν θέλουμε να καθορίσουμε την επιρροή των κόμβων προς την κατεύθυνση των ακμών, δηλαδή από γονέα σε παιδί, υπάρχουν δύο επιρροές: αυτή του A στον B

και αυτή του B στον C. Για να καθορίσουμε το μέγεθος της επιρροής θα υπολογίζαμε την $D(P(B), P(B|A))$ και $D(P(C), P(C|B))$. Τώρα, υποθέτοντας ότι όλοι οι κόμβοι είναι δυαδικοί, αν $P(B) = [0, 5, 0, 5]$, $P(B|A) = [0.9, 0.1]$, $P(C) = [0.9, 0.1]$ και $P(C|B) = [0.5, 0.5]$, θα μπορούσαμε να συναντήσουμε πρόβλημα κατά τη χρήση ενός ασύμμετρου μέτρου, γιατί ουσιαστικά θα καταλήγαμε στην παρακάτω σύγκριση:

$$D([0.5, 0.5], [0.9, 0.1]) == D([0.9, 0.1], [0.5, 0.5]) \quad (4.8)$$

Απαιτείται να χρησιμοποιήσουμε ένα συμμετρικό μέτρο, διαφορετικά θα μπορούσαμε να λάβουμε δύο διαφορετικές τιμές ενώ θέλουμε να θεωρήσουμε αυτές τις δύο διαφορές ίσες.



Σχήμα 4.7: Απλό δίκτυο τριών κόμβων

Τέλος, θα θέλαμε να έχουμε ένα μέτρο που μας δίνει τιμές σε ένα συγκεκριμένο περιορισμένο εύρος, κατά προτίμηση από το 0 έως το 1, αφού εκφράζουμε πιθανότητες. Αυτό θα παρέχει μια εύκολη άμεση χαρτογράφηση στο πάχος μιας ακμής. Θα μπορούσαμε φυσικά να χρησιμοποιήσουμε ένα απεριόριστο μέτρο και απλώς να καταχωρήσουμε την υψηλότερη διαφορά στο παχύτερο τόξο και να καθορίσουμε το πάχος από όλα τα άλλα τόξα σε σχέση με το παχύτερο με κανονικοποίηση. Αλλά έτσι δεν θα μπορούμε, για παράδειγμα, να συγκρίνουμε το πάχος των ακμών σε δύο διαφορετικά δίκτυα μεταξύ τους, επειδή οι δύο υψηλότερες τιμές δύο δικτύων μπορεί να διαφέρουν σημαντικά και ως εκ τούτου ένα ορισμένο πάχος ενός τόξου σε ένα δίκτυο μπορεί να υποδηλώνει μια πολύ διαφορετική ποσότητα επιρροής από το ίδιο πάχος σε άλλο δίκτυο.

4.5 Μετρικές Αποστάσεων

Σε αυτήν την ενότητα θα αναλύσουμε τις σημαντικότερες μετρικές αποστάσεων. Για τον σκοπό αυτό θα ορίσουμε τις P και Q ως δύο διακριτές κατανομές πιθανότητας:

$$P, Q \in \left\{ (p_1, p_2, \dots, p_n) \mid p_i > 0, \sum_{i=1}^n p_i = 1 \right\}, n > 1 \quad (4.9)$$

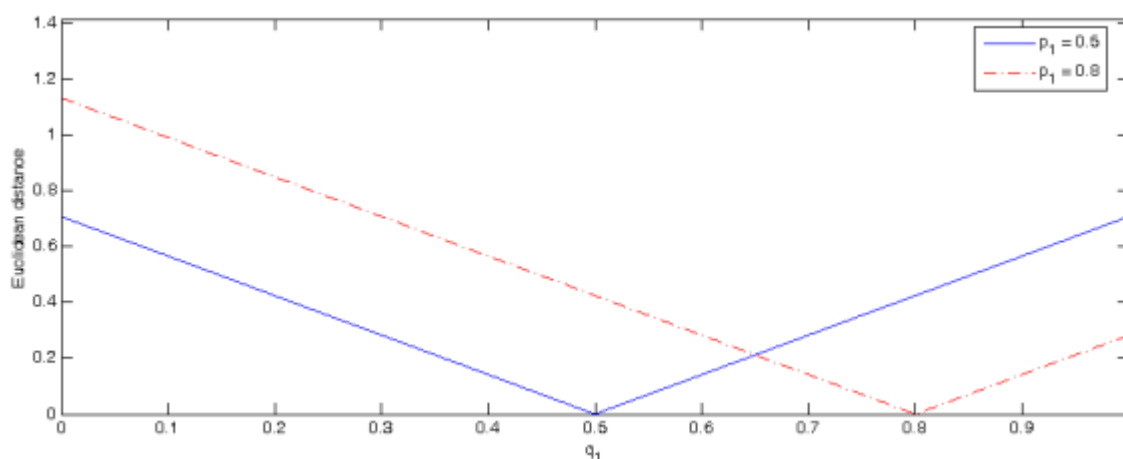
4.5.1 Ευκλείδεια Απόσταση

Μια από τις κατεξοχήν διαδεδομένες αποστάσεις είναι η Ευκλείδεια, και ορίζεται ως εξής:

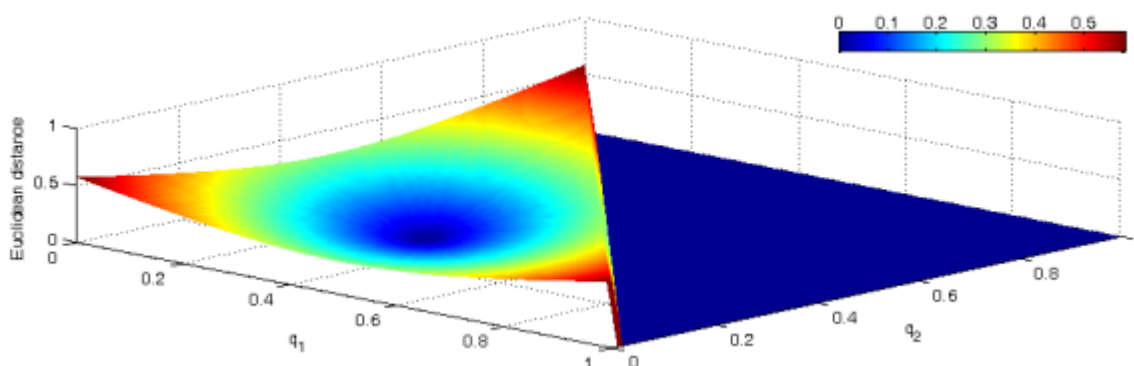
$$E(P, Q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (4.10)$$

Αν P και Q είναι δύο σημεία του N -διάστατου χώρου, η παραπάνω έκφραση υπολογίζει την πραγματική (φυσική) χωρική απόσταση μεταξύ των δύο σημείων. Όταν χρησιμοποιείται με διακριτές κατανομές πιθανοτήτων, όπου το άθροισμα όλων των στοιχείων είναι πάντα ίσο με 1, η τιμή της μέτρησης κυμαίνεται πάντα από το 0, όπου δηλαδή δεν υπάρχει διαφορά, μέχρι το $\sqrt{2}$, την μέγιστη διαφορά. Η Ευκλείδεια απόσταση είναι συμμετρική εξ' ορισμού. Στα σχήματα 4.6-4.8 φαίνεται η συμπεριφορά της Ευκλείδειας σε διάφορες καταστάσεις. Μπορούμε πολύ εύκολα να μετατρέψουμε το εύρος τιμών της Ευκλείδειας σε $[0, 1]$ με την ακόλουθη νόρμα:

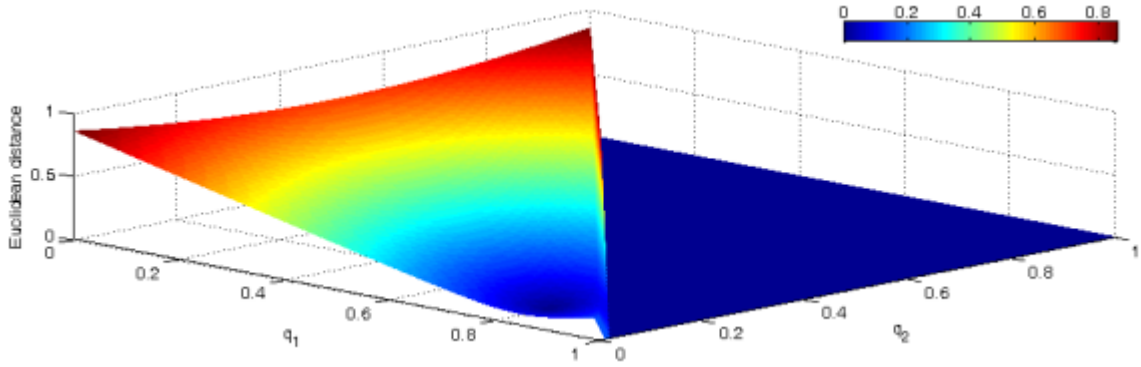
$$E_{norm}(P, Q) = \frac{E(P, Q)}{\sqrt{2}} \tag{4.11}$$



Σχήμα 4.8: Ευκλείδειες αποστάσεις της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$



Σχήμα 4.9: Ευκλείδεια απόσταση της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0



Σχήμα 4.10: Ευκλείδεια απόσταση της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0

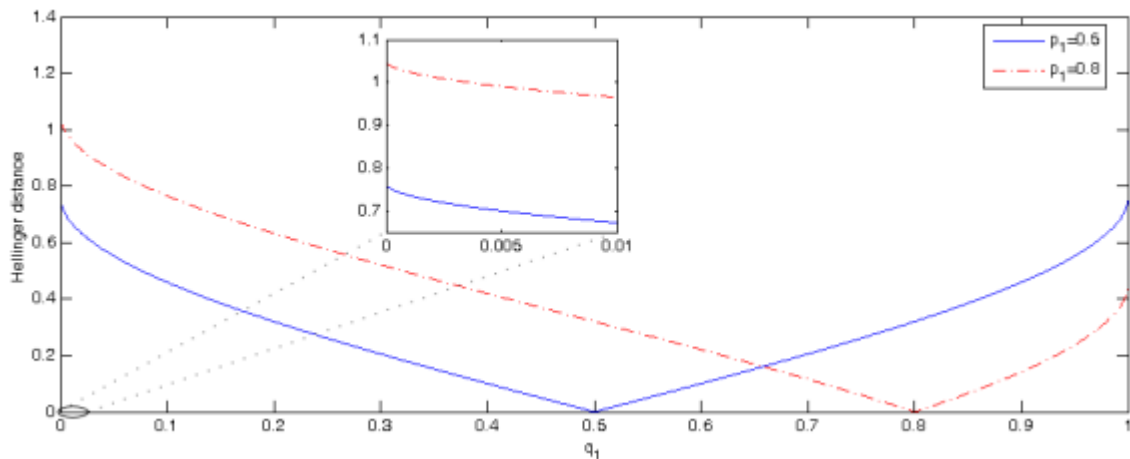
4.5.2 Απόσταση Hellinger

Στον συμπερασμό των μπεϊσιανών δικτύων χρησιμοποιείται πολύ συχνά η απόσταση Hellinger[113], η οποία ορίζεται ως εξής:

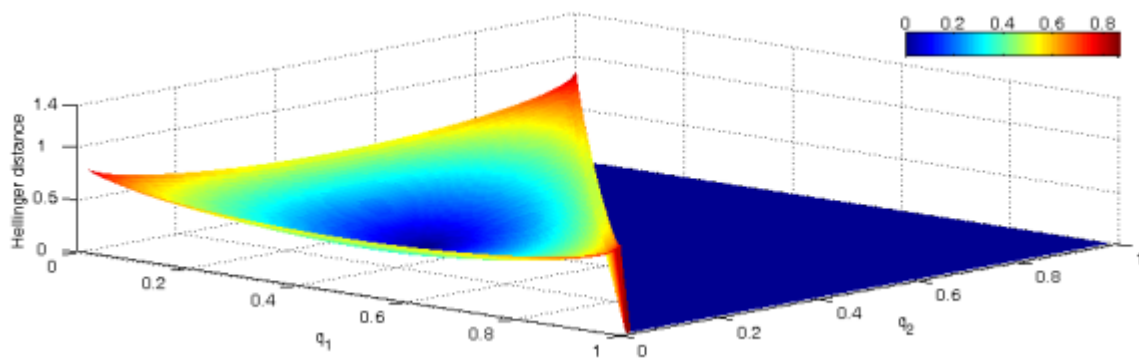
$$H(P, Q) = \sqrt{\sum_{i=1}^n (\sqrt{p_i} - \sqrt{q_i})^2} \quad (4.12)$$

Η τιμές αυτής της μετρικής κυμαίνονται από 0 έως $\sqrt{2}$, όπως και η Ευκλείδεια. Στα σχήματα 4.9-4.11 δείχνουν την συμπεριφορά της απόστασης Hellinger. Γίνεται αντιληπτό ότι η Hellinger είναι πιο ευαίσθητη όταν πλησιάζει τις ακραίες τιμές 0 ή 1, συγκριτικά με την Ευκλείδεια απόσταση. Η Hellinger είναι επίσης συμμετρική και μπορούμε, πάλι, να ορίσουμε έναν γραμμικό μετασχηματισμό για να λάβουμε εύρος τιμών $[0, 1]$:

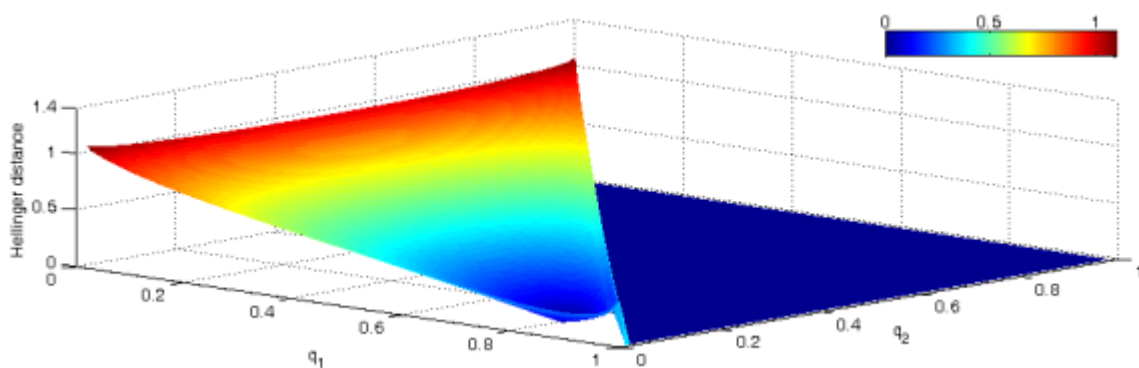
$$H_{norm}(P, Q) = \frac{H(P, Q)}{\sqrt{2}} \quad (4.13)$$



Σχήμα 4.11: Αποστάσεις Hellinger της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$



Σχήμα 4.12: Απόσταση Hellinger της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $\mathcal{Q} = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0



Σχήμα 4.13: Απόσταση Hellinger της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $\mathcal{Q} = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0

4.5.3 Απόσταση Kullback - Leibler

Αποτελώντας την μετρική που χρησιμοποιείται περισσότερο στην βιβλιογραφία[23], η απόσταση Kullback - Leibler, ή αλλιώς απόκλιση Kullback - Leibler, προέρχεται από την περιοχή της Θεωρίας Πληροφορίας (Information Theory) και δίνεται από τον τύπο:

$$K(P, \mathcal{Q}) = \sum_{i=1}^n (p_i \log_2(\frac{p_i}{q_i})) \quad (4.14)$$

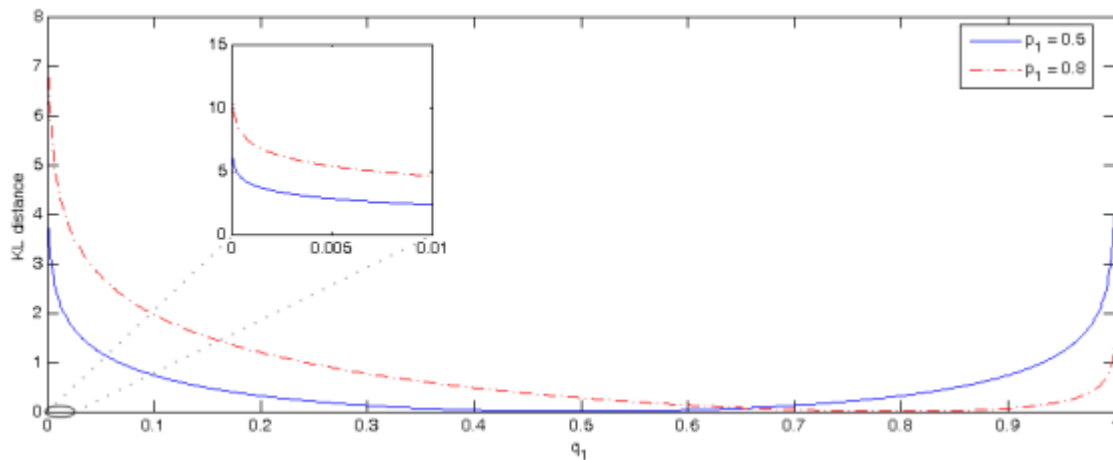
Από τον ορισμό και χρησιμοποιώντας τις λογαριθμικές ιδιότητες γίνεται εύκολα αντιληπτό ότι:

$$K(P, \mathcal{Q}) = - \sum_{i=1}^n p_i \log_2(q_i) + \sum_{i=1}^n p_i \log_2(p_i) = H(P, \mathcal{Q}) - H(P) \quad (4.15)$$

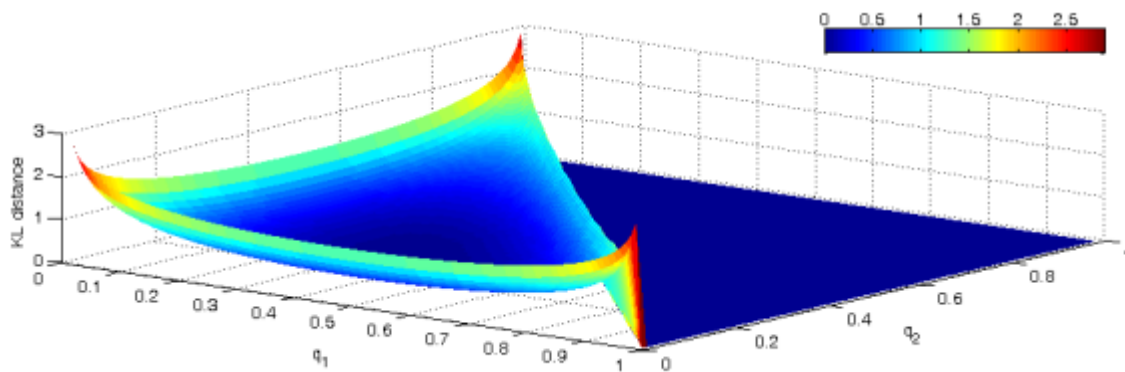
όπου $H(P, \mathcal{Q})$ είναι η Διασταυρώμενη Εντροπία (Cross-Entropy) των P και \mathcal{Q} , που εκφράζει την συνολική διαφορά μεταξύ δύο κατανομών, και $H(P)$ είναι η Εντροπία του P , που εκφράζει την ποσότητα της πληροφορίας που κατέχει το P . Η τιμή αυτής της μετρικής

κυμαίνεται από το 0 ως το άπειρο (∞). Στα σχήματα 4.12-4.14 απεικονίζουν την συμπεριφορά της Kullback - Leibler. Όπως και η Hellinger, οι διακυμάνσεις κοντά στο 0 και στο 1 μεταβάλλονται διαφορετικά από την εναπομείνουσα περιοχή.

Τα κυριότερα προβλήματα με την συγκεκριμένη μετρική σχετικά με την εφαρμογή μας είναι τρία: Πρώτον, δεν είναι συμμετρική (προκύπτει από τον ορισμό), δεύτερον, το εύρος τιμών πάει στο άπειρο και επομένως πρέπει να ορίσουμε έναν πολύπλοκο γραμμικό μετασχηματισμό για να περιορίσουμε το εύρος αυτό, και τέλος, όταν $q_i = 0$, έχουμε διαίρεση με το 0 κι έτσι πρέπει να αποκλείσουμε αυτήν την τιμή ως αποδεκτή. Μπορούμε να επιλύσουμε αυτά τα εμπόδια με την επόμενη μετρική, την J-Divergence.



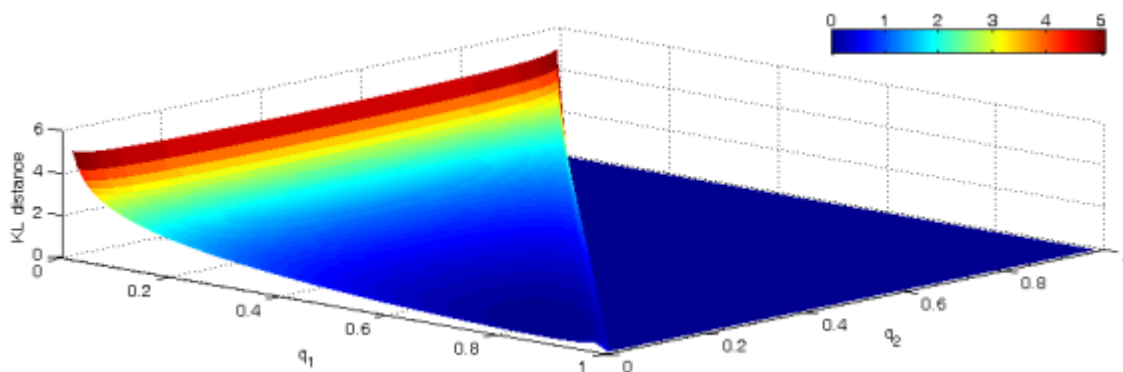
Σχήμα 4.14: Αποστάσεις Kullback-Leibler της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$



Σχήμα 4.15: Απόσταση Kullback-Leibler της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0

4.5.4 J - Απόκλιση

Αντί να μετασχηματίσουμε την Kullback - Leibler σε συμμετρική, μπορούμε να επιλέξουμε την J-Απόκλιση (J-Divergence) που ουσιαστικά αποτελεί την μέση τιμή δύο τιμών



Σχήμα 4.16: Απόσταση Kullback-Leibler της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0

των αποστάσεων Kullback - Leibler. Γενικά στη στατιστική, η απόκλιση (Divergence) είναι μια συνάρτηση που καθορίζει ένα μέτρο διαφοράς της μιας κατανομής πιθανότητας από την άλλη. Η απόκλιση είναι μια ασθενέστερη έννοια από αυτήν της απόστασης, κυρίως στο ότι η απόκλιση δεν πρέπει να είναι συμμετρική (δηλαδή, η απόκλιση από το p στο q δεν χρειάζεται να είναι ίση με την απόκλιση από το q στο p) και δεν χρειάζεται να ικανοποιεί την τριγωνική ανισότητα.[114]. Ορίζουμε την J-Απόκλιση ως:

$$J(P, Q) = \frac{K(P, Q) + K(Q, P)}{2} \tag{4.16}$$

Έτσι λύνουμε το πρόβλημα της συμμετρίας, αλλά εξακολουθεί το εύρος τιμών να φτάνει στο άπειρο. Παρόλα αυτά η νόρμα που ορίζουμε τώρα είναι πιο απλή και εκφράζεται ως:

$$J_{norm}(P, Q) = \frac{J(P, Q)}{\sqrt{J(P, Q)^2 + a}} \tag{4.17}$$

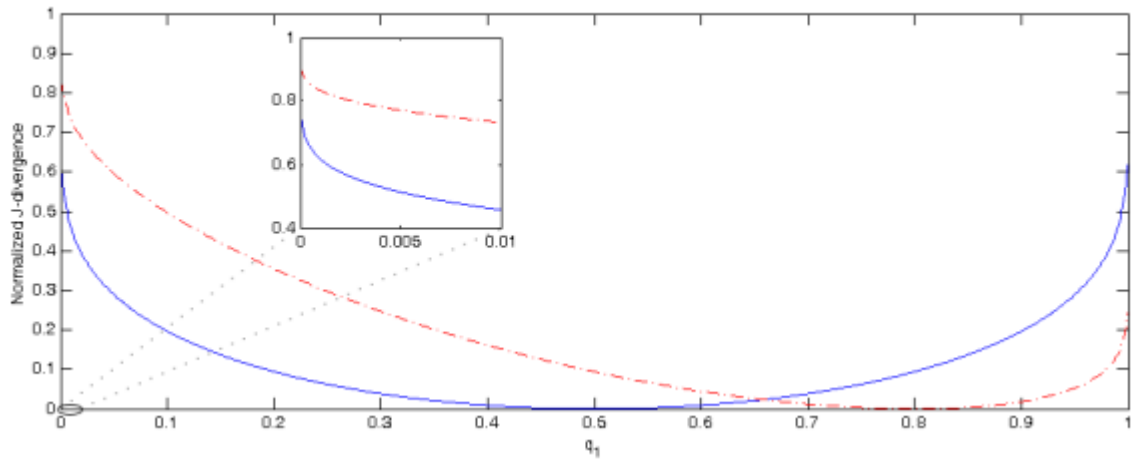
όπου το $a \geq 0$ είναι υπερπαράμετρος που ελέγχει το μέγεθος της κανονικοποίησης. Στα σχήματα 4.15-4.17 φαίνεται η κανονικοποιημένη J-Απόκλιση όταν $a = 10$. Αν την συγκρίνουμε με την Kullback - Leibler, είναι εμφανές ότι έχουμε μια πολύ ικανοποιητική προσέγγιση, η οποία όμως είναι συμμετρική και έχει τιμές από το 0 ως το 1.

Τέλος, η πιθανότητα διαίρεσης με το μηδέν δεν έχει ακόμα αποκλειστεί, αλλά πλέον ο καινούργιος ορισμός συγκλίνει στο 1 όταν το $q_i \rightarrow 0$, μπορούμε να ανανεώσουμε τον αρχικό ορισμό της νόρμας ως εξής:

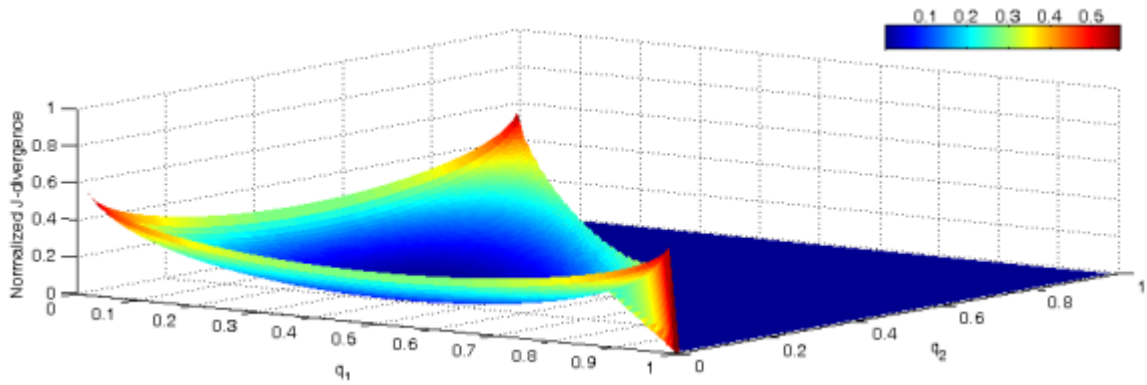
$$J_{norm}(P, Q) = \begin{cases} 1 & \exists i \in (1, \dots, n) : q_i = 0 \\ \frac{J(P, Q)}{\sqrt{J(P, Q)^2 + a}} & \text{αλλιού} \end{cases} \tag{4.18}$$

4.5.5 Απόσταση Συναρτήσεων Αθροιστικής Κατανομής

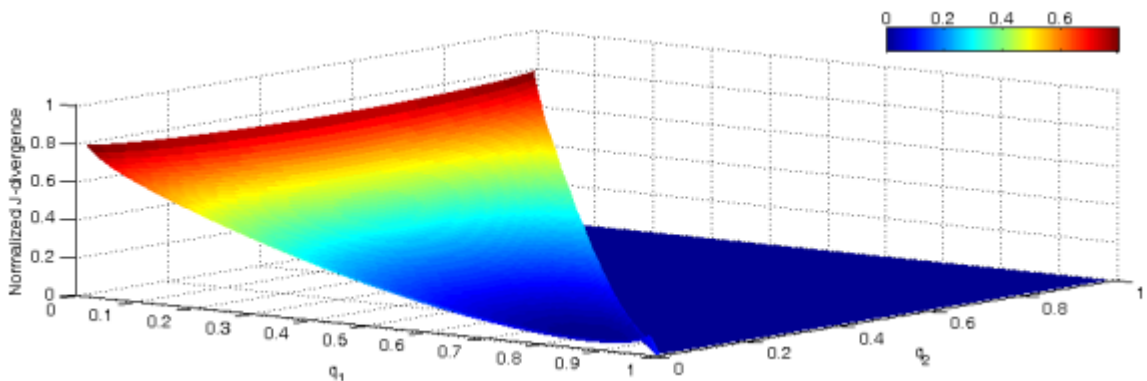
Αυτή η μετρική βασίζεται στην μέθοδο που πρωτοχρησιμοποιήθηκε από τον Kraaijeveld το 2005[115]. Στοχεύει σε διατεταγμένες κατανομές, πράγμα που σημαίνει ότι αν οι καταστάσεις των κόμβων είναι διατεταγμένες από τα αριστερά προς τα δεξιά, από τα λιγότερα



Σχήμα 4.17: Κανονικοποιημένη J -Απόκλιση με $a = 10$ της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$



Σχήμα 4.18: Κανονικοποιημένη J -Απόκλιση με $a = 10$ της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0

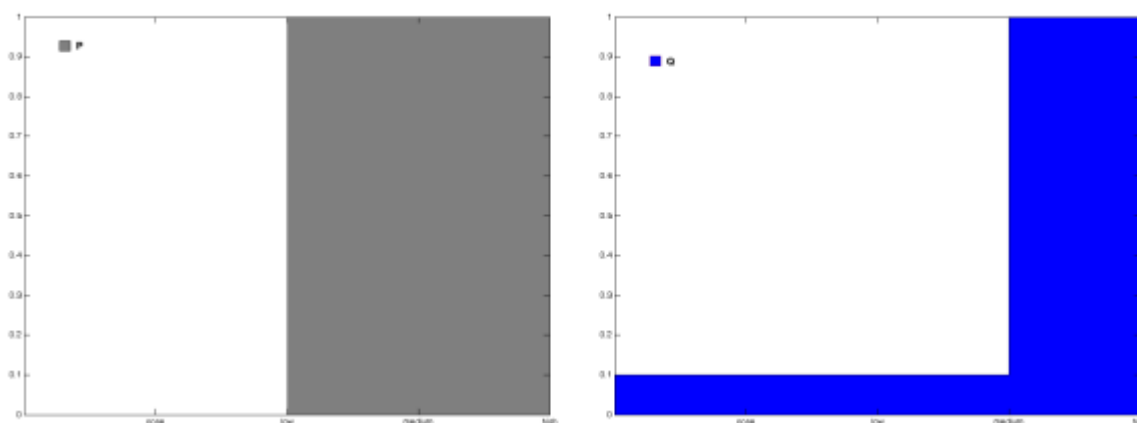


Σχήμα 4.19: Κανονικοποιημένη J -Απόκλιση με $a = 10$ της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0

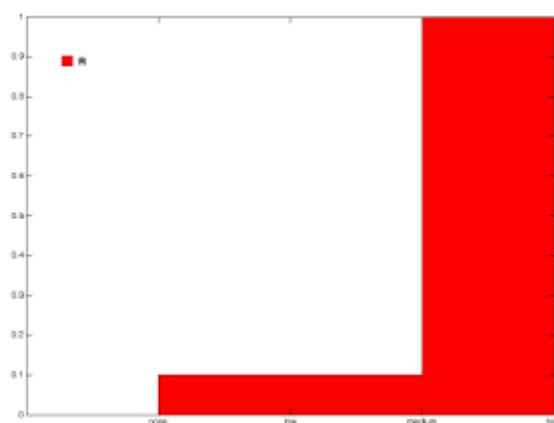
σημαντικά στα περισσότερο σημαντικά, τόσο περισσότερο η πιθανότητα θα μεταβάλλεται από την μία πλευρά της κατανομής στην άλλη, δηλαδή τόσο σημαντικότερη θα είναι και η

επίδραση της αλλαγής. Συγκρίνει τις Συναρτήσεις Αθροιστικών Κατανομών (ΣΑΚ, cumulative distribution functions (CDF)) των δύο κατανομών που θέτουμε υπό σύγκριση. Γενικεύουμε αυτή την μετρική ούτως ώστε να μπορεί να χρησιμοποιηθεί ως μετρική της απόστασης μεταξύ δύο κατανομών πιθανότητας. Ορίζεται ως:

$$C(P, Q) = \frac{1}{n-1} \sum_{i=1}^n |P(P \leq p_i) - P(Q \leq q_i)| \quad (4.19)$$



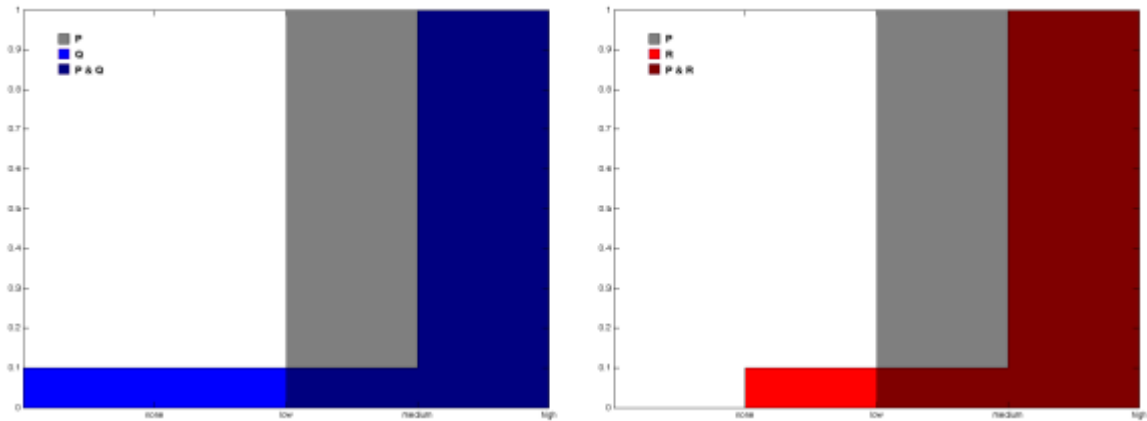
Σχήμα 4.20: ΣΑΚ των κατανομών P (αριστερά) και Q (δεξιά)



Σχήμα 4.21: ΣΑΚ της κατανομής R

Το εύρος της κυμαίνεται από το 0, όπου δεν υπάρχει καμία διαφορά, ως το 1, την μέγιστη διαφορά. Θα αποτυπώσουμε την λειτουργία αυτής της μετρικής με το εξής παράδειγμα. Δεδομένου τριών διακριτών κατανομών: $P = [0, 0, 1, 0]$, $Q = [0.1, 0, 0, 0.9]$ και $R = [0, 0.1, 0, 0.9]$, με τέσσερις καταστάσεις η κάθε μία, διατεταγμένες από το χαμηλότερο στο υψηλότερο (καθόλου, λίγο, μέτριο, πολύ). Η ΣΑΚ για κάθε μία από αυτές τις κατανομές φαίνεται στα σχήματα 4.18 και 4.19.

Αν θέλουμε να συγκρίνουμε τις P και Q καθώς και την R, μπορούμε να τις συναθροίσουμε όπως φαίνεται στο σχήμα 4.20. Διαισθητικά, η διαφορά μεταξύ P και Q είναι μεγαλύτερη



Σχήμα 4.22: Συνδυασμένες κατανομές P και Q (αριστερά) και P και R (δεξιά)

από αυτή μεταξύ P και R , επειδή στην Q η πιθανότητα της πρώτης της κατάστασης αυξάνεται συγκριτικά με την P , όσο στην R η πιθανότητα της δεύτερης κατάστασης αυξάνεται. Η πιθανότητα της P συγκεντρώνεται στην τρίτη κατάσταση, επομένως μία αύξηση στην πρώτη κατάσταση είναι πιο σημαντική από μια αύξηση στην δεύτερη κατάσταση, επειδή η δεύτερη κατάσταση τοποθετείται ακριβώς δίπλα στην τρίτη κατάσταση, σε αντίθεση με την πρώτη, που τοποθετείται μία θέση πιο μακριά από την τρίτη. Αυτό φαίνεται στο σχήμα 4.20. Θεωρούμε το μέγεθος των μη επικαλυπτόμενων μερών ως την διαφορά μεταξύ των δύο κατανομών. Με αυτόν τον τρόπο, από το σχήμα 4.20 (αριστερά) λαμβάνουμε, εφαρμόζοντας την 4.18:

$$\frac{1}{4-1}(0.1+0.1+0.9+0) = \frac{11}{30} \approx 0.37 \quad (4.20)$$

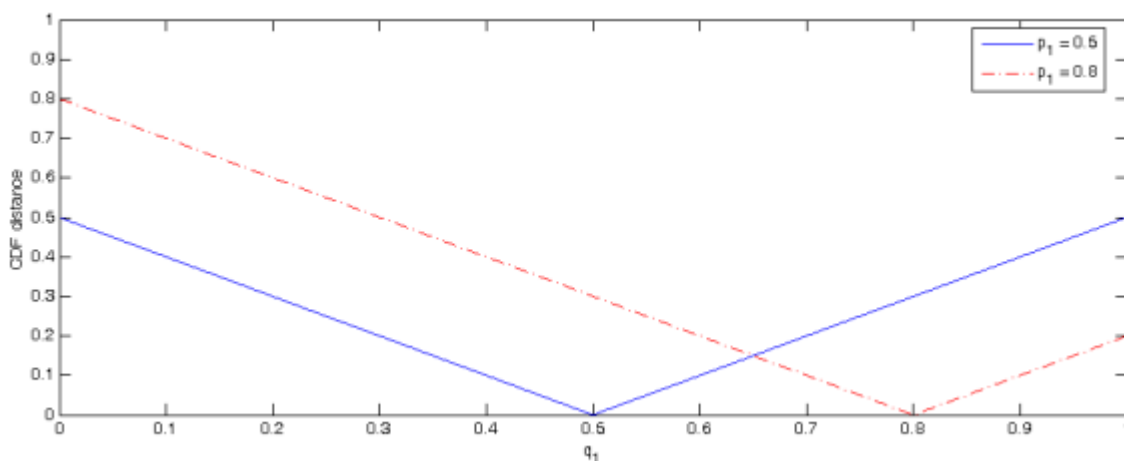
Για το σχήμα 4.20 (δεξιά) έχουμε:

$$\frac{1}{4-1}(0+0.1+0.9+0) = \frac{10}{30} \approx 0.33 \quad (4.21)$$

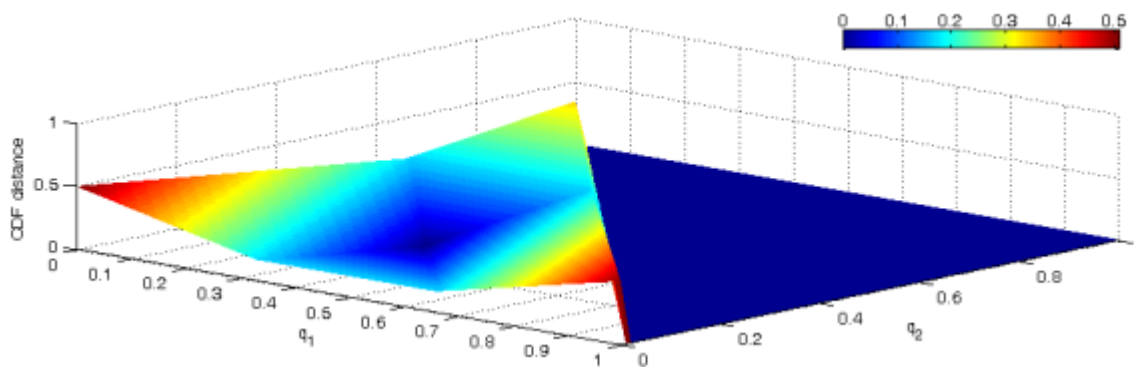
παρατηρούμε λοιπόν ότι η διαφορά μεταξύ P και Q είναι πράγματι μεγαλύτερη από αυτή μεταξύ P και R . Είναι μια μικρή διαφορά (0.04), κι αυτό διότι η διαφορά μεταξύ των δύο κατανομών είναι πραγματικά πολύ μικρή.

Τα σχήματα 4.21-4.23 δείχνουν την συμπεριφορά της απόστασης ΣΑΚ. Στα ακραία σημεία της εικόνας 4.22, παρατηρούμε πως όταν $[q_1, q_2, q_3] = [0, 1, 0]$, η διαφορά με την κατάσταση $[0.33, 0.33, 0.34]$ είναι μικρότερη από όταν το $[q_1, q_2, q_3] = [0, 0, 1]$ ή $[q_1, q_2, q_3] = [1, 0, 0]$, όπου η διαφορά είναι μέγιστη. Αυτό συμβαίνει διότι στις δύο τελευταίες περιπτώσεις, η ομοιόμορφα διασκορπισμένη κατανομή πιθανότητας $[0.33, 0.33, 0.34]$ συγκεντρώνεται ολοκληρωτικά είτε στην πιο αριστερή, είτε στην πιο δεξιά κατάσταση της Q , προκαλώντας μια μέγιστη αλλαγή, ενώ στην άλλη περίπτωση η μάζα της πιθανότητας μετακινήθηκε στην κεντρική κατάσταση, που είναι λιγότερη ακραία αλλαγή.

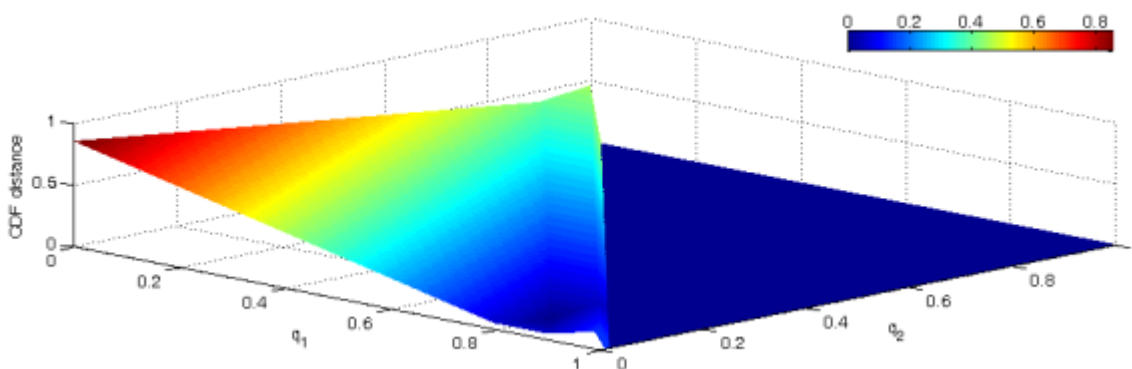
Αν παρατηρήσουμε το σχήμα 4.23, βλέπουμε ότι όταν $[q_1, q_2, q_3] = [0, 0, 1]$ η διαφορά με την κατάσταση $[0.8, 0.1, 0.1]$ είναι μέγιστη, διότι η πιθανότητα μεταβαίνει από τα αριστερά στο άκρως δεξί άκρο. Είναι επίσης εμφανές ότι η αλλαγή στην κατάσταση $[q_1, q_2, q_3] = [0, 1, 0]$ θεωρείται μεγαλύτερη από την αλλαγή σε $[q_1, q_2, q_3] = [1, 0, 0]$, που εξηγείται από



Σχήμα 4.23: Αποστάσεις ΣΑΚ της δυαδικής κατανομής $P = (p_1, p_2)$ έναντι της κατανομής $Q = (q_1, q_2)$ όπου $p_2 = 1 - p_1$ και $q_2 = 1 - q_1$



Σχήμα 4.24: Αποστάσεις ΣΑΚ της κατανομής τριών καταστάσεων $P = (0.33, 0.33, 0.34)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0



Σχήμα 4.25: Αποστάσεις ΣΑΚ της κατανομής τριών καταστάσεων $P = (0.8, 0.1, 0.1)$ έναντι της κατανομής $Q = (q_1, q_2, q_3)$ όπου $q_3 = 1 - q_1 - q_2$. Όταν $q_1 + q_2 > 1$ η κατανομή πιθανότητας γίνεται μη έγκυρη και τίθεται 0

την μεγαλύτερη μεταβολή της πιθανότητας στην πρώτη περίπτωση.

Πίνακας 4.3: Σύνοψη μετρικών αποστάσεων.

Μετρική	Εύρος	Συμμετρία	Σχόλια
Ευκλείδεια	$[0, \sqrt{2}]$	ναι	Η πιο απλή.
Hellinger	$[0, \sqrt{2}]$	ναι	Ευαίσθητη κοντά στο 0 και στο 1.
Kullback-Leibler	$[0, \infty)$	όχι	Ευαίσθητη κοντά στο 0 και στο 1.
J-Απόκλιση	$[0, \infty)$	ναι	Ευαίσθητη κοντά στο 0 και στο 1. Διαίρεση αν $q_i = 0$.
ΣΑΚ	$[0, 1]$	ναι	Διατεταγμένες κατανομές.

4.5.6 Συμπεράσματα

Ο πίνακας 4.3 συνοψίζει τις ιδιότητες των μετρικών αποστάσεων που συζητήσαμε παραπάνω. Οι Kullback - Leibler, J-Απόκλιση και Hellinger είναι πιο ευαίσθητες κοντά στις ακραίες τιμές (0 και 1), κάτι που επιθυμούμε επειδή έτσι αποτυπώνονται σχετικές διαφορές. Από τις τρεις, μόνο οι J-Απόκλιση και Hellinger είναι συμμετρικές, μια ιδιότητα που απαιτείται. Η ΣΑΚ είναι ιδανική επιλογή όταν έχουμε διατεταγμένους κόμβους, διότι υποδηλώνει την μεταβολή της πιθανότητας ανάλογα με την συνάρτηση αθροιστικής πιθανότητας των δύο κατανομών. Το μειονέκτημά της είναι ότι χειρίζεται τις τιμές κοντά στο 0 και στο 1 όπως την εναπομείνασα περιοχή. Ποιο μέτρο θα δώσει τα καλύτερα αποτελέσματα στην εφαρμογή μας είναι δύσκολο να προσδιοριστεί, ιδιαίτερα επειδή η απόδοση κάθε μετρικής μπορεί να διαφέρει για μπεϊσιανά δίκτυα με διαφορετικά χαρακτηριστικά σε κάθε εκτέλεση. Στην εφαρμογή μας εκτελέσαμε με χρήση της Ευκλείδειας απόστασης, της Hellinger, της ΣΑΚ και της J-Απόκλισης. Δεν πρόκειται να χρησιμοποιήσουμε την απόσταση Kullback-Leibler, γιατί δεν είναι συμμετρική και επομένως δεν είναι κατάλληλη για τις ανάγκες μας.

Κεφάλαιο 5

Μεθοδολογία και Υλοποίηση

Στο κεφάλαιο αυτό γίνεται περιγραφή των δεδομένων που τροφοδότησαν τα μπεισιανά δίκτυα και καταγράφονται τα αποτελέσματα της μελέτης της Π. Χατζηδάκη[77], ευρήματα που παίζουν εξέχοντα ρόλο στην μετέπειτα επιχείρηση του Συμπερασμού. Αναλύονται οι διαδικασίες προ επεξεργασίας των δεδομένων, η διακριτοποίηση (Discretization) και η διαχείριση των χαμένων τιμών (Missing Values), η παραμετροποίηση (Parameterization) του νευρωνικού, η αλγόριθμοι κατασκευής των μπεισιανών δικτύων (Structural Learning) από τα δεδομένα και οι μετρικές που χρησιμοποιούνται για την ερμηνεία και επεξήγηση των αποτελεσμάτων.

5.1 Λογισμικό SMILE

Η ανάλυση του μοντέλου πραγματοποιήθηκε χρησιμοποιώντας το λογισμικό GeNIe Academic της BayesFusion[116]. Το SMILE (Structural Modeling, Inference, and Learning Engine) είναι μια πλήρως ανεξάρτητη βιβλιοθήκη λειτουργιών που εφαρμόζει γραφικά πιθανολογικά και θεωρητικά μοντέλα, όπως μπεισιανά δίκτυα, διαγράμματα επιρροής και μοντέλα δομικών εξισώσεων. Οι επιμέρους λειτουργίες του, που ορίζονται στο SMILE Applications Programmer Interface (API), επιτρέπουν τη δημιουργία, επεξεργασία, αποθήκευση και φόρτωση γραφικών μοντέλων και τη χρήση τους για πιθανολογικούς συλλογισμούς και λήψη αποφάσεων υπό αβεβαιότητα. Το SMILE εφαρμόζεται σε C++ με τρόπο ανεξάρτητο από την πλατφόρμα του οικοδεσπότη. Παρέχει Java wraps (jSMILE), .NET (SMILE.NET), Python (PySMILE) και R (rSMILE) για χρήστες που θέλουν να χρησιμοποιήσουν το SMILE με γλώσσες διαφορετικές από την C++. Το SMILE είναι εξοπλισμένο με εξωτερικό κέλυφος, περιβάλλον προγραμματιστή για τη δημιουργία μοντέλων γραφικών αποφάσεων, γνωστό ως GeNIe. Το GeNIe εξαρτάται από την πλατφόρμα και λειτουργεί σε υπολογιστές Windows. Το SMILE μπορεί να ενσωματωθεί σε προγράμματα που χρησιμοποιούν γραφικά πιθανολογικά μοντέλα ως μηχανές συλλογισμού τους. Τέτοια προγράμματα μπορούν να διανεμηθούν στους τελικούς χρήστες ή να τοποθετηθούν σε διακομιστές για χρήση στο cloud. Τα μοντέλα που αναπτύχθηκαν στο SMILE μπορούν να εξοπλιστούν με μια διεπαφή χρήστη που ταιριάζει περισσότερο στον χρήστη της εφαρμογής που προκύπτει. Τα GeNIe, QGeNIe και SMILE έχουν αρχικά αναπτυχθεί για να είναι σημαντικά εργαλεία διδασκαλίας και έρευνας σε ακαδημαϊκά περιβάλλοντα και έχουν χρησιμοποιηθεί σε εκατοντάδες πανεπιστήμια παγκοσμίως. Οι περισσότερες έρευνες που πραγματοποιήθηκαν στο Decision Systems Laboratory, Univer-

sity of Pittsburgh, βρήκαν το δρόμο τους στο GeNIe και το SMILE. Λόγω της ευελιξίας και της αξιοπιστίας τους, τα GeNIe, QGeNIe και SMILE έγιναν πολύ δημοφιλή και έγιναν de facto πρότυπα στον ακαδημαϊκό χώρο, ενώ αγκαλιάστηκαν από πολλούς κυβερνητικούς και εμπορικούς χρήστες. Το ισχυρότερο στοιχείο του SMILE, που το διακρίνει από ένα μεγάλο αριθμό άλλων εργαλείων γραφικής μοντελοποίησης, είναι η ευκολία χρήσης του από την πλευρά του προγραμματιστή (προσφέρει σύγχρονο API που βασίζεται σε αντικείμενα), διαθεσιμότητα για πολλαπλές πλατφόρμες, η αξιοπιστία του (έχει δοκιμαστεί σε μεγάλο βαθμό σε πρακτικές έρευνες και εμπορικές εφαρμογές από το 1998) και ταχύτητα (έχει δείξει εξαιρετικές επιδόσεις σε αγώνες ταχύτητας UAI). Η ταχύτητα είναι ιδιαίτερα σημαντική, καθώς οι περισσότεροι υπολογισμοί σε πιθανολογικά γραφικά μοντέλα είναι εκθετικής τάξης[116].

5.2 Σύνολο Δεδομένων

Τα δεδομένα προέρχονται από μια συγχρονική μελέτη επιπολασμού παραγόντων που σχετίζονται με αυξημένο βάρος παιδιών ηλικίας 6 έως 11 ετών που προσέρχονται σε Ιατρείο Παχυσαρκίας. Συγκεκριμένα, η έρευνα αυτή διεξήχθη στο Ιατρείο Παχυσαρκίας της Μονάδας Αναπτυξιακής και Συμπεριφορικής Παιδιατρικής του Νοσοκομείου Παίδων «Αγία Σοφία» κατά το χρονικό διάστημα 01/11/2017 έως 31/12/2018. Στη μελέτη συμπεριλήφθηκαν διαδοχικά όλα τα παιδιά που επισκέφτηκαν το Ιατρείο για πρώτη φορά το χρονικό διάστημα 01/11/2017 έως 31/12/2018, πληρούσαν τα κριτήρια εισαγωγής και οι γονείς τους έδωσαν τη συγκατάθεσή τους (n=48).

Βασικά κριτήρια εισαγωγής ήταν τα παιδιά να είναι ηλικίας από 6 έως 11 ετών, το στάδιο σεξουαλικής ανάπτυξης (ή αλλιώς ενήθωσης κατά Tanner) να είναι από I έως III[117], και το BMI z-score να είναι μεγαλύτερο των 2 SD για το αντίστοιχο φύλο και την ηλικία, σύμφωνα με τον ορισμό της παχυσαρκίας στα παιδιά ηλικίας 5-19 ετών του Παγκόσμιου Οργανισμού Υγείας[118]. Από την άλλη, κριτήρια αποκλεισμού από την έρευνα αποτέλεσαν τα δευτεροπαθή αίτια παχυσαρκίας όπως το σύνδρομο Cushing, ο υποθυρεοειδισμός, γενετικά σύνδρομα σχετιζόμενα με αυξημένο σωματικό βάρος (π.χ. σύνδρομο Down, σύνδρομο Prader Willi) ή γενετικά αίτια παχυσαρκίας[77]. Οι κατηγορίες που εστιάζουμε είναι η Γονεϊκή Τυπολογία και οι Γονεϊκές Πρακτικές Σίτισης.

5.2.1 Γονεϊκή Τυπολογία Μητέρων

Στη μελέτη χρησιμοποιήθηκε το ερωτηματολόγιο γονεϊκής τυπολογίας των Robinson et al (Parenting Styles and Dimensions Questionnaire - PSDQ[119]) σταθμισμένο για τις Ελληνίδες μητέρες από τους Αντωνοπούλου και Τσίτσα το 2011[120]. Ενώ το αρχικό ερωτηματολόγιο των Robinson et al περιλαμβάνει τρεις γονεϊκούς τύπους, την υποστηρικτική, την επιτρεπτική και την αυταρχική μητέρα, η στάθμισή του στις Ελληνίδες μητέρες ανέδειξε και έναν τέταρτο τύπο, την αυστηρή μητέρα, η οποία εκπορεύεται από την ελληνική πραγματικότητα και έχει χαρακτηριστικά τόσο της υποστηρικτικής (νοιάζεται για τις επιθυμίες και τα συναισθήματα του παιδιού της) αλλά και της αυταρχικής μητέρας (ασκεί αυστηρό έλεγχο και θέτει ξεκάθαρα όρια στη συμπεριφορά του παιδιού της)[120]. Συγκεντρωτικά οι γονεϊκοί τύποι που αναλύουμε είναι:

1. Υποστηρικτική (Supportive),
2. Επιτρεπτική (Permissive),
3. Αυταρχική (Authoritarian),
4. Αυστηρή (Authoritative)

5.2.2 Γονεϊκές Πρακτικές Σίτισης

Οι μητέρες των παιδιών συμπλήρωσαν το ερωτηματολόγιο γονεϊκών πρακτικών σίτισης (Comprehensive Feeding Practices Questionnaire - CFPQ) των Musher-Eizenman και Holub. Το ερωτηματολόγιο περιλαμβάνει 49 ερωτήσεις που αξιολογούν 12 πρακτικές σίτισης[121]. Επειδή όμως το συγκεκριμένο ερωτηματολόγιο δε βρισκόταν στην τελική του μορφή κατά τη διάρκεια της μελέτης χρησιμοποιήθηκαν μόνο 29 από τις τελικές 42 ερωτήσεις της ελληνικής εκδοχής του ερωτηματολογίου[122]. Οι γονεϊκές πρακτικές σίτισης που μελετήθηκαν είναι:

1. Παρακολούθηση των ανθυγιεινών τροφίμων που καταναλώνει το παιδί (Monitoring),
2. Συμμετοχή του παιδιού στην προετοιμασία των γευμάτων (Involvement),
3. Λειτουργία των γονέων ως πρότυπα υγιεινής διατροφής (Modeling),
4. Περιορισμός της τροφής για αποφυγή της αύξησης βάρους του παιδιού (Restriction),
5. Το να επιτρέπουν στο παιδί οι γονείς να έχει ένα βαθμό ελέγχου της διατροφικής του πρόσληψης (Child Control) και
6. Η πίεση για φαγητό (Pressure to Eat)[122].

5.2.3 Άλλες Κατηγορίες Δεδομένων

Εκτός των προαναφερθέντων κατηγοριών, στα δεδομένα περιλαμβάνονταν πληθώρα μετρήσεων από άλλες κατηγορίες όπως:

1. Η αξιολόγηση περιγεννητικών παραγόντων: προγεννητικών (π.χ. διαβήτη κύησης, κάπνισμα) και μεταγεννητικών (π.χ. μητρικός θηλασμός)
2. Παράγοντες του τρόπου ζωής των παιδιών και των γονέων (π.χ. φυσική δραστηριότητα, ύπνος, ενασχόληση με οθόνες)
3. Καταγραφή κοινωνικοδημογραφικών και οικονομικών δεδομένων των μητέρων (π.χ. εκπαίδευση, εργασία, εθνικότητα)
4. Οι κλινικές εξετάσεις των παιδιών και των γονέων (ανθρωπομετρικά χαρακτηριστικά των παιδιών και των γονέων)

Όλοι οι παράγοντες καταγράφηκαν από το ιστορικό των παιδιών, από μεμονωμένες ή εβδομαδιαίας συχνότητας συνεντεύξεις με το παιδί και τους γονείς ύστερα από συμπλήρωση ειδικών φορμών συλλογής δεδομένων και μέσω κλινικών εξετάσεων[77].

5.2.4 Αποτελέσματα Στατιστικής Ανάλυσης

Στην μελέτη συμμετείχαν 48 παιδιά, 28 αγόρια (58,3%) και 20 κορίτσια (41,7%) με μέσο όρο ηλικίας τα $8,9 \pm 1,3$ έτη. Το επίπεδο σημαντικότητας ορίστηκε ως $p < 0.05$ [77].

1. Τα αγόρια του δείγματος σε σύγκριση με τα κορίτσια ήταν βαρύτερα, με μέσο βάρος $55,7 \pm 13,3$ kg, έναντι $48,1 \pm 10,9$ kg ($p = 0,028$) και είχαν μεγαλύτερο BMI, $28,4 \pm 4,7$ kg/m² έναντι $25,3 \pm 3,8$ kg/m² ($p = 0,016$).
2. Τα αγόρια είχαν πιο συχνά (25%) και τους δυο γονείς παχύσαρκους έναντι 10% των κοριτσιών, τα κορίτσια είχαν πιο συχνά (70%) τον ένα γονέα παχύσαρκο έναντι 32,1% των αγοριών, ενώ κανέναν γονιό παχύσαρκο είχε το 42,9% των αγοριών και το 20% των κοριτσιών ($p = 0,035$).
3. Οι μητέρες των παχύσαρκων κοριτσιών είναι πιο επιτρεπτικές σε σχέση με τις μητέρες των παχύσαρκων αγοριών ($p = 0.03$).
4. Οι μητέρες των παχύσαρκων κοριτσιών είναι πιο πιεστικές με το φαγητό σε σχέση με τις μητέρες των παχύσαρκων αγοριών ($p = 0,029$)[77].

Όπως θα δούμε και στην συνέχεια, η "ανατομία" των δεδομένων παίζει σημαντικότατο ρόλο στα μετέπειτα αποτελέσματα.

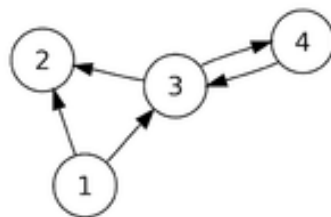
5.3 Προεπεξεργασία Δεδομένων

5.3.1 Ελλιπείς Τιμές

Πρώτο βήμα αποτέλεσε η διαχείριση των ελλিপών τιμών των δεδομένων. Λόγω των ιδιαιτεροτήτων κάποιων χαρακτηριστικών, διάφορες τιμές απουσίαζαν διασκορπισμένες στο σύνολο των δεδομένων. Εννέα χαρακτηριστικά είχαν τουλάχιστον ένα ελλιπές στοιχείο πληροφορίας μη κρίσιμο, δηλαδή η απουσία τους προσέδιδε πληροφορία στο σύστημα κι άρα συμπληρώθηκαν με μηδενική τιμή. Δέκα χαρακτηριστικά είχαν τουλάχιστον ένα ελλιπές στοιχείο πληροφορίας κρίσιμο, και επομένως συμπληρώθηκαν με την μέση τιμή της συνολικής στήλης για να μην χαθεί η πληροφορία.

5.3.2 Διακριτοποίηση

Πολύ βασική διαδικασία, αν όχι η σημαντικότερη, αποτέλεσε η Διακριτοποίηση (Discretization) πολλών χαρακτηριστικών. Οι περισσότεροι αλγόριθμοι για κατασκευή μπεισιανών δικτύων και διαγραμμάτων επιρροής έχουν σχεδιαστεί για διακριτές μεταβλητές. Για να επωφεληθούμε από αυτούς τους αλγόριθμους, τα περισσότερα δίκτυα περιλαμβάνουν διακριτές μεταβλητές ή εννοιολογικά συνεχής μεταβλητές που έχουν διακριθεί συλλογιστικά. Ενώ η διάκριση μεταξύ διακριτών και συνεχών μεταβλητών είναι σαφώς ορισμένη, η διάκριση μεταξύ διακριτών και συνεχών ποσοτήτων είναι μάλλον ασαφής. Πολλές ποσότητες μπορούν να αναπαρασταθούν ως διακριτές και συνεχείς. Οι διακριτές μεταβλητές είναι συνήθως βολικές προσεγγίσεις των πραγματικών ποσοτήτων, επαρκείς για το σκοπό της συλλογιστικής. Η εμπειρία στην αναλυτική μοντελοποίηση αποφάσεων έχει διδάξει ότι η αναπαράσταση των



Σχήμα 5.1: Κατευθυνόμενο Ακυκλικό Γράφημα (ΚΑΓ)

συνεχών μεταβλητών με διακριτές προσεγγίσεις τριών έως πέντε σημείων, λειτουργούν εξαιρετικά καλά στις περισσότερες περιπτώσεις (πχ η μεταβλητή "Θερμοκρασία Σώματος" μπορεί να θεωρηθεί συνεχής, αλλά μπορεί και να διακριτοποιηθεί σε πολύ χαμηλή, χαμηλή, μέτρια, υψηλή, πολύ υψηλή κλπ).

Η διαδικασία που ακολουθήθηκε για την διακριτοποίηση των μεταβλητών, που ήταν απαραίτητο ούτως ώστε να μην χαθεί πληροφορία, ήταν η εξής: πρώτα καταμετρούμε την πληθικότητα των διαφορετικών τιμών ανά χαρακτηριστικό και την διαφορά των ακραίων τιμών. Ύστερα ορίζουμε τον αριθμό των διαστημάτων διακριτοποίησης *discretization intervals* ανάλογα τα παραπάνω ποσά ώστε να δημιουργούμε διαστήματα που να ακολουθούν κανονική κατανομή ως προς το πλήθος των συνεχών τιμών που καταλήγουν να περιέχουν. Με αυτόν τον τρόπο διαφυλάξαμε ότι μετά το πέρας της διακριτοποίησης θα παραμείνει ως έχει η ποικιλία των αρχικών συνεχών μεταβλητών. Τέλος χρησιμοποιήσαμε την Ιεραρχική (Hierarchical) μέθοδο, δηλαδή πραγματοποιούμε μη επιβλεπόμενη ομαδοποίηση (*clustering*) ανάλογη της πληθικότητας των συνεχών τιμών, καταλήγοντας έτσι η τελική μορφή των κατηγορικών μεταβλητών.

Διακριτοποίηση πραγματοποιήθηκε όπου ήταν απολύτως αναγκαίο, και πιο συγκεκριμένα στις τιμές μέσων όρων με πολλά δεκαδικά ψηφία, στα Ύψη, Βάρη, Ηλικίες και ΔΜΣ παιδιών και γονέων, καθώς και στην συστολική και διαστολική αρτηριακή πίεση. Για να υπάρχει ομοιομορφία μεταξύ των χαρακτηριστικών, και επομένως, καλύτερα και πιο ακριβή αποτελέσματα, τα χαρακτηριστικά: Ηλικίες (παιδιών-γονέων), Ύψη (παιδιών-γονέων) και Βάρη (παιδιών-γονέων) αναπαραστάθηκαν με διακριτές προσεγγίσεις έξι (6) σημείων, ενώ ο ΔΜΣ (παιδιών-γονέων), η συστολική και διαστολική αρτηριακή πίεση, οι γονεϊκοί τύποι της μητέρας καθώς και οι γονεϊκές πρακτικές σίτισης με διακριτές προσεγγίσεις τεσσάρων (4) σημείων.

5.4 Κατασκευή Μπεϊσιανού Δικτύου

Ένα από τα σημαντικότερα στοιχεία της πιθανοτικής μοντελοποίησης είναι η ικανότητα των κατευθυνόμενων πιθανολογικά γράφων να αναπαριστούν την αιτιώδη δομή του τομέα εφαρμογής που μας ενδιαφέρει. Η ίδια η δομή που προκύπτει ως αποτέλεσμα είναι πολύτιμη και αποτελεί σημαντική πηγή διορατικότητας. Ένα σημαντικό στοιχείο αυτής της ανάλυσης είναι η Ισχύς Επιρροής (Strength of Influence) και των Διαδρομών (Pathways) μέσα στον γράφο. Δεδομένης μιας συλλογής δειγμάτων, η διαδικασία της "μάθησης μπεϊσιανού δικτύου" καταλήγει πάντα σε ένα ΚΑΓ που αναπαριστά τις συσχετίσεις μεταξύ των μεταβλητών.

Η δομή του δικτύου είναι πολύτιμη και αποτελεί σημαντικότερη πηγή γνώσης. Τα κατασκευασμένα μοντέλα μπορούν να αναλυθούν ενδελεχώς ως προς την δομή τους. Προκειμένου να μάθουμε για την δομή του δικτύου, πρέπει να υλοποιήσουμε διάφορους Αλγόριθμους Μάθησης (Learning Algorithms) με τις αντίστοιχες παραμέτρους τους και να καταλήξουμε σε εκείνον με το καλύτερο αποτέλεσμα. Υπάρχουν τρία σημαντικά εμπόδια στην μάθηση που δοκιμάζονται στην αρχή της εκτέλεσης καθενός από τους αλγόριθμους. Καθένας από τους αλγόριθμους είναι ικανός για τη μάθηση όταν όλες οι μεταβλητές είναι κατηγορηματικές. Ακόμα, κανένας από τους αλγόριθμους εκμάθησης δομής δεν μπορεί να μάθει τη δομή ενός μοντέλου όταν λείπουν τιμές στις εγγραφές. Τέλος, κανένας από τους αλγόριθμους δεν επιτρέπει την εκμάθηση με σταθερές τιμές, δηλαδή, μεταβλητές (στήλες στα δεδομένα) που περιέχουν την ίδια τιμή σε όλες τις εγγραφές. Με την προεπεξεργασία που υλοποιήσαμε και περιγράψαμε παραπάνω καλύψαμε όλα τα πιθανά αυτά εμπόδια. Οι παρακάτω αλγόριθμοι αποσκοπούν στην Δομική Μάθηση (Structural Learning) του μπειϊσιανού δικτύου από τα δεδομένα.

5.4.1 Αλγόριθμοι Μπειϊσιανής Μάθησης

Μπειϊσιανή Αναζήτηση

Ο αλγόριθμος δομικής μάθησης της Μπειϊσιανής Αναζήτησης Bayesian Search είναι ένας από τους πρώτους και πιο διαδεδομένους αλγόριθμους που χρησιμοποιείται σε πληθώρα εφαρμογών μέχρι και σήμερα. Εισήχθη από τους Cooper και Herkovitz[123] και βελτιώθηκε αργότερα από τον Heckerman[124]. Ακολουθεί ουσιαστικά μια διαδικασία ανάβασης σε λόφο (Hill Climbing) με τυχαία επανεκκίνηση, που καθοδηγείται από μια ευριστική βαθμολογία, η οποία στο GeNIe είναι η συνάρτηση λογαριθμικής πιθανότητας (log-likelihood function). Ο αλγόριθμος παράγει ένα ΚΑΓ (βλ. Σχήμα 5.1) που επιτυγχάνει την υψηλότερη βαθμολογία. Η βαθμολογία αυτή είναι ανάλογη με την πιθανότητα των δεδομένων δεδομένης της δομής, η οποία, υποθέτοντας ότι εκχωρούμε την ίδια εκ των προτέρων πιθανότητα σε οποιοδήποτε χαρακτηριστικό, είναι ανάλογη στην πιθανότητα της δομής δεδομένου των δεδομένων.

P.C.

Ο αλγόριθμος δομικής μάθησης PC είναι ένας από τους πρώτους και πιο δημοφιλείς αλγόριθμους, που εισήχθησαν από τον Spirtes et al.[125]. Χρησιμοποιεί ανεξαρτησίες που παρατηρούνται στα δεδομένα (που δημιουργήθηκαν από μέσες τιμές κλασικών δοκιμών ανεξαρτησίας) για να συμπεράνουμε τη δομή που τις έχει δημιουργήσει. Το θέμα μιας εργασίας από τους Druzdzel και Glymour[126], οι οποίοι κατέληξαν στο συμπέρασμα ότι η μόνη άμεση αιτία χαμηλής διατήρησης φοιτητών στα πανεπιστήμια των ΗΠΑ είναι η ποιότητα των εισερχόμενων φοιτητών, είναι ένα από τα επιτυχημένα παραδείγματα συσχετιστικής ανακάλυψης και το συμπέρασμά της επαληθεύτηκε εμπειρικά αργότερα σε ένα πραγματικό πείραμα από το Πανεπιστήμιο Carnegie Mellon[126].

Greedy Thick Thinning

Ο αλγόριθμος δομικής μάθησης Greedy Thick Thinning (GTT) βασίζεται στην προσεγγιστική Μπεϊσιανή Αναζήτηση και έχει περιγραφεί από την Cheng et al.[127]. Ο GTT ξεκινά με ένα κενό γράφημα και προσθέτει επανειλημμένα ακμές (χωρίς τη δημιουργία κύκλου) που αυξάνει στο μέγιστο την οριακή πιθανότητα $P(D|S)$ έως ότου να μην υπάρχει καμία προσθήκη τόξου με θετική προσαύξηση (αυξητική φάση). Στη συνέχεια, αφαιρεί επανειλημμένα τις ακμές έως ότου καμία διαγραφή να έχει ως αποτέλεσμα θετική αύξηση του $P(D|S)$ (μειωτική φάση). Είναι ένας προσεγγιστικός αλλά πολύ γρήγορος αλγόριθμος που αποδίδει αρκετά καλά σε δομές μεγάλου μεγέθους που δεν μπορούν να επιλυθούν με ακρίβεια.

Tree Augmented Naive Bayes

Ο αλγόριθμος δομικής μάθησης Tree Augmented Naive Bayes (TAN) είναι μια μέθοδος βασισμένη στην προσεγγιστική Μπεϊσιανή Αναζήτηση που περιγράφεται και αξιολογήθηκε διεξοδικά από τους Friedman et al.[128]. Ο αλγόριθμος ξεκινά με την δομή του Naive Bayes, στην οποία η μεταβλητή κλάσης είναι ο μόνος γονέας όλων των εναπομείναντων χαρακτηριστικών και προσθέτει συνδέσεις μεταξύ των χαρακτηριστικών λαμβάνοντας υπόψη την πιθανή εξάρτηση μεταξύ τους, δεσμευμένης της μεταβλητής κλάσης. Ο αλγόριθμος επιβάλλει το όριο μόνο ενός επιπλέον γονέα για κάθε μεταβλητή χαρακτηριστικών (επιπλέον της μεταβλητής κλάσης, η οποία είναι γονέας κάθε μεταβλητής χαρακτηριστικών). Αξίζει να σημειωθεί ότι η δομή του Naive Bayes υποθέτει ότι τα χαρακτηριστικά είναι ανεξάρτητα μεταξύ τους δεσμευμένης της μεταβλητής κλάσης, η οποία οδηγεί σε ανακρίβειες όταν δεν είναι ανεξάρτητες. Ο αλγόριθμος TAN είναι απλός και λειτουργεί πιο αξιόπιστα από τον Naive Bayes, αλλά ενδείκνυται σε λίγες περιστάσεις.

5.4.2 Παραμετροποίηση

Οι παράμετροι του αλγορίθμου της Μπεϊσιανής Αναζήτησης είναι οι εξής. Αρχικά επιλέγουμε τον μέγιστο αριθμό γονέων ανά χαρακτηριστικό (Max Parent Count). Επειδή το μέγεθος των πινάκων δεσμευμένης πιθανότητας ενός κόμβου αυξάνεται εκθετικά στον αριθμό των γονέων του κόμβου, είναι καλή ιδέα να θέσουμε ένα όριο στον αριθμό αυτό, ώστε η κατασκευή του δικτύου να μην εξαντλήσει όλη την διαθέσιμη μνήμη του υπολογιστή. Μετά από πολλές δοκιμές, καταλήξαμε σε 4 (με άνω όριο το 8).

Οι επαναλήψεις (Iterations) ορίζουν τον αριθμό επανεκκινήσεων του αλγορίθμου. Γενικά, ο αλγόριθμος πραγματοποιεί αναζήτηση μέσω ενός υπερεκθετικού χώρου αναζήτησης και ο στόχος του μπορεί να είναι σαν την αναζήτηση βελόνας στα άχυρα. Οι επανεκκινήσεις επιτρέπουν την ανίχνευση περισσότερων περιοχών του χώρου αναζήτησης και αυξάνουν την πιθανότητα εύρεσης μιας δομής που θα ταιριάζει στα δεδομένα καλύτερα. Ο υπολογιστικός χρόνος είναι πρακτικά γραμμικός στον αριθμό των επαναλήψεων.

Η Πιθανότητα Σύνδεσης (Link Probability) είναι μια παράμετρος που χρησιμοποιείται κατά τη δημιουργία ενός τυχαίου δικτύου εκκίνησης στην έναρξη κάθε μιας από τις επαναλήψεις. Επηρεάζει ισχυρά την συνδεσιμότητα του αρχικού δικτύου, και την θέτουμε στο 0,5, ώστε να δώσουμε ίση αρχική πιθανότητα μεταξύ των κόμβων να έχουν ή όχι συσχέτιση.

Η εκ των προτέρων Πιθανότητα Σύνδεσης (Prior Link Probability) επηρεάζει το BDeu score, προσφέροντας μια εκ των προτέρων πιθανότητα σε όλες τις ακμές. Υλοποιεί τον τύπο με τον ακόλουθο τρόπο :

$$\log PS = \log ML + |p| * \log(pll) + (|n| - |p| - 1) * \log(1 - pll)$$

όπου :

PS = Posterior Score,

ML = Marginal likelihood (BDeu),

p = parents,

n = nodes

Το BDeu score στοχεύει στη μεγιστοποίηση της εκ των υστέρων πιθανότητας των δεδομένων του ΚΑΓ, ενώ ταυτόχρονα υποθέτει μια ομοιόμορφη κατανομή των προηγούμενων πιθανοτήτων των δεδομένων αυτών έναντι όλων των πιθανών ΚΑΓ[116].

5.5 Μπείσιανός Συμπερασμός

Υλοποιώντας όλα τα παραπάνω προκύπτει το τελικό ΚΑΓ, το οποίο και θα αναλύσουμε προκειμένου να εκμαιεύσουμε όποια πληροφορία χρειαζόμαστε. Το βασικότερο που μας ενδιαφέρει στην παρούσα φάση είναι να καταγράψουμε την αλληλεπίδραση των διαφόρων χαρακτηριστικών και την επιρροή τους στα κύρια χαρακτηριστικά που μας ενδιαφέρουν, κυρίως τα: Βάρος, ΔΜΣ και ΔΜΣ ζ-βαθμολογία, αλλά και γενικότερα τις επιδράσεις όλων των χαρακτηριστικών. Η βασικότερη μετρική ώστε να ποσοτικοποιήσουμε μαθηματικά την αλληλεπίδραση αυτή είναι η Ισχύς της Επιρροής (Strength of Influence).

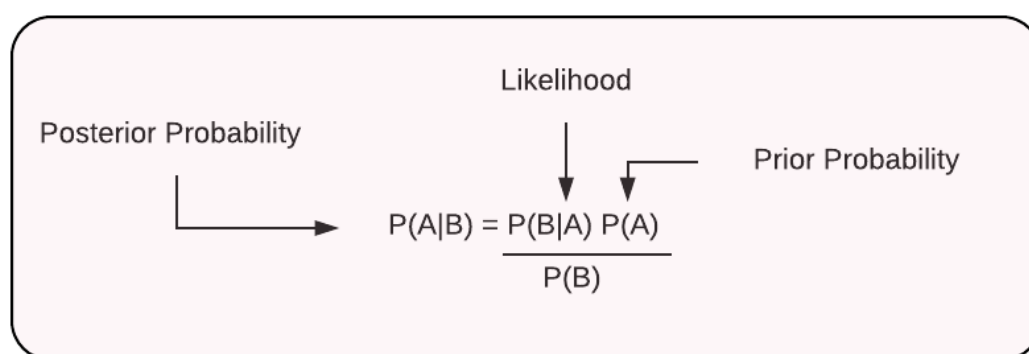
5.5.1 Ισχύς Επιρροής

Μια από τις σημαντικότερες λειτουργίες των μπείσιανών δικτύων είναι η οπτική αναπαράσταση της αλληλεπίδρασης των μεταβλητών και χαρακτηριστικών του δικτύου. Κάθε ακμή έχει διαφορετικό πάχος, ανάλογα με την ισχύ της επιρροής μεταξύ των κόμβων που συνδέονται. Η ισχύς της επιρροής υπολογίζεται από τις Κατανομές Δεσμευμένης Πιθανότητας, ΚΔΠ (conditional probability distribution, CPT) του παιδιού κόμβου και ουσιαστικά εκφράζει την απόσταση μεταξύ των διαφόρων ΚΔΠ του θυγατρικού κόμβου δεσμευμένο στις καταστάσεις του γονικού κόμβου.

Το πάχος των ακμών μπορεί να βασιστεί σε μία από τις τρεις μετρικές: Μέση, Μέγιστη και Σταθμισμένη (Average, Maximum, Weighted). Το Μέγιστο χρησιμοποιεί την μεγαλύτερη απόσταση μεταξύ των κατανομών, το Μέσο παίρνει τον μέσο όρο όλων των αποστάσεων ενώ το Σταθμισμένο ζυγίζει τις αποστάσεις από την περιθωριακή (Marginal) πιθανότητα του γονικού κόμβου. Σε όλα τα παραπάνω μπορούμε να υπολογίσουμε την κανονικοποιημένη τιμή. Σε κανονικοποιημένη λειτουργία, η πιο παχιά ακμή δίνεται σε εκείνη την ακμή που έχει την υψηλότερη Ισχύ Επιρροής. Υπολογίζονται τα πάχη όλων των άλλων ακμών ανάλογα προς την παχύτερη ακμή. Στη μη κανονικοποιημένη λειτουργία, το πάχος υπολογίζεται με βάση την απόλυτη τιμή της απόστασης. Οι ακμές έχουν διαφορετικό πάχος, ανάλογα με τη Ισχύ της

Επιρροής μεταξύ των κόμβων που συνδέουν. Η ισχύς της επιρροής υπολογίζεται από την ΚΔΠ του θυγατρικού κόμβου και ουσιαστικά εκφράζει την απόσταση μεταξύ διαφόρων κατανομών δεσμευμένης πιθανότητας επί του κόμβου-παιδιού υπό τις καταστάσεις του γονικού κόμβου. Αυτό εξυπηρετεί το σκοπό της οπτικής ανάλυσης και επαλήθευσης των μοντέλων κατά τη διάρκειά της κατασκευής[116].

5.5.2 Ανάλυση Ευαισθησίας Δικτύου



Εικόνα 5.1: Εκ των Υστέρων Πιθανότητα

Η Ανάλυση Ευαισθησίας Δικτύου (Sensitivity Analysis) είναι μια τεχνική του Castillo et al.[129] που μπορεί να βοηθήσει στην επικύρωση των παραμέτρων ενός μπεϊσιανού δικτύου. Αυτό γίνεται με τη διερεύνηση του αποτελέσματος μικρών αλλαγών στις αριθμητικές παραμέτρους (π.χ. στις πιθανότητες) ως προς τις παραμέτρους εξόδου (π.χ. εκ των υστέρων πιθανότητες (posterior probabilities)). Οι εξαιρετικά ευαίσθητες παράμετροι επηρεάζουν προφανώς το αποτέλεσμα πιο σημαντικά. Δίνεται ένα σύνολο κόμβων-στόχων κι ο αλγόριθμος υπολογίζει αποτελεσματικά ένα πλήρες σύνολο παραγώγων των εκ των υστέρων πιθανοτήτων πάνω στους κόμβους-στόχους σε κάθε μία από τις αριθμητικές παραμέτρους του δικτύου. Αυτές οι παράγωγοι δίνουν μια ένδειξη σημασίας της ακρίβειας των αριθμητικών παραμέτρων δικτύου για τον υπολογισμό των εκ των υστέρων πιθανοτήτων των στόχων. Εάν η παράγωγος είναι μεγάλη για μια παράμετρο P , τότε έστω και μια μικρή αλλαγή στο P μπορεί να οδηγήσει σε μια σημαντική αλλαγή στις εκ των υστέρων πιθανότητες του. Εάν η παράγωγος είναι μικρή, τότε ακόμη και μεγάλες αλλαγές στην παράμετρο έχουν μικρή διαφορά για την τελευταία[130]. Σχηματικά αυτό μπορεί να παρουσιαστεί στην εικόνα 5.1.

5.6 Αξιολόγηση Μοντέλου

5.6.1 Συνάρτηση Log Likelihood

Στην στατιστική, η Συνάρτηση Πιθανότητας (Likelihood Function) μετρά την θετική προσαρμογή ενός στατιστικού μοντέλου σε ένα δείγμα δεδομένων για δεδομένες τιμές των άγνωστων παραμέτρων. Διαμορφώνεται από την "άπό κοινού" κατανομή πιθανότητας του δείγ-

ματος, αλλά θεωρείται και χρησιμοποιείται μόνο ως συνάρτηση των παραμέτρων, αντιμετωπίζοντας έτσι τις τυχαίες μεταβλητές ως σταθερές στις παρατηρούμενες τιμές. Η Likelihood Function περιγράφει μια υπερεπιφάνεια της οποίας η κορυφή, εάν υπάρχει, αντιπροσωπεύει τον συνδυασμό των τιμών των παραμέτρων του μοντέλου που μεγιστοποιούν την πιθανότητα του ληφθέντος δείγματος. Η διαδικασία για την απόκτηση αυτών των τιμών για την μεγιστοποίηση της συνάρτησης πιθανότητας είναι γνωστή ως maximum likelihood estimation, η οποία για υπολογιστική ευκολία γίνεται συνήθως χρησιμοποιώντας τον φυσικό λογάριθμο της πιθανότητας, γνωστή ως Log-Likelihood Function. Η συνάρτηση φυσικού λογαρίθμου είναι αρνητική για τιμές μικρότερες από 1 και θετική για τιμές μεγαλύτερες από 1. Για διακριτές τυχαίες μεταβλητές θα ισχύει πάντα ότι η συνάρτηση θα παίρνει αρνητικές τιμές.

5.6.2 Διασταυρωμένη Επικύρωση

Σύμφωνα με την τεχνική της διασταυρωμένης επικύρωσης K-πτυχών (K fold cross-validation), τα δεδομένα χωρίζονται σε K ισομερή υποσύνολα, όπου K φυσικός αριθμός. Ύστερα κάθε υποσύνολο χρησιμοποιείται διαδοχικά για έλεγχο (test set) και τα υπόλοιπα αξιοποιούνται για την εκπαίδευση του μοντέλου (training set). Έτσι, μία αξιόπιστη εκτίμηση για την ακρίβεια του μοντέλου προκύπτει από τον μέσο όρο των εκτιμήσεων.

Όταν χρησιμοποιούμε την Ακρίβεια (Accuracy) ως συνάρτηση βαθμολογίας (scoring function), ο αλγόριθμος θα χρησιμοποιήσει την ακρίβεια αυτή ως λειτουργία βαθμολογίας για να αναζητήσει τον βέλτιστο γράφο. Καθορίζουμε μια μεταβλητή κλάσης και την τεχνική της διασταυρωμένης επικύρωσης (cross-validation), leave one out ή K-fold cross-validation, με το πρώτο να είναι μια ειδική περίπτωση του τελευταίου, όταν ο αριθμός είναι ίσος με τον αριθμό εγγραφών στο σύνολο δεδομένων. Χρησιμοποιείται ακρίβεια διαστρωμάτωσης (classification accuracy) μόνο για την σύγκριση των δικτύων που δημιουργούνται από τις (ανεξάρτητες) επαναλήψεις της Μπεϊσιανής Αναζήτησης. Προφανώς, η πραγματική εσωτερική διεργασία του αλγορίθμου δεν επηρεάζεται από αυτήν την επιλογή. Ο αλγόριθμος παράγει τελικά ένα ακυκλικό κατευθυνόμενο γράφημα (ΚΑΓ) που επιτυγχάνει την υψηλότερη βαθμολογία[116].

Κεφάλαιο 6

Αποτελέσματα και Συζήτηση

Τα μπεισιανά δίκτυα αντιπροσωπεύουν κατευθυνόμενα ακυκλικά γραφήματα (ΚΑΓ), όπου κόμβοι και ακμές υποδηλώνουν σχέση αίτιου - αιτιατού μεταξύ διαφορετικών μεταβλητών[102]. Η τοπογραφική δομή των δικτύων αυτών αντικατοπτρίζει την εξάρτηση των μεταβλητών και απεικονίζει την κατανομή πιθανότητας συγκεκριμένων εφαρμογών που εκτελούνται υπό καθορισμένες συνθήκες. Θεωρώντας το $X = X_1, X_2, X_3, \dots, X_n$ ένα σύνολο μεταβλητών n διαστάσεων, το μπεισιανό δίκτυο ορίζεται ως ένα σύνολο ζευγών $X = \langle G, P \rangle$ όπου G συμβολίζει το ΚΑΓ στο οποίο κάθε κόμβος εκφράζει μια μεταβλητή $X_1, X_2, X_3, \dots, X_n$ και κάθε ακμή αντιπροσωπεύει τη σχέση άμεσης εξάρτησης μεταξύ αυτών των μεταβλητών. Επιπλέον, το P δηλώνει το σύνολο των παραμέτρων που ποσοτικοποιούν το δίκτυο, περιέχει δηλαδή τις πιθανότητες κάθε πιθανής τιμής x_i για κάθε μεταβλητή X_i . Αποσυνθέτοντας την από κοινού πιθανότητα P , υπό την υπόθεση ότι κάθε κόμβος είναι ανεξάρτητος από τους μη γονικούς του κόμβους, μπορεί αυτή να υπολογιστεί, χρησιμοποιώντας την συνάρτηση κατανομής πιθανότητας του μπεισιανού δικτύου ως εξής:

$$P(X) = P(X_1, X_2, X_3, \dots, X_n) = \prod_{i=1}^n P\left(\frac{X_i}{X_{j(i)}}\right) \quad (6.1)$$

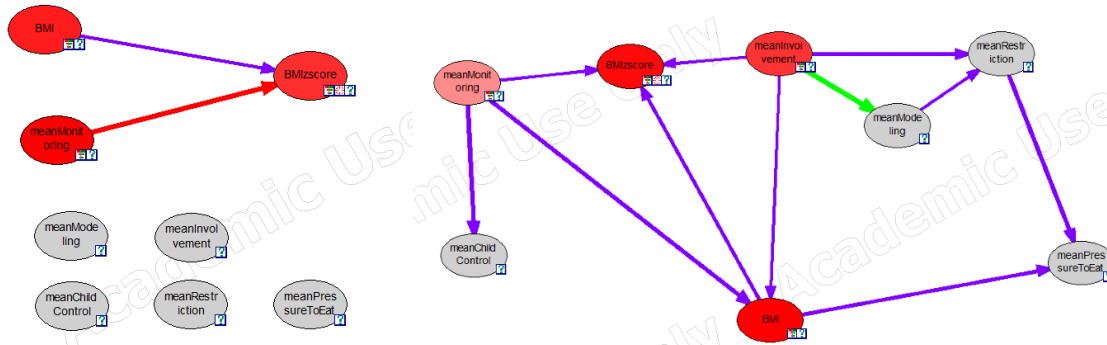
όπου $X_{j(i)}$ το σύνολο των γονικών μεταβλητών του X_i στο G . Από το κεφάλαιο 4 γνωρίζουμε ότι μπορούμε να εξάγουμε την εκ των προτέρων (posterior) πιθανότητα μέσω συμπερασμού στην μεταβλητή ενδιαφέροντος[23].

Οι μεταβλητές ενδιαφέροντος στην παρούσα εργασία έχουν εξαχθεί ως ποσοτικές τιμές από μια συγχρονική μελέτη επιπολασμού παραγόντων που σχετίζονται με το αυξανόμενο βάρος παιδιών ηλικίας έξι έως έντεκα ετών που προσέρχονται από το Ιατρείο Παχυσαρκίας του Νοσοκομείου Παιδών "Αγία Σοφία"[77]. Η Μπεισιανή Ανάλυση Δικτύου επιτεύχθηκε χρησιμοποιώντας το πρόγραμμα BayesFusion GeNIe[117], εφαρμόζοντας ένα σύνολο αλγορίθμων επιβλεπόμενης μάθησης για την εύρεση του βέλτιστου μοντέλου.

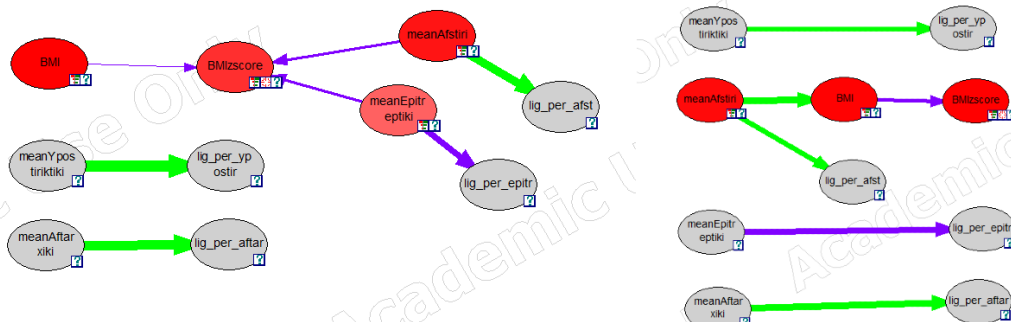
Η J-Απόκλιση και οι Ευκλείδεια, Hellinger και CDF αποστάσεις είναι στατιστικές μέθοδοι που μετρούν την κατεύθυνση και την ισχύ γραμμικών συσχετίσεων μεταξύ δύο τυχαίων μεταβλητών[112]. Η Ανάλυση Ευαισθησίας (Sensitivity Analysis) του δικτύου εκφράζει την επίδραση μικρών ή μεγαλύτερων αλλαγών των μεταβλητών πιθανοτήτων στις μεταβλητές εξόδου του δικτύου, δηλαδή στις εκ των υστέρων πιθανότητες των μεταβλητών ενδιαφέροντος (εδώ: ΔΜΣ και ΔΜΣ ζ-βαθμολογία).

6.1 Παραγόμενα Μπείσιανά Δίκτυα

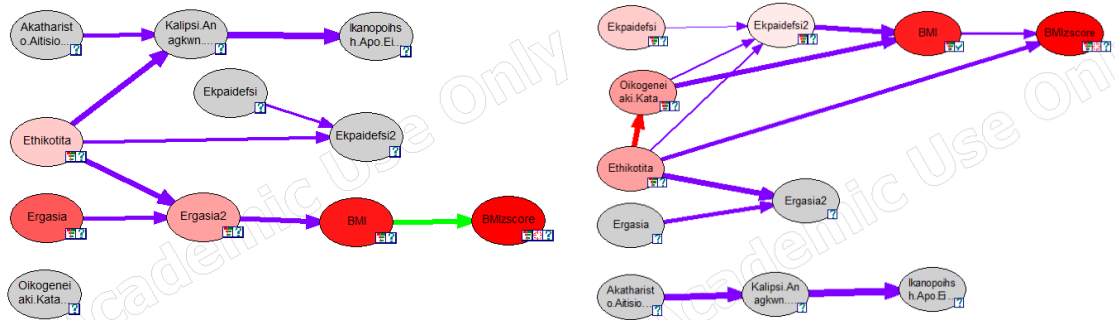
Τα μπείσιανά δίκτυα που παράχθηκαν μέσω του Αλγορίθμου της Μπείσιανής Αναζήτησης (Κεφάλαιο 5) φαίνονται στις Εικόνες 6.1-7.



Εικόνα 6.1: Μπείσιανό Δίκτυο Γονεϊκών Πρακτικών Σίτισης στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)

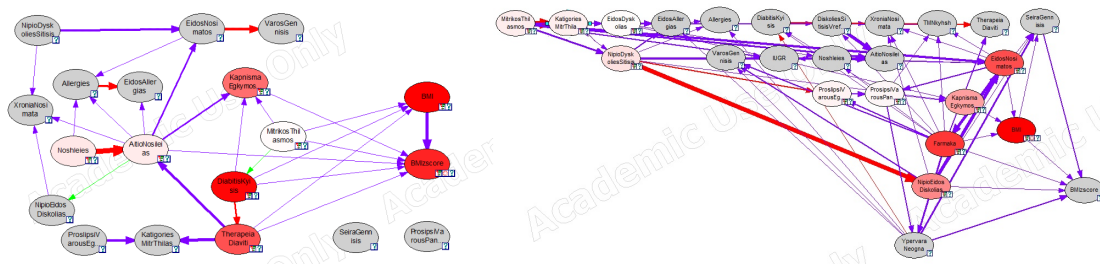


Εικόνα 6.2: Μπείσιανά Δίκτυα Γονεϊκών Τύπων στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)

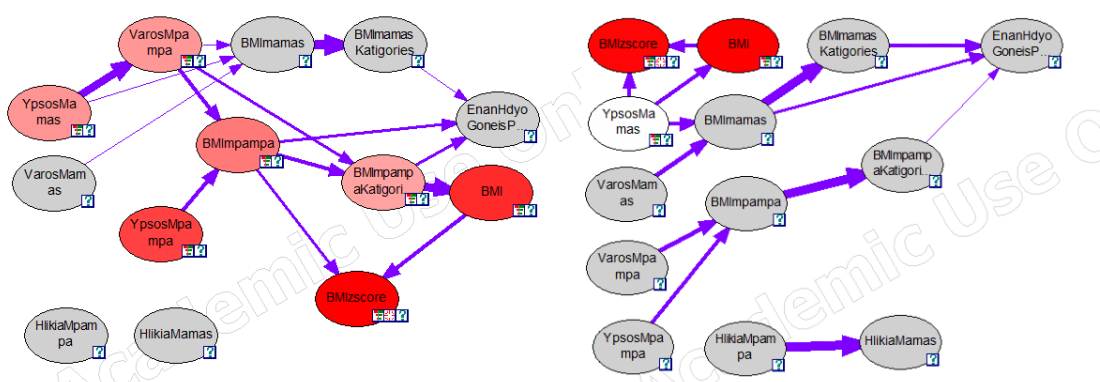


Εικόνα 6.3: Μπείσιανά Δίκτυα Κοινωνικοδημογραφικών Χαρακτηριστικών των Μητέρων στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)

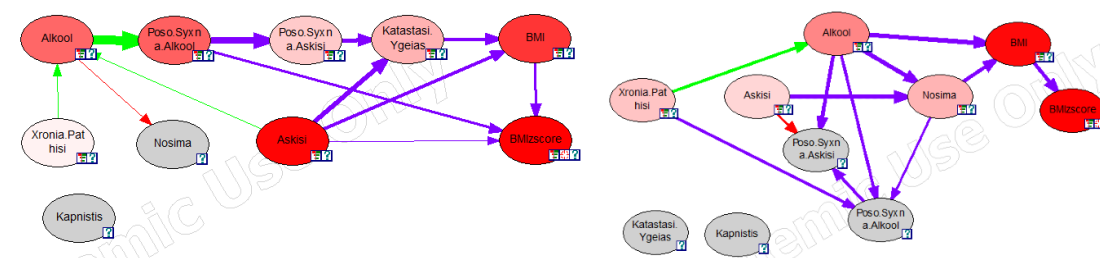
Εφαρμόστηκε Συμπερασμός με βάση την προσέγγιση Μπείσιανής Ανάλυσης Δικτύου με το να αντιπροσωπεύεται η από κοινού κατανομή πιθανότητας ως Μπείσιανό δίκτυο. Οι κόμβοι του δικτύου αντιπροσωπεύουν μεταβλητές ενδιαφέροντος και αιτιώδης και οι ακμές αντιπροσωπεύουν στατιστικές εξαρτήσεις μεταξύ των μεταβλητών. Αυτές οι εξαρτήσεις στο δίκτυο



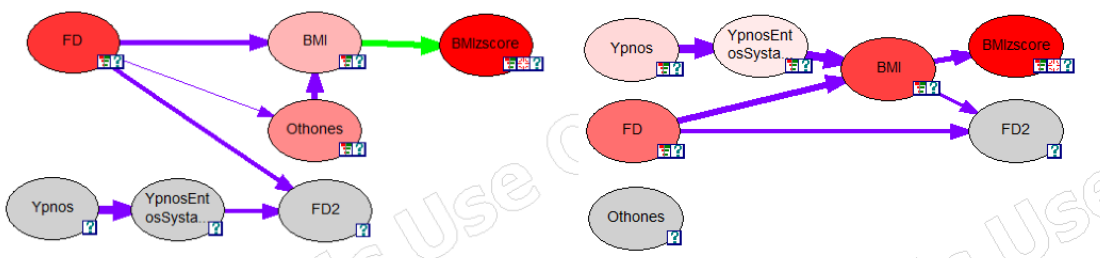
Εικόνα 6.4: Μπείσιανά Δίκτυα Περιγεννητικών Παραγόντων Κινδύνου στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)



Εικόνα 6.5: Μπείσιανά Δίκτυα Σωματομετρικών Χαρακτηριστικών των Γονέων στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)



Εικόνα 6.6: Μπείσιανά Δίκτυα του Τρόπου Ζωής των Γονέων στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)



Εικόνα 6.7: Μπείσιανά Δίκτυα του Τρόπου Ζωής των Παιδιών στο σύνολο των Κοριτσιών (αριστερά) και Αγοριών (δεξιά)

ποσοτικοποιούνται από τις δεσμευμένες πιθανότητες για κάθε έναν κόμβο όπως προκύπτουν από τους γονικούς του κόμβους. Μια εκ των υστέρων πιθανότητα μπορεί να υποστηριχθεί από το δίκτυο οποιοδήποτε υποσυνόλου μεταβλητών δεδομένου στοιχείων για οποιαδήποτε

άλλο υποσύνολο. Ο συμπερασμός μπορεί να επιτευχθεί με εισαγωγή στοιχείων σε σύνολα μεταβλητών με γνωστές καταστάσεις. Συνδυάζοντας τον κανόνα Πιθανοτήτων με τον κανόνα του Μπέις, δημιουργείται ένα πλήρες σύστημα συλλογισμού, το οποίο περιλαμβάνει την παραδοσιακή επαγωγική λογική ως ειδική περίπτωση[23].

Τα Μπεϊσιανά δίκτυα προτάθηκαν αργά στην δεκαετία του 1970 με κίνητρο την μοντελοποίηση συνδυασμών των στοιχείων από κάτω προς τα πάνω (bottom-up perceptual) και από πάνω προς τα κάτω (top-down semantic) ανάγνωσης στοιχείων. Αυτή η αμφίδρομη δυνατότητα συμπεράσματος σε συνδυασμό με το αυστηρό πιθανολογικό θεμέλιο κατέστησαν τα δίκτυα αυτά μια ζαφνική ανάγκη και χρήσιμη επιλογή για αβέβαιη συλλογιστική σε ειδικά συστήματα και στην τεχνητή νοημοσύνη, αντικαθιστώντας παλαιότερα μη γενικευμένα, βασισμένα σε κανόνες συστήματα ad-hoc rule-based systems[23].

6.2 Αποτελέσματα Ισχύος Επιρροής στα Μπεϊσιανά Δίκτυα

Στην παρούσα εργασία επικεντρωνόμαστε στην J-Απόκλιση (με $a = 1$) καθώς αποτελεί την καταλληλότερη μετρική για το σκοπό μας (Πίνακες 6.1-2), και παρέχει τα σημαντικότερα αποτελέσματα. Στους πίνακες 6.1 και 6.2 περιλαμβάνουμε την μέγιστη (Max) και την μέση τιμή (Avg) της συσχέτισης που εμφανίζει μια μεταβλητή γονικού κόμβου σε μια μεταβλητή κόμβου παιδιού. Καταγράφουμε τις υπόλοιπες μετρικές (πίνακες 6.3-4) για λόγους πληρότητας και σύγκρισης. Η Ευκλείδεια απόσταση, όντας η πιο απλή αναφορικά με τις υπόλοιπες κι αυτή που απαιτεί την λιγότερη πληροφορία για τον υπολογισμό της, παρέχει και τα λιγότερο καλά αποτελέσματα, γεγονός που περιμέναμε. Ομοίως και η ΣΑΚ, αφού σε πολλές περιπτώσεις δεν πληρούταν το προαπαιτούμενο της διάταξης για όλους τους κόμβους των δικτύων. Τέλος η Hellinger προσδίδει πιο ενδιάμεσα αποτελέσματα, διότι παρέχει περισσότερη πληροφορία από την Ευκλείδεια, αλλά συνεχίζει να είναι λιγότερο ευαίσθητη από την J-Απόκλιση. Οι τιμές στους πίνακες 6.3 και 6.4 αφορούν μόνο την μέγιστη τιμή της συσχέτισης (Max) που εμφανίζει μια μεταβλητή γονικού κόμβου σε μια μεταβλητή κόμβου παιδιού, κι όχι την μέση τιμή ή την σταθμισμένη μέση τιμή. Το εύρος των τιμών αυτών μπορεί να κυμανθεί από το 0 (δεν υπάρχει ακμή και άρα, συσχέτιση) μέχρι το 1 (μέγιστη συσχέτιση). Κάθε συσχέτιση εκφράζει την θετική επίδραση του γονικού κόμβου στον κόμβο παιδί, κι αφού οι κόμβοι παιδιά αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία, ουσιαστικά εκφράζεται η επίδραση του εκάστοτε χαρακτηριστικού στην αύξηση του ΔΜΣ και της ΔΜΣ ζ-βαθμολογίας των παιδιού.

Από το σύνολο των συσχετίσεων που παράχθηκαν από τα Μπεϊσιανά Δίκτυα (ακμές), καταγράφουμε εδώ μόνο εκείνες που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία των παιδιών, καθώς αποτελούν τις μεταβλητές κύριου ενδιαφέροντος. Τα μπεϊσιανά δίκτυα παράχθηκαν ξεχωριστά σε αγόρια και κορίτσια, καθώς στις ηλικίες των 6 με 11 ετών που αφορούν τα δεδομένα, οι σωματομετρικές διαφορές των δύο φύλων χρήζουν διαφορετικής αντιμετώπισης. Χαρακτηριστικό παράδειγμα, ο ίδιος ο υπολογισμός του ΔΜΣ και της ΔΜΣ ζ-βαθμολογίας αποκλίνει αρκετά μεταξύ δύο παιδιών ίδιας ηλικίας, βάρους και ύψους, αλλά διαφορετικού φύλου.

Το BayesFusion GeNIe μας διευκολύνει να απεικονίσουμε την ποσοτική πλευρά του δικτύου πέρα από την οπτική επιθεώρηση της δομής του δικτύου. Τα δίκτυα παρέχουν μια

Πίνακας 6.1: Καταγραφή της *J*-Απόκλισης (με $\alpha = 1$) που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία για το σύνολο των Κοριτσιών. Οι συσχετίσεις έχουν υπολογιστεί για την μέγιστη και μέση τιμή.

Γονέας	Παιδί	Max	Avg
Οθόνες	ΔΜΣ	0.95851	0.542261
Παρακολούθηση στο Φαγητό	ΔΜΣ ζ-β.	0.914723	0.584837
Φυσική Δραστηριότητα	ΔΜΣ	0.895181	0.401655
Επιτρεπτική Μητέρα	ΔΜΣ ζ-β.	0.886069	0.170399
ΔΜΣ Πατέρα	ΔΜΣ ζ-β.	0.870445	0.279441
Συχνότητα Αλκοόλ	ΔΜΣ ζ-β.	0.849394	0.213901
Άσκηση Γονέα	ΔΜΣ ζ-β.	0.849394	0.14702
Μητρικός Θηλασμός	ΔΜΣ ζ-β.	0.843927	0.0921878
ΔΜΣ ζ-β. Πατέρα	ΔΜΣ	0.837128	0.722609
Αυστηρή Μητέρα	ΔΜΣ ζ-β.	0.826238	0.165734
Διαβήτης Κύησης	ΔΜΣ ζ-β.	0.824173	0.133944
Τύπος Εργασίας	ΔΜΣ	0.770504	0.573801
Κατάσταση Υγείας Γονέα	ΔΜΣ	0.741598	0.33517
Θεραπεία Διαβήτη	ΔΜΣ ζ-β.	0.737925	0.125117
Αίτιο Νοσηλείας	ΔΜΣ ζ-β.	0.668078	0.0568489
Κάπνισμα Εγκυμοσύνη	ΔΜΣ ζ-β.	0.606943	0.0986555
Άσκηση Γονέα	ΔΜΣ	0.569676	0.32127
Διαβήτης Κύησης	ΔΜΣ	0.556688	0.195272
Θεραπεία Διαβήτη	ΔΜΣ	0.473253	0.143871
Μητρικός Θηλασμός	ΔΜΣ	0.203783	0.0931944

λεπτομερή ανάλυση του γνωσιακού τομέα για μεταβλητές ενδιαφέροντος μέσω συμπαγούς αναπαράστασης της από κοινού κατανομής πιθανότητας. Αυτά τα δίκτυα είναι εγγενώς πιθανολογικά. Επιπλέον, μπεϊσιανά δίκτυα μπορούν να αναπαραστήσουν με μαθηματικό τρόπο κατευθύνσεις επιδράσεων που η κλασική Άλγεβρα δεν μπορεί. Στις Εικόνες (6.1-7), οι συσχετίσεις μεταξύ των χαρακτηριστικών (δεδομένα) ποσοτικοποιούνται υπό τον όρο της *J*-Απόκλισης. Το πάχος κάθε ακμής είναι ανάλογο της Ισχύος της Επιρροής μεταξύ δύο κόμβων. Τα κόκκινα και πράσινα χρώματα συμβολίζουν αρνητική και θετική συσχέτιση αντίστοιχα, ενώ το ιώδες (μωβ) δηλώνει συσχέτιση και προς τις δύο κατευθύνσεις.

Αναφορικά με το σύνολο δεδομένων των κοριτσιών, η μέγιστη ισχύς επιρροής εμφανίστηκε στο ζεύγος (Οθόνες, ΔΜΣ) με μέγιστη τιμή 0.95851 χρησιμοποιώντας *J*-Απόκλιση (Εικόνα 6.7), ενώ επίσης πολύ υψηλή συσχέτιση εμφάνισε το ζεύγος (Γονεϊκή Παρακολούθηση στο Φαγητό, ΔΜΣ ζ-β.) με 0.914723 (Εικόνα 6.1). Πάνω από 0.85 εμφάνισαν τα ζεύγη (Φυσική Δραστηριότητα, ΔΜΣ) με 0.895181, (Επιτρεπτική Μητέρα, ΔΜΣ ζ-β.) με 0.886069 και (ΔΜΣ Πατέρα, ΔΜΣ ζ-β.) με 0.870445. Ακόμα, στο εύρος [0.8 – 0.85] έχουμε τα ζεύγη (Συχνότητα Αλκοόλ Γονέα, ΔΜΣ ζ-β.) και (Άσκηση Γονέα, ΔΜΣ ζ-β.) με 0.849394, (Μητρικός Θηλασμός, ΒΜΙ ζ-β.) με 0.843927, (ΔΜΣ ζ-β. Πατέρα, ΔΜΣ) με 0.837128, (Αυστηρή Μητέρα, ΒΜΙ ζ-β.) με 0.826238 και το ζεύγος (Διαβήτης Κύησης, ΔΜΣ ζ-β.) με 0.824173. Τέλος, αξίζουν να αναφερθούν τα ζεύγη (Τύπος Εργασίας, ΔΜΣ) με 0.770504, (Κατάσταση Υγείας Γονέα, ΔΜΣ) με 0.741598 και (Θεραπεία Διαβήτη, ΔΜΣ ζ-β.) με 0.737925. Πάνω από το φράγμα του 0.6 καταλήγουν τα ζεύγη (Αίτιο Νοσηλείας, ΔΜΣ ζ-β.) και (Κάπνισμα στην Εγκυμοσύνη, ΔΜΣ ζ-β.), ενώ πάνω από 0.5 (Άσκηση Γονέα, ΔΜΣ) και (Διαβήτης Κύησης, ΔΜΣ) (Πίνακας 6.1.).

Για το σύνολο δεδομένων των αγοριών η μέγιστη ισχύς επιρροής εμφανίστηκε στο ζεύγος

Πίνακας 6.2: Καταγραφή της J-Απόκλισης (με $a = 1$) που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία για το σύνολο των Αγοριών. Οι συσχετίσεις έχουν υπολογιστεί για την μέγιστη και μέση τιμή.

Γονέας	Παιδί	Max	Avg
ΔΜΣ	Πίεση στο Φαγητό	0.9582	0.445098
Φυσική Δραστηριότητα	ΔΜΣ	0.952806	0.522708
Τύπος Εκπαίδευσης	ΔΜΣ	0.941632	0.531103
Συμμετοχή στο Φαγητό	ΔΜΣ ζ-β.	0.929828	0.345264
Ύπνος Εντός Συστάσεων	ΔΜΣ	0.917243	0.358179
Χρόνιο Νόσημα Παιδιού	ΔΜΣ	0.88992	0.438004
Ύψος Μητέρας	ΔΜΣ	0.887694	0.395488
Οικογενειακή Κατάσταση	ΔΜΣ	0.879093	0.533525
Παρακολούθηση στο Φαγητό	ΔΜΣ	0.877907	0.586838
Παρακολούθηση στο Φαγητό	ΔΜΣ ζ-β.	0.867202	0.358179
Αυστηρή Μητέρα	ΔΜΣ	0.864425	0.786711
Κατανάλωση Αλκοόλ	ΔΜΣ	0.858849	0.506545
Φάρμακα	ΔΜΣ ζ-β.	0.858558	0.141117
Εθνικότητα	ΔΜΣ ζ-β.	0.857872	0.486319
Σειρά Γέννησης	ΔΜΣ ζ-β.	0.849394	0.267338
Συμμετοχή στο Φαγητό	ΔΜΣ	0.840062	0.370986
Ύψος Μητέρας	ΔΜΣ ζ-β.	0.823833	0.304539
Είδος Νοσήματος	ΔΜΣ	0.7798	0.160429
Νήπιο Είδος Δυσκολίας	ΔΜΣ ζ-β.	0.754127	0.122632
ΔΜΣ	Φυσική Δραστηριότητα	0.622769	0.234338
Υπέρβαρα Νεογνά	ΔΜΣ ζ-β.	0.620423	0.242309
ΔΜΣ	Σειρά Γέννησης	0.618278	0.0225346
Φάρμακα	ΔΜΣ	0.563809	0.174694
Κάπνισμα στην Εγκυμοσύνη	ΔΜΣ	0.532767	0.143451

(Πίεση για Φαγητό, ΔΜΣ) με 0.9582 (Εικόνα 6.1), ενώ πολύ κοντά ήταν και το ζεύγος (Φυσική Δραστηριότητα, ΔΜΣ) με 0.952806 (Εικόνα 6.7). Πάνω από 0.9 είχαν επίσης τα ζεύγη (Τύπος Εκπαίδευσης Μητέρας, ΔΜΣ) με 0.941632 (Εικόνα 6.3), (Επιρροή στο Φαγητό, ΔΜΣ ζ-β.) με 0.929828 (Εικόνα 6.1) και (Ύπνος Εντός Συστάσεων, ΔΜΣ) με 0.917243 (Εικόνα 6.7). Πάνω από 0.85 εμφάνισαν τα ζεύγη (Χρόνιο Νόσημα Παιδιού, ΔΜΣ) με 0.88992, (Ύψος Μαρμάς, ΔΜΣ) με 0.88769, (Οικογενειακή Κατάσταση, ΔΜΣ) με 0.879093, (Παρακολούθηση Φαγητού, ΔΜΣ) με 0.877907, (Παρακολούθηση Φαγητού, ΔΜΣ ζ-β.) με 0.867202, (Αυστηρή, ΔΜΣ) με 0.864425, (Κατανάλωση Αλκοόλ Γονέα, ΔΜΣ) με 0.858849, (Φάρμακα, ΔΜΣ ζ-β.) με 0.858558 και (Εθνικότητα, ΔΜΣ ζ-β.) με 0.857872. Στο εύρος [0.8 – 0.85] έχουμε τα ζεύγη (Σειρά Γέννησης, ΔΜΣ ζ-β.) με 0.849394, (Συμμετοχή του Παιδιού στο φαγητό, ΔΜΣ) με 0.840062 και (Ύψος Μαρμάς, ΔΜΣ ζ-β.) με 0.823833. Αξίζει να αναφερθούν τα ζεύγη (Είδος Νοσήματος, ΔΜΣ) με 0.7798 και (Νήπιο Είδος Δυσκολίας, ΔΜΣ ζ-β.) με 0.754127. Πάνω από 0.6 εντοπίζουμε στα (Φυσική Δραστηριότητα, ΔΜΣ) με 0.622769, (Υπέρβαρα Νεογνά, ΔΜΣ ζ-β.) με 0.620423 και (Σειρά Γέννησης, ΔΜΣ) με 0.618278. Πάνω από 0.5 έχουμε στα (Φάρμακα, ΔΜΣ) με 0.563809 και (Κάπνισμα στην Εγκυμοσύνη, ΔΜΣ) με 0.532767 (Πίνακας 6.2).

Αν επικεντρωθούμε στην μέση τιμή των συσχετίσεων που προκύπτουν από την J-Απόκλιση,

Πίνακας 6.3: Καταγραφή των Συσχετίσεων που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία για το σύνολο των Κοριτσιών. Οι συσχετίσεις έχουν υπολογιστεί για την μέγιστη τιμή.

Γονέας	Παιδί	Hellinger	Ευκλείδεια	ΣΑΚ
Οθόνες	ΔΜΣ	0.689562	0.825592	0.824966
Παρακολούθηση στο Φαγητό	ΔΜΣ ζ-β.	0.564811	0.7	0.7
Φυσική Δραστηριότητα	ΔΜΣ	0.561747	0.646945	0.430333
Επιτρεπτική Μητέρα	ΔΜΣ ζ-β.	0.54348	0.708333	0.78333
ΔΜΣ Πατέρα	ΔΜΣ ζ-β.	0.535474	0.678571	0.678571
Συχνότητα Αλκοόλ	ΔΜΣ ζ-β.	0.504623	0.666667	0.666667
Άσκηση Γονέα	ΔΜΣ ζ-β.	0.504623	0.666667	0.666667
Μητρικός Θηλασμός	ΔΜΣ ζ-β.	0.49956	0.66092	0.66092
ΔΜΣ ζ-β. Πατέρα	ΔΜΣ	0.499131	0.560536	0.318182
Αυστηρή Μητέρα	ΔΜΣ ζ-β.	0.475507	0.625	0.625
Διαβήτης Κύησης	ΔΜΣ ζ-β.	0.482473	0.641026	0.641026
Τύπος Εργασίας	ΔΜΣ	0.443869	0.5	0.391667
Κατάσταση Υγείας Γονέα	ΔΜΣ	0.421677	0.450925	0.25
Θεραπεία Διαβήτη	ΔΜΣ ζ-β.	0.412742	0.546798	0.546798
Αίτιο Νοσηλείας	ΔΜΣ ζ-β.	0.378452	0.511918	0.511918
Κάπνισμα Εγκυμοσύνη	ΔΜΣ ζ-β.	0.356551	0.47619	0.47619
Άσκηση Γονέα	ΔΜΣ	0.333333	0.333333	0.25
Διαβήτης Κύησης	ΔΜΣ	0.332236	0.429811	0.215056
Θεραπεία Διαβήτη	ΔΜΣ	0.299156	0.386877	0.206738
Μητρικός Θηλασμός	ΔΜΣ	0.189568	0.22304	0.128378

η ανάγνωση των αποτελεσμάτων είναι διαφορετική. Για το σύνολο των δεδομένων των κοριτσιών η υψηλότερη μέση τιμή εμφανίζεται στο ζεύγος (ΔΜΣ ζ-β. Πατέρα, ΔΜΣ) με 0.722609. Πάνω από το φράγμα του 0.5 έχουμε τα ζεύγη (Γονεϊκή Παρακολούθηση στο Φαγητό, ΔΜΣ ζ-β.) με 0.584837, (Τύπος Εργασίας Γονέα, ΔΜΣ) με 0.57380 και (Οθόνες, ΔΜΣ) με 0.542261. Το ζεύγος (Φυσική Δραστηριότητα, ΔΜΣ) παρατηρείται πάνω από 0.4, ενώ τα ζεύγη (Κατάσταση Υγείας Γονέα, ΔΜΣ) και (Άσκηση Γονέα, ΔΜΣ) είναι πάνω από 0.3.

Για το σύνολο των δεδομένων των αγοριών, η υψηλότερη μέση τιμή εμφανίζεται στο ζεύγος (Αυστηρή Μητέρα, ΔΜΣ) με 0.786711. Πάνω από 0.5 έχουν κατά σειρά τα ζεύγη (Παρακολούθηση στο Φαγητό, ΔΜΣ), (Οικογενειακή Κατάσταση, ΔΜΣ), (Τύπος Εκπαίδευσης, ΔΜΣ), (Φυσική Δραστηριότητα, ΔΜΣ) και (Κατανάλωση Αλκοόλ Γονέα, ΔΜΣ). Μέση ισχύ πάνω από 0.4 παρατηρούμε στα ζεύγη (Εθνικότητα, ΔΜΣ), (ΔΜΣ, Πίεση στο Φαγητό) και (Χρόνιο Νόσημα Παιδιού, ΔΜΣ), ενώ πάνω από 0.3 στα ζεύγη (Ύψος Μητέρας, ΔΜΣ), (Συμμετοχή στο Φαγητό, ΔΜΣ), (Παρακολούθηση στο Φαγητό, ΔΜΣ ζ-β.), (Ύπνος Εντός Συστάσεων, ΔΜΣ), (Συμμετοχή στο Φαγητό, ΔΜΣ ζ-β.) και τέλος (Ύψος Μητέρας, ΔΜΣ ζ-β.).

Στους Πίνακες 6.3 και 6.4 καταγράφουμε τα αποτελέσματα των συσχετίσεων των υπόλοιπων τριών μετρικών αποστάσεων (Hellinger, Ευκλείδειας και ΣΑΚ) για λόγους πληρότητας. Είναι προφανές ότι προσδίδουν αρκετά χαμηλότερη κατά μέτρο συσχέτιση σε σχέση με την J-Απόκλιση. Αυτό, εκτός των άλλων, επιβεβαιώνει την αρχική υπόθεσή μας ότι καθώς αποτελούν απλούστερες μετρικές που απαιτούν λιγότερη πληροφορία κατά των υπολογισμών, χάνουν σε ακρίβεια και ευαισθησία.

Συνολικά, οι ισχυρότερες συσχετίσεις βρέθηκαν μεταξύ διαφόρων ζευγών μεταβλητών. Τα ζεύγη (Οθόνες, ΔΜΣ) και (Γονεϊκή Παρακολούθηση στο Φαγητό, ΔΜΣ ζ-βαθμολογία) για το

Πίνακας 6.4: Καταγραφή των Συσχετίσεων που αφορούν τον ΔΜΣ και την ΔΜΣ ζ-βαθμολογία για το σύνολο των Αγοριών. Οι συσχετίσεις έχουν υπολογιστεί για την μέγιστη τιμή.

Γονέας	Παιδί	Hellinger	Ευκλείδεια	ΣΑΚ
ΔΜΣ	Πίεση στο Φαγητό	0.68559	0.774344	0.694444
Φυσική Δραστηριότητα	ΔΜΣ	0.688351	0.738323	0.408333
Τύπος Εκπαίδευσης	ΔΜΣ	0.619776	0.665525	0.472222
Συμμετοχή στο Φαγητό	ΔΜΣ ζ-β.	0.628101	0.7875	0.7875
Ύπνος Εντός Συστάσεων	ΔΜΣ	0.58726	0.625694	0.283333
Χρόνιο Νόσημα Παιδιού	ΔΜΣ	0.554548	0.589256	0.388889
Ύψος Μητέρας	ΔΜΣ	0.5395	0.59797	0.374206
Οικογενειακή Κατάσταση	ΔΜΣ	0.540064	0.512239	0.403846
Παρακολούθηση στο Φαγητό	ΔΜΣ	0.478	0.511695	0.46517
Παρακολούθηση στο Φαγητό	ΔΜΣ ζ-β.	0.529659	0.685714	0.685714
Αυστηρή Μητέρα	ΔΜΣ	0.525001	0.624319	0.357143
Κατανάλωση Αλκοόλ	ΔΜΣ	0.511343	0.586302	0.25
Φάρμακα	ΔΜΣ ζ-β.	0.514021	0.677217	0.677217
Εθνικότητα	ΔΜΣ ζ-β.	0.520852	0.666667	0.666667
Σειρά Γέννησης	ΔΜΣ ζ-β.	0.504623	0.666667	0.666667
Συμμετοχή στο Φαγητό	ΔΜΣ	0.489545	0.587862	0.233549
Ύψος Μητέρας	ΔΜΣ ζ-β.	0.462084	0.535566	0.59797
Είδος Νοσήματος	ΔΜΣ	0.44681	0.5	0.314815
Νήπιο Είδος Δυσκολίας	ΔΜΣ ζ-β.	0.42855	0.576923	0.576923
ΔΜΣ	Φυσική Δραστηριότητα	0.360045	0.428571	0.428571
Υπέρβαρα Νεογνά	ΔΜΣ ζ-β.	0.353299	0.471631	0.471631
ΔΜΣ	Σειρά Γέννησης	0.355654	0.416667	0.277778
Φάρμακα	ΔΜΣ	0.334863	0.306186	0.25
Κάπνισμα στην Εγκυμοσύνη	ΔΜΣ	0.315109	0.306186	0.25

σύνολο των κοριτσιών, και τα ζεύγη (Πίεση στο Φαγητό, ΔΜΣ) και (Φυσική Δραστηριότητα, ΔΜΣ) για το σύνολο των αγοριών, καταλήγουν στις υψηλότερες παρατηρούμενες μέγιστες τιμές. Αν εστιάσουμε στην μέση τιμή των συσχετίσεων, οι υψηλότερες μέσες τιμές παρατηρούνται στο ζεύγος (ΔΜΣ ζ-βαθμολογία Πατέρα, ΔΜΣ) στο σύνολο των κοριτσιών και στο ζεύγος (Αυστηρή Μητέρα, ΔΜΣ) στο σύνολο των Αγοριών.

6.3 Αποτελέσματα Ανάλυσης Ευαισθησίας στα Μπείσιανά Δίκτυα

Στις Εικόνες (6.1-7) οι κόμβοι που είναι χρωματισμένοι με κόκκινο χρώμα περιέχουν καταστάσεις που είναι σημαντικές για τον υπολογισμό της εκ των υστέρων κατανομής πιθανότητας των κόμβων στόχων (εδώ κόμβοι στόχοι είναι ο ΔΜΣ και η ΔΜΣ ζ-βαθμολογία). Προφανώς η ένταση του κόκκινου χρώματος υποδηλώνει και την αντίστοιχη επιρροή στον υπολογισμό. Οι υπόλοιποι μην χρωματισμένοι κόμβοι (γκρι χρώμα) δεν περιέχουν καταστάσεις που εμπειρίζονται στον παραπάνω υπολογισμό, και επομένως η Ευαισθησία τους είναι μηδενική και καθορίζεται ποιοτικά, πριν από κάθε άλλο υπολογισμό κατά την δημιουργία του μπείσιανού δικτύου, σύμφωνα με ένα επίπεδο σχετικού υπολογισμού στο GeNIe. Είναι σημαντικό να

Πίνακας 6.5: Καταγραφή της Ανάλυσης Ευαισθησίας, της Ακρίβειας και της συνάρτησης EM Log Likelihood για το σύνολο των Κοριτσιών. Η ευαισθησία των ΔΜΣ και ΔΜΣ ζ-βαθμολογίας έχει υπολογιστεί για την μέγιστη τιμή.

	ΔΜΣ	ΔΜΣ ζ-β.	Ακρίβεια %	EM Log L.
Γονεϊκές Πρακτικές Σίτισης	0.408	0.35	0.8	-112.63
Γονεϊκοί Τύποι Μητέρων	0.316	0.18	0.75	-112.273
Κοινωνικοδημογραφικά Μητέρων	0.621	0.472	0.8	-172.798
Σωματομετρικά Γονέων	0.444	0.234	0.85	-213.501
Τρόπος Ζωής Γονέων	0.292	0.172	0.8	-122.183
Τρόπος Ζωής Παιδιών	0.132	0.453	0.8	-92.6183
Περιγεννητικοί Παράγοντες	0.368	0.164	0.9	-185.97

Πίνακας 6.6: Καταγραφή της Ανάλυσης Ευαισθησίας, της Ακρίβειας και της συνάρτησης EM Log Likelihood για το σύνολο των Αγοριών. Η ευαισθησία των ΔΜΣ και ΔΜΣ ζ-βαθμολογίας έχει υπολογιστεί για την μέγιστη τιμή.

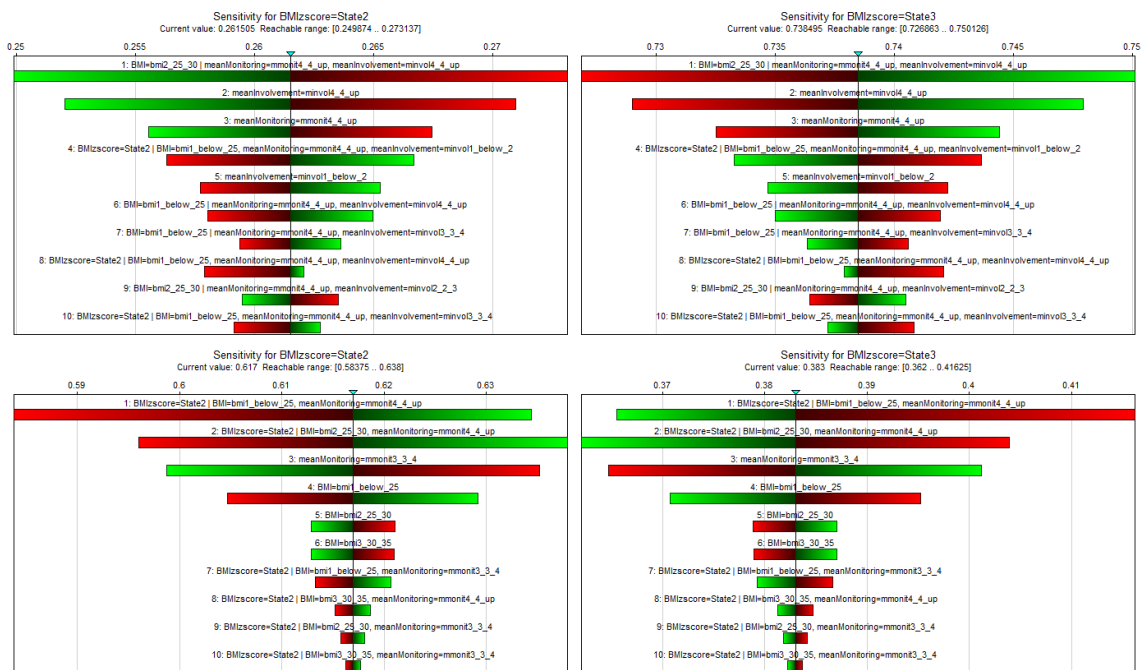
	ΔΜΣ	ΔΜΣ ζ-β.	Ακρίβεια %	EM Log L.
Γονεϊκές Πρακτικές Σίτισης	0.362	0.276	0.964286	-139.1
Γονεϊκοί Τύποι Μητέρων	0.697	0.577	0.928571	-164.098
Κοινωνικοδημογραφικά Μητέρων	0.405	0.499	0.892857	-252.043
Σωματομετρικά Γονέων	0.464	0.315	0.928571	-314.701
Τρόπος Ζωής Γονέων	0.527	0.494	0.928571	-179.973
Τρόπος Ζωής Παιδιών	0.302	0.463	0.928571	-145.098
Περιγεννητικοί Παράγοντες	0.451	0.143	0.964286	-308.725

υπογραμμίσουμε εδώ ότι ο αλγόριθμος Ανάλυσης Ευαισθησίας είναι άμεσα συνδεδεμένος με την δομή του δικτύου και τις συσχετίσεις των μεταβλητών.

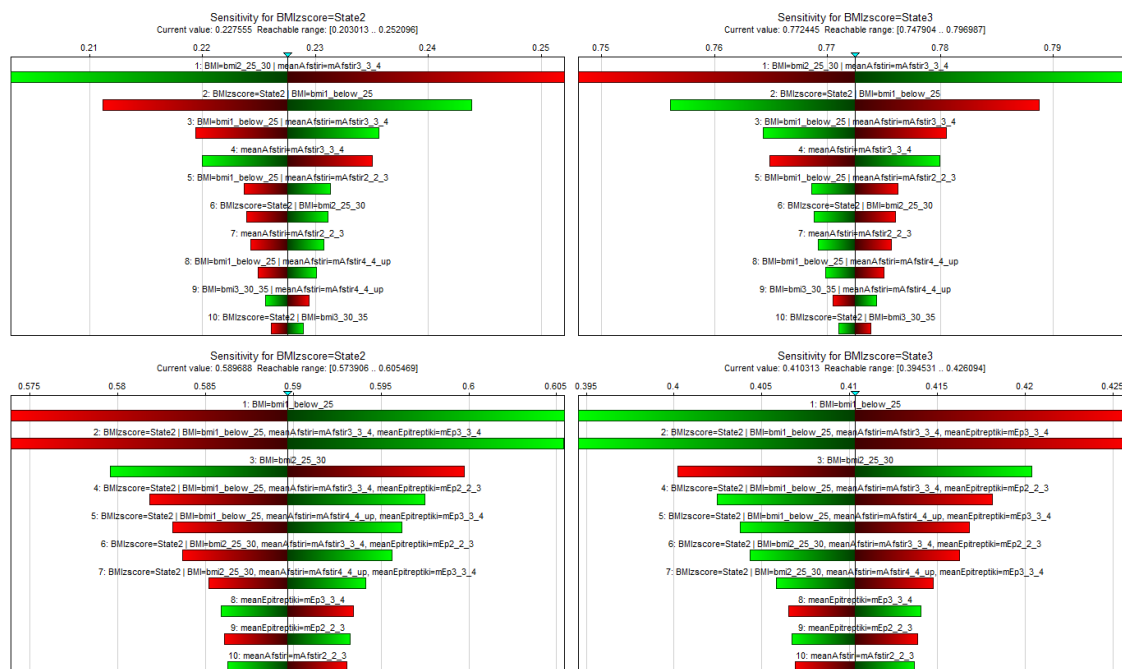
Στην πρώτη στήλη των Πινάκων 6.5 και 6.6 καταγράφουμε την μέγιστη τιμή της Ευαισθησίας που εντοπίζεται στους κόμβους ΔΜΣ και ΔΜΣ ζ-βαθμολογία. Αυτό ουσιαστικά σημαίνει ότι στο συγκεκριμένο δίκτυο υπήρξε τουλάχιστον μία κατάσταση ενός γονικού κόμβου του στόχου που είχε την αναγραφόμενη επίδραση (με μέγιστο το 1) σε κάποια κατάσταση του κόμβου-παιδιού, πράγμα που ουσιαστικά εκφράζει ποσοτικά και ποιοτικά την επίδραση των κόμβων στον ΔΜΣ και στην ΔΜΣ ζ-βαθμολογία των παιδιών.

Το ακόλουθο σημείο χρήζει ιδιαίτερης προσοχής: Η Ανάλυση Ευαισθησίας του δικτύου αναπαριστά την επίδραση του ρυθμού μεταβολής κάθε μεμονωμένης κατάστασης ενός γονικού (και γενικότερα προγονικού) κόμβου σε μια συγκεκριμένη κατάσταση του κόμβου στόχου. Με αυτόν τον τρόπο, ενώ η Συσχέτιση μεταξύ δύο κόμβων μπορεί να είναι μεγάλη (ή μικρή), η Ευαισθησία μεταξύ δύο καταστάσεων μπορεί να είναι αντίστοιχα μικρή (ή μεγάλη) μεμονωμένα αλλά όχι συνολικά. Κοινώς, με την Ευαισθησία εντοπίζουμε συγκεκριμένες καταστάσεις από γονικούς κόμβους του κόμβου στόχου που επιδρούν σε αυτόν, ενώ η Συσχέτιση υπολογίζει την εκ προοιμίου σχέση αίτιου - αιτιατού (και την ποσοτικοποιεί) μεταξύ αυστηρά δύο συνδεδεμένων κόμβων.

Στις Εικόνες 6.8-14 απεικονίζονται οι πιο ευαίσθητες παράμετροι για τις επιλεγμένες καταστάσεις ενδιαφέροντος του κόμβου στόχου (εδώ: sd 2 και sd 3 της ΔΜΣ ζ-βαθμολογίας), διαταγμένες από την περισσότερο επιδραστική στην λιγότερο. Για κάθε παράμετρο βλέπουμε την ακριβή τοποθεσία της στο δίκτυο (τον κόμβο και την κατάστασή του δεσμευμένη από τις



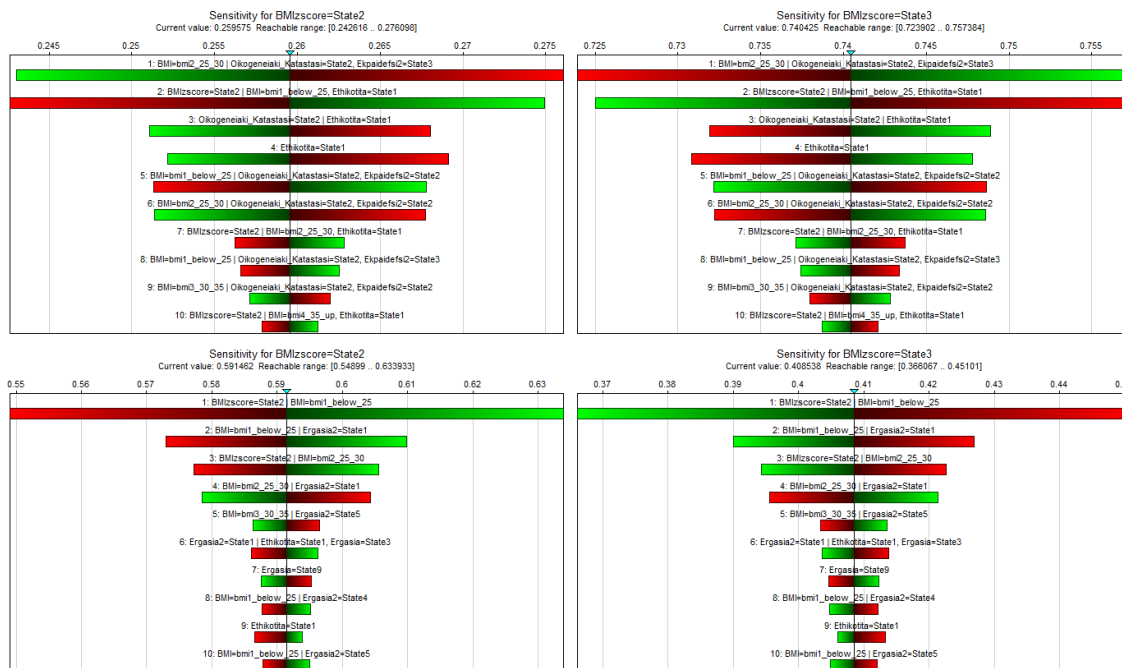
Εικόνα 6.8: Διαγράμματα Ευαισθησίας των Γονεϊκών Πρακτικών Σίτισης στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)



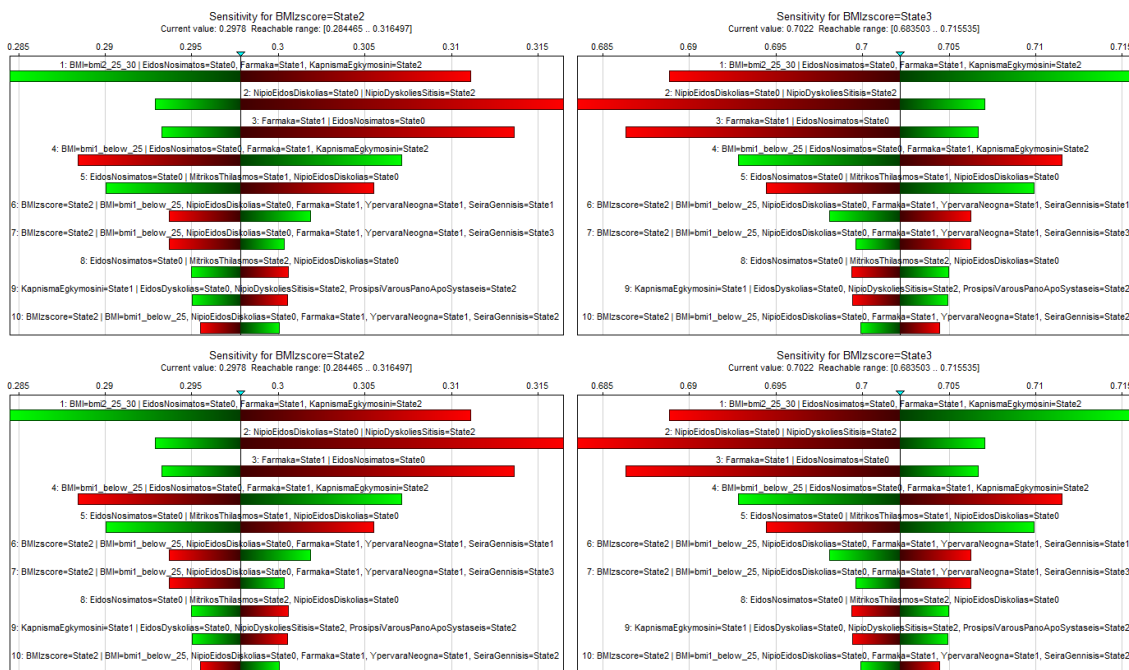
Εικόνα 6.9: Διαγράμματα Ευαισθησίας των Γονεϊκών Συμπεριφορών στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)

καταστάσεις των γονέων του). Ο άξονας των x δείχνει το εύρος της απόλυτης τιμής της αλλαγής της κατάστασης στόχου όταν αλλάξει η εκάστοτε παράμετρος 10% προς τα κάτω (αριστερό ημιεπίπεδο) και 10% προς τα πάνω (δεξιό ημιεπίπεδο). Το κόκκινο χρώμα εκφράζει

αρνητική κατεύθυνση της αλλαγής και το πράσινο θετική.



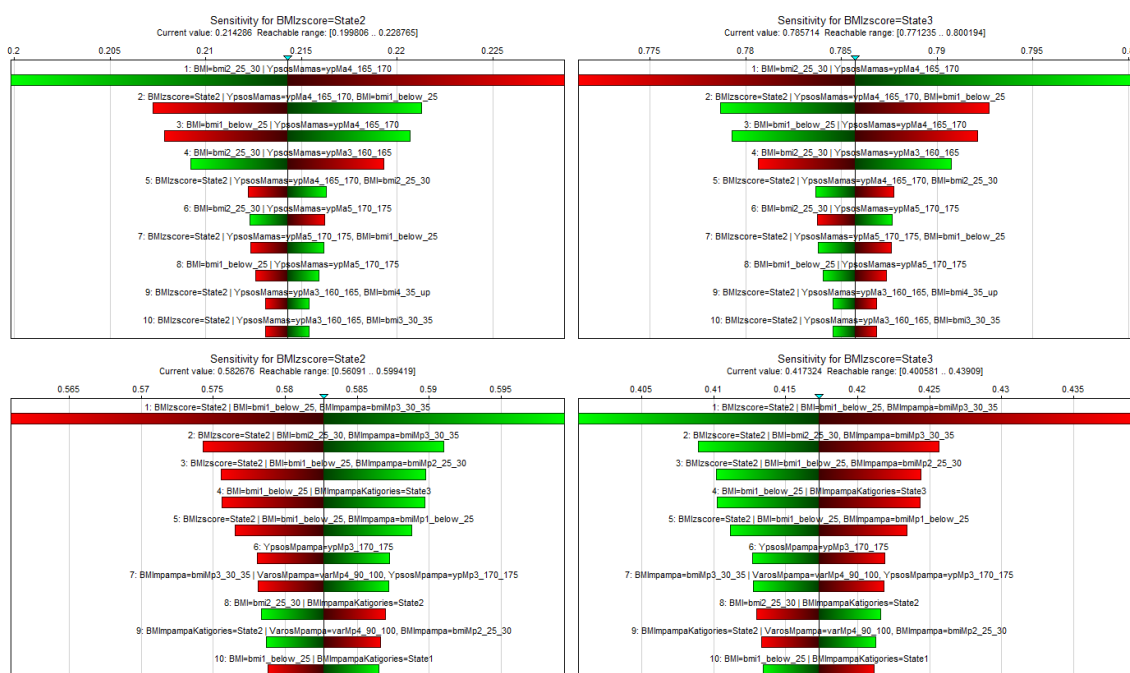
Εικόνα 6.10: Διαγράμματα Ευαισθησίας των Κοινωνικοδημογραφικών Χαρακτηριστικών των Γονέων στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)



Εικόνα 6.11: Διαγράμματα Ευαισθησίας των Περιγεννητικών Παραγόντων Κινδύνου στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)

Για το σύνολο των κοριτσιών για παράδειγμα, στην Εικόνα 6.9 φαίνεται ότι η παράμετρος “ΔΜΣ κάτω του 25” προκαλεί αύξηση της τάξης του 60% στην κατάσταση στόχο “sd 2 της

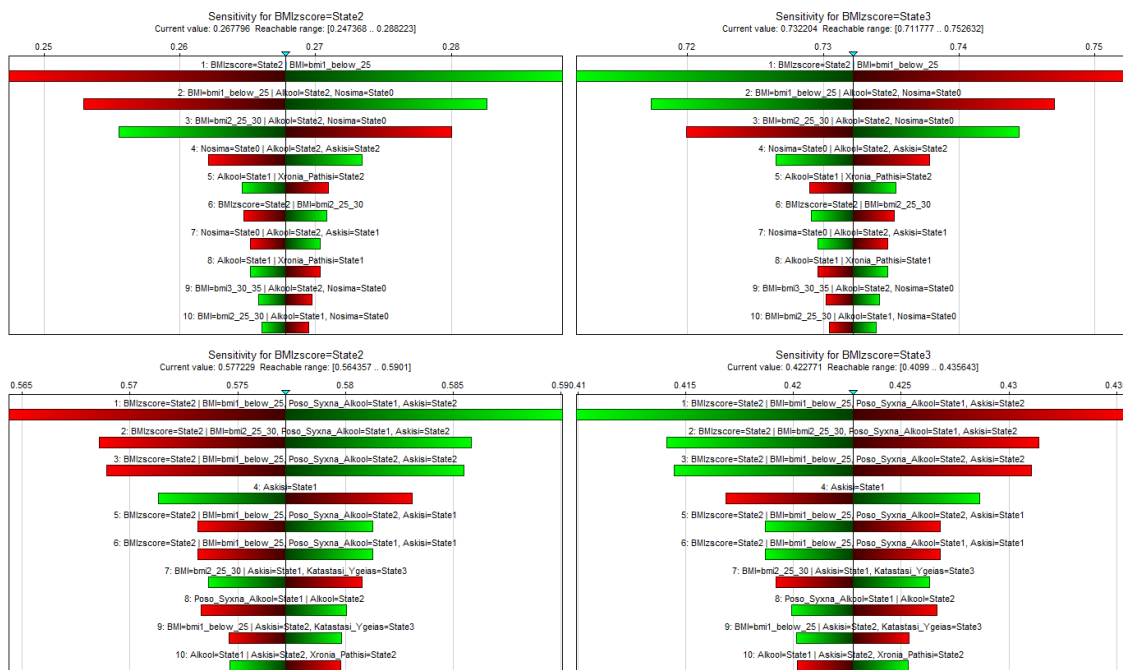
ΔΜΣ ζ-β” και μείωση της τάξης περίπου του 57%. Με το ίδιο σκεπτικό, στην Εικόνα 6.10, ο ΔΜΣ των κοριτσιών μεταξύ 25-30 δεδομένου του Είδος Νοσήματος στον βαθμό 0, των Φαρμάκων στον βαθμό 1 και του Καπνίσματος στην Εγκυμοσύνη στον βαθμό 2 προκαλεί αύξηση της τάξης του 70% και μείωση της τάξης του 68% στο sd 3 της ΔΜΣ ζ-β. Ο ΔΜΣ κάτω του 25 δεδομένου του ΔΜΣ του Πατέρα μεταξύ 30 και 35 προκαλεί αυξομείωση της τάξης του 60% (Εικόνα 6.12). Στην Εικόνα 6.13, ο ΔΜΣ των κοριτσιών κάτω του 25 δεδομένης της συχνότητας κατανάλωσης Αλκοόλ των Γονέων στον βαθμό 1 και της Άσκησης των Γονέων στο βαθμό 2 προκαλεί αυξομείωση της τάξης του 59% περίπου στο sd 3 της ΔΜΣ ζ-β. Τέλος ο ΔΜΣ των κοριτσιών κάτω του 25 προκαλεί αύξηση της τάξης του 63% στο sd 2 της ΔΜΣ ζ-β και μείωση της τάξης περίπου του 55% (Εικόνα 6.14).



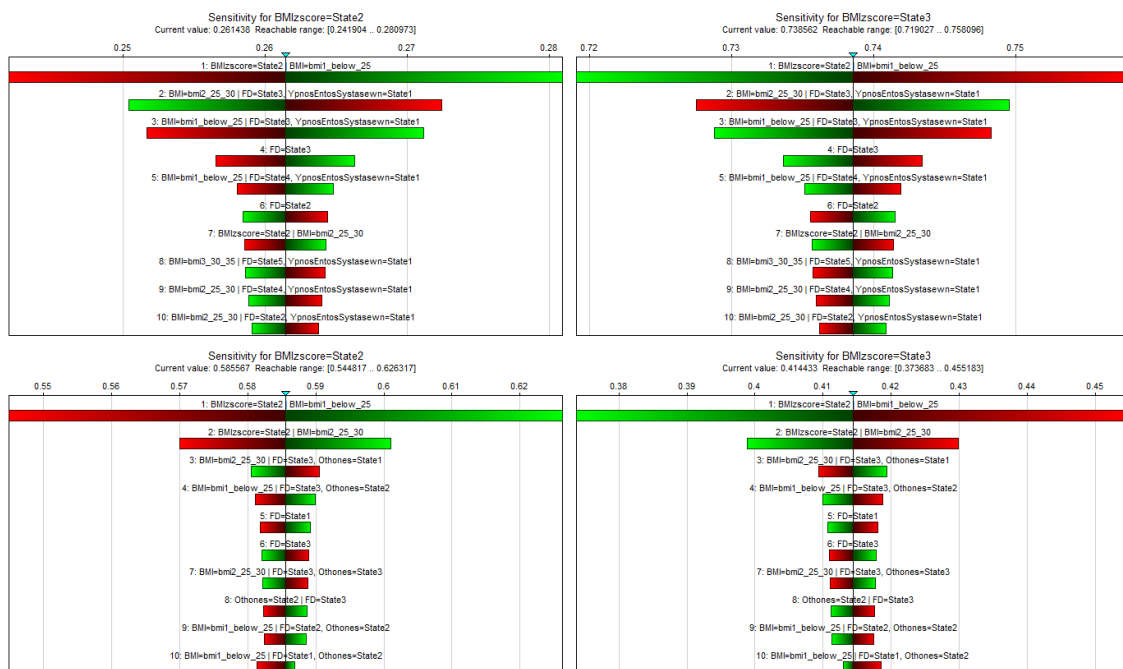
Εικόνα 6.12: Διαγράμματα Ευαισθησίας των Σωματομετρικών Χαρακτηριστικών των Γονέων στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)

Ο ΔΜΣ των αγοριών μεταξύ 25-30 δεδομένης Αυστηρής Μητέρας μεταξύ 3 και 4 προκαλεί αυξομείωση της τάξης του 80% στο sd 3 της ΔΜΣ ζ-β (Εικόνα 6.9). Ομοίως, ο ΔΜΣ των αγοριών μεταξύ 25-30 δεδομένης της Παρακολούθησης πάνω από 4 και της Ενασχόλησης πάνω από 4, προκαλεί αυξομείωση της τάξης του 75% στο sd 3 της ΔΜΣ ζ-β (Εικόνα 6.8), ενώ η ίδια παράμετρος δεδομένης της Οικογενειακής Κατάστασης στον βαθμό 2 και της Εκπαίδευσης τύπου 2 στον βαθμό 3 προκαλεί αυξομείωση της τάξης του 75% στο sd 3 της ΔΜΣ ζ-β (Εικόνα 6.10). Στην εικόνα 6.11 φαίνεται επίσης ότι ο ΔΜΣ των αγοριών μεταξύ 25-30 δεδομένου του Είδος Νοσήματος στον βαθμό 0, των Φαρμάκων στον βαθμό 1 και του Καπνίσματος στην Εγκυμοσύνη στον βαθμό 2 προκαλεί αύξηση της τάξης του 70% και μείωση της τάξης του 68% στο sd 3 της ΔΜΣ ζ-β. Τέλος στην Εικόνα 6.12, ο ΔΜΣ των αγοριών μεταξύ 25-30 δεδομένου του Ύψους της Μητέρας μεταξύ 165 και 170 εκ. καταλήγει στην μέγιστη παρατηρούμενη αύξηση του sd 3 της ΔΜΣ ζ-β με 80% και μείωση περίπου 77%.

Οι πολύ υψηλές τιμές της Ακρίβειας στους Πίνακες 6.3 και 6.4 επιβεβαιώνουν ότι πράγ-



Εικόνα 6.13: Διαγράμματα Ευαισθησίας του Τρόπου Ζωής των Γονέων στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)



Εικόνα 6.14: Διαγράμματα Ευαισθησίας του Τρόπου Ζωής των Παιδιών στις καταστάσεις της ΔΜΣ ζ-βαθμολογίας sd 2 (αριστερά) και sd 3 (δεξιά) στο σύνολο των Κοριτσιών (κάτω) και Αγοριών (πάνω)

ματι κατασκευάσαμε Μπείσιανά Δίκτυα για κάθε κατηγορία πολύ κοντά στο βέλτιστο δυνατό. Στο σύνολο των κοριτσιών το εύρος κυμαίνεται μεταξύ 75% και 90%, ενώ για τα αγόρια από 89% μέχρι 96% περίπου. Πρέπει να σημειωθεί ότι ως μεταβλητή κλάσης χρησιμοποιήθηκε

η ΔΜΣ ζ-βαθμολογία και η διασταυρωμένη επικύρωση ήταν δέκα πτυχών (10-fold cross validation). Με αυτόν το τρόπο η Ακρίβεια αφορά την σωστή κατηγοριοποίηση (classification) των μεταβλητών σε σχέση με την μεταβλητή κλάσης. Ομοίως, η συνάρτηση Log Likelihood κυμαίνεται σε ικανοποιητικά επίπεδα.

Κεφάλαιο **7**

Επίλογος

7.1 Συμπεράσματα

Ο κύριος στόχος της παρούσας διπλωματικής εργασίας είναι η διεξαγωγή μιας βαθύτερης ανάλυσης για να βρεθεί η σχέση και ο βαθμός συσχέτισης μεταξύ των μεταβλητών που εξάγονται από σωματομετρικά, ψυχομετρικά και κοινωνικο-οικονομικά δεδομένα γονέων και παχύσαρκων παιδιών. Στο πλαίσιο της μελέτης αυτής, αναπτύχθηκαν μοντέλα διερεύνησης των παραγόντων που επηρεάζουν την εμφάνιση και μεταβολή των επιπέδων παχυσαρκίας στα παιδιά, ώστε η επιστημονική κοινότητα, οι αρμόδιοι φορείς αλλά και οι ίδιοι οι γονείς και τα παιδιά τους να προθούν σε κατάλληλες παρεμβάσεις για την διαχείριση και εν τέλη καταπολέμηση της παιδικής παχυσαρκίας. Η αναζήτηση των παραγόντων επιτεύχθηκε μέσω της μεθόδου της Μπεϊσιανής Ανάλυσης Δικτύου, η οποία εμπεριέχει την κατασκευή του δικτύου και την εξαγωγή Συμπερασμού με εφαρμογή μαθηματικών αναλύσεων. Τα αποτελέσματα έδειξαν ότι τα μοντέλα έχουν ικανοποιητική ερμηνευτική ικανότητα και είναι σε θέση να εντοπίζουν ενδιαφέρουσες επιδράσεις, άλλες προφανείς κι άλλες όχι, όπως φαίνεται κι από τα αποτελέσματα του κεφαλαίου 6.

Οι κόμβοι του συστήματος ΚΑΓ, δηλαδή τα σωματομετρικά και ψυχομετρικά χαρακτηριστικά των γονέων και των παιδιών, αντιπροσωπεύουν κλινικούς προγνωστικούς παράγοντες, ενώ οι ακμές μεταξύ διαφορετικών κόμβων αντιπροσωπεύουν πιθανοτικές σχέσεις. Αυτή η προσέγγιση έχει πλεονεκτήματα έναντι των παραδοσιακών προσεγγίσεων καθώς προσδίδει μια εικόνα της πιθανολογικής σχέσης αντικατοπτρίζοντας την επιρροή των μεταβλητών μεταξύ τους. Σε αυτή την έρευνα, η σχέση αναπτύσσεται μεταξύ των χαρακτηριστικών που εξάγονται από υπέρβαρα και παχύσαρκα παιδιά όπως συζητείται στην εργασία που διεξήχθη στο Νοσοκομείο Παιδών "Η Αγία Σοφία" [77]. Οι προσεγγίσεις του Μπεϊσιανού δικτύου έχουν ένα ακόμα χαρακτηριστικό γνώρισμα το οποίο είναι εμπνευσμένο από τη βιολογική δομή των νευρικών ιστών, στην οποία οι νευρώνες επεξεργάζονται την συναπτική (synaptic) είσοδο και στη συνέχεια επικοινωνούν μεταξύ τους. Προσεγγίσεις Μπεϊσιανού δικτύου έχουν χρησιμοποιηθεί ευρέως σε διάφορους τομείς όπως η επεξεργασία εικόνας, η χειρογραφία και η αναγνώριση φωνής. Ωστόσο, αυτή η προσέγγιση έχει το πλεονέκτημα του να μοντελοποιεί ταυτόχρονα τις μεταβλητές για την κατανόηση της συσχέτισης μεταξύ πολλαπλών μεταβλητών και να επιλέγει μεταβλητής για περαιτέρω εξατομικευμένη ανάλυση. Η ισχύς της επιρροής μεταξύ οποιουδήποτε ζεύγους μεταβλητών μπορεί να υπολογιστεί χρησιμοποιώντας ανάλυση συσχέτισης. Η συνδιακύμανση εξαρτάται από τη μεταβλητότητα καθεμιάς από τις δύο με-

ταβλητές. Όλες αυτές οι μέθοδοι εξάγουν αριθμούς στην περιοχή $[0, +1]$ και εκφράζουν το πόσο στενά συνδέονται οι δύο μεταβλητές. Μια τέλεια συσχέτιση αντιπροσωπεύεται από το 1, ενώ το 0 υποδηλώνει ότι δεν υπάρχει συστηματική σχέση μεταξύ των δύο μεταβλητών. Μια μέση τιμή της συσχέτισης της τάξης $< 0,5$ θεωρείται σημαντική σχέση[23].

Για να συγκριθεί η δομή του παραγόμενου μπεισιανού δικτύου, το BayesFusion GeNIe χρησιμοποιεί το ελάχιστο μήκος περιγραφής του υποψηφίου δικτύου στον αλγόριθμο που βασίζεται στο *BD-eu* σκορ[116]. Η πολυπλοκότητα του δικτύου τροποποιείται από τον δομικό συντελεστή (structural coefficient, SC) μέσω της στάθμισης της δομής που κωδικοποιείται (σε bits) του μπεισιανού δικτύου.

Εκτελούμε μια ολοκληρωμένη ανάλυση δικτύου για να βρεθεί η σχέση μεταξύ των περιγραφικών παραγόντων στο επιλεγμένο δίκτυο και να μελετηθεί η επίδραση των μεμονωμένων περιγραφικών παραγόντων αυτών στην πρόβλεψη της παιδικής παχυσαρκίας. Για να εξεταστεί σε βάθος η σχέση μεταξύ των μεταβλητών, οι μέγιστες και μέσες τιμές της *J*-Απόκλισης, οι μέγιστες τιμές των Hellinger, ΣΑΚ και Ευκλείδειας, καθώς και οι υψηλότερες τιμές της ευαισθησίας δικτύου αναπαριστούν την ισχύ μεταξύ των μεταβλητών που παρατηρούνται συνολικά στο δίκτυο και ουσιαστικά υπολογίζουν την πιθανολογική εξάρτηση μεταξύ των κόμβων. Η από κοινού σχέση μεταξύ δύο μεταβλητών στο δίκτυο σε σύγκριση με μια υπόθεση ανεξαρτησίας προσδιορίζεται χρησιμοποιώντας την *J*-Απόκλιση για την μέτρηση του κέρδους της πληροφορίας. Η υψηλότερη τιμή της δείχνει ότι υπάρχει περισσότερη άμεση σχέση και μεγαλύτερη εξάρτηση του κόμβου γονέα με τον κόμβο παιδί.

Συμπερασματικά, οι ισχυρότερες συσχετίσεις βρέθηκαν μεταξύ ποικίλων ζευγών μεταβλητών. Στο σύνολο δεδομένων των κοριτσιών, η ζ-βαθμολογία του ΔΜΣ του πατέρα επιδρά στην αύξηση του ΔΜΣ (Avg: 0.722609). Η παρακολούθηση στο φαγητό (Monitoring) από την άλλη, επιδρά στην μείωση του βάρους (Avg: 0.584837). Ο τύπος εργασίας της μητέρας επιδρά στην ζ-βαθμολογία του ΔΜΣ (Avg: 0.57380) και τέλος στον ΔΜΣ επιδρά η ενασχόληση με οθόνες (Avg: 0.542261). Στο σύνολο δεδομένων των αγοριών, η αυστηρότητα των μητέρων (Authoritative) οδηγεί σε αυξημένο ΔΜΣ (Avg: 0.786711). Η παρακολούθηση στο φαγητό (Monitoring) επιδρά στον ΔΜΣ τους (Avg: 0.586838), μαζί με την οικογενειακή κατάσταση των αγοριών (Avg: 0.533525), τον τύπο εκπαίδευσης της μητέρας (Avg: 0.531103), την φυσική δραστηριότητα (Avg: 0.522708) και την κατανάλωση αλκοόλ του γονέα (Avg: 0.506545).

Είναι πολύ σημαντικό να υπογραμμισθούν τα χαρακτηριστικά που φαίνεται να μην έχουν κάποια επιρροή στο βάρος των αγοριών και κοριτσιών. Όσον αφορά τις γονεϊκές πρακτικές σίτισης, οι γονείς ως πρότυπα (Modeling), ο περιορισμός της τροφής (Restriction) και ο έλεγχος της διατροφής από το παιδί (Child Control) δεν φαίνεται να επιδρούν θετικά στο αυξημένο βάρος των παιδιών. Στην κατηγορία των γονεϊκών τύπων της μητέρας, μόνο η επιτρεπτική και η αυστηρή αντιμετώπιση στο φαγητό φαίνεται να επιδρά στο αυξημένο βάρος των παιδιών, κι όχι ο υποστηρικτικός ή αυταρχικός τύπος. Ο ύπνος των παιδιών δεν συσχετίστηκε με αυξημένο βάρος, ενώ στον αντίποδα, διάφορες συνήθειες των γονέων όπως η κατανάλωση αλκοόλ φαίνεται να είναι αρκετά πιο επιδραστικές στα παιδιά από όσο πιστεύουμε μέχρι σήμερα. Αξίζει να σημειωθεί ότι οι περιγεννητικοί παράγοντες όπως ο διαβήτης κύησης κλπ δεν φαίνεται να έχουν την προσδοκώμενη μεγάλη συσχέτιση στην αύξηση του βάρους των παιδιών. Αντίθετα, ο μητρικός θηλασμός επίσης δεν συσχετίστηκε με αυξημένο βάρος, γεγονός μάλλον θετικό στην επικράτηση του μητρικού θηλασμού στις πολύ νεαρές

ηλικίες. Τέλος, τα κοινωνικο-δημογραφικά χαρακτηριστικά των γονέων όπως η εθνικότητα, η οικογενειακή κατάσταση (έγγαμοι, άγαμοι κλπ), η εκπαίδευση των γονέων, η εργασία και το εισόδημα έχουν μέτρια προς χαμηλή επίδραση στο αυξημένο βάρος των παιδιών.

Σε αυτό το σημείο οφείλουμε να αναφέρουμε τους περιορισμούς της μελέτης μας. Στα δεδομένα συμπεριλήφθηκαν παιδιά υπέρβαρα και παχύσαρκα, των οποίων οι γονείς έτρεψαν ανησυχία για το βάρος τους και απευθύνθηκαν σε αντίστοιχα ιατρεία. Εν τέλη ο αριθμός τους έφτασε τα 48 (28 αγόρια και 20 κορίτσια), επομένως γίνεται εύκολα κατανοητό ότι αποτελούν στοιχεία ενός πολύ ειδικού μέρους του πληθυσμού, όπου αφενός τα παιδιά ενδεχομένως να είχαν φτάσει σε ακραία επίπεδα παχυσαρκίας, και αφετέρου οι γονείς είχαν το γνωσιακό υπόβαθρο και χαρακτήρα να διερευνήσουν περαιτέρω το ζήτημα, γεγονός που να προκαθορίζει προς πολλές κατευθύνσεις τα κοινωνικο-δημογραφικά χαρακτηριστικά, τους γονεϊκούς τύπους και τις γονεϊκές πρακτικές σίτισης. Επιπλέον, το μικρό δείγμα της πληθικότητας των δεδομένων, κάνει τα μοντέλα μας μη εφαρμόσιμα στον γενικό πληθυσμό. Στην κατάληξη αυτή ευθύνονται και πολλές πρακτικές δυσχέρειες, όπως η έλλειψη ομάδας ελέγχου από το ιατρείο, καθώς και η συγχρονική φύση της μελέτης από όπου προέρχονται τα δεδομένα. Ακόμα, πριν την συλλογή των δεδομένων δεν εφαρμόστηκε εκτενής γενετικός έλεγχος (για οικονομικούς λόγους), ο οποίος θα αποδεικνυόταν καθοριστικής σημασίας για την διερεύνησή μας.

Παρόλα αυτά, η μελέτη από την οποία προήλθαν τα δεδομένα αποτελεί την πρώτη συγχρονική μελέτη στην Ελλάδα που διερεύνησε την σχέση γονεϊκής τυπολογίας και γονεϊκών πρακτικών σίτισης με την παιδική παχυσαρκία[77] και κατ' επέκταση, η παρούσα μελέτη πρωτοπορεί σε επίπεδο μοντελοποίησης των συμπερασμάτων με την εφαρμογή αυστηρών μαθηματικών μεθόδων και δεν εμμένει σε μια στείρα και επιφανειακή στατιστική παρατήρηση των αποτελεσμάτων. Οι ερμηνεύσεις μας είναι πολύτιμες διότι μπορούν να αποτελέσουν την βάση για το σχεδιασμό ειδικών συμπεριφορικών παρεμβάσεων που να περιλαμβάνουν πέρα από το παιδί, ολόκληρη την οικογένεια για την αντιμετώπιση της παιδικής παχυσαρκίας. Γίνεται εύκολα αντιληπτό ότι επιδραστικές αλλαγές δεν απαιτούνται μόνο στον τρόπο ζωής των παιδιών, αλλά και στον τρόπο ζωής των γονέων και την γονεϊκή αντιμετώπιση του φαγητού των παιδιών σαν ολότητα, με κατάλληλες ενέργειες.

7.2 Μελλοντικές Επεκτάσεις

Χρησιμοποιώντας την προσέγγιση του Συμπερασμού Μπεϊσιανών Δικτύων, παρατηρήσαμε σημαντικές συσχετίσεις μεταξύ των χαρακτηριστικών που εξήχθησαν από την[77] και οι οποίες μπορούν περαιτέρω να χρησιμοποιηθούν για τη δημιουργία και διερεύνηση νέων υποθέσεων. Νέες μελέτες απαιτούνται για την διερεύνηση της σημασίας αυτών των νέων συσχετισμών. Μια πρώτη έγκριση των παραγόμενων σχέσεων υποστηρίζεται από την παρουσία διαφόρων τεκμηριωμένων και αυταπόδεικτων συνδέσεων στο πλαίσιο των συνολικών ερευνών πάνω στο συγκεκριμένο ζήτημα. Η παρούσα έρευνα θα δώσει μια εικόνα για την βαθύτερη ανάλυση της συσχέτισης και της δύναμης της σχέσης μεταξύ των χαρακτηριστικών που μπορεί να οδηγήσουν σε υψηλότερα ποσοστά ανίχνευσης της παιδικής παχυσαρκίας και υπερβαρότητας, και να βελτιώσει αποτελεσματικά την απόδοση της αξιολόγησης.

Εν κατακλείδι, το γεγονός ότι η παιδική παχυσαρκία, όπως αποδείχθηκε και από τις

στατιστικές μελέτες που καταγράψαμε στα κεφάλαια 2 και 3, αφορά ένα όλο και μεγαλύτερο μέρος του πληθυσμού, σε συνδυασμό με την αυξανόμενη εξοικείωση των ιατρών με την τεχνολογία και την εμπιστοσύνη στα μοντέλα, καθιστά επιτακτική την ανάγκη περαιτέρω εξέλιξης και βελτιστοποίησης των μοντέλων που σχετίζονται με την συσχέτιση παραγόντων και την ερμηνευσιμότητα των αποτελεσμάτων. Πιο συγκεκριμένα, είναι σημαντικό να εξετασθούν οι ερμηνείες των προβλέψεων του μοντέλου από επαγγελματίες υγείας προκειμένου να αξιολογηθεί ο τρόπος λήψης αποφάσεων. Πολύ σημαντικό είναι να χρησιμοποιηθεί ένα νέο σύνολο δεδομένων ως είσοδος στο μοντέλο ώστε να αξιολογηθεί η απόδοσή του και να διορθωθούν οι ελλείψεις του. Επιπλέον ένα πλουσιότερο σύνολο δεδομένων με περισσότερα δείγματα και μεγαλύτερη ποικιλία παρατηρήσεων θα αποτελούσε καθοριστικό παράγοντα γενίκευσης των αποτελεσμάτων σε μεγαλύτερο εύρος πληθυσμού. Από την άλλη οπτική, μια ολοκληρωμένη επαγωγική διαδικασία προϋποθέτει την εφαρμογή του συνόλου δεδομένων εισόδου και σε διαφορετικά μοντέλα ώστε να συγκριθεί η απόδοση και τα αποτελέσματα. Τέλος, απαιτείται να εφαρμοστούν και άλλες μέθοδοι ερμηνευσιμότητας πέραν της μπεϊσιανής ανάλυσης δικτύου προκειμένου να συγκριθούν τα αποτελέσματα, να αξιολογηθεί η απόδοση της τεχνικής και να πραγματοποιηθεί μία εμβάθυνση στην εσωτερική λειτουργία των μπεϊσιανών δικτύων.

Παράρτημα

Χρήσιμες Έννοιες

A'.1 P-Value

Στην στατιστική, η τιμή p είναι η πιθανότητα λήψης αποτελεσμάτων, τουλάχιστον τόσο ακραία όσο τα παρατηρούμενα αποτελέσματα ενός στατιστικού τεστ υποθέσεων (hypothesis test), υποθέτοντας ότι η μηδενική υπόθεση είναι σωστή (null hypothesis). Το hypothesis test είναι μια πράξη στην στατιστική σύμφωνα με την οποία ένας αναλυτής ελέγχει μια υπόθεση σχετικά με μια παράμετρο πληθυσμού. Η μεθοδολογία που χρησιμοποιεί ο αναλυτής εξαρτάται από τη φύση των δεδομένων που χρησιμοποιούνται και τον λόγο της ανάλυσης. Το hypothesis test χρησιμοποιείται για την εκτίμηση της λογικής μιας υπόθεσης χρησιμοποιώντας δείγματα δεδομένων. Τέτοια δεδομένα μπορεί να προέρχονται από μεγαλύτερο πληθυσμό ή από διαδικασία δημιουργίας δεδομένων.

Η τιμή p χρησιμοποιείται ως εναλλακτική λύση στα σημεία απόρριψης για να παρέχει το μικρότερο επίπεδο σημασίας στο οποίο θα απορρίπτεται η μηδενική υπόθεση. Μια μικρή τιμή της p σημαίνει ότι υπάρχουν ισχυρότερα στοιχεία υπέρ της εναλλακτικής υπόθεσης. Η τιμή p είναι ένα μέτρο της πιθανότητας ότι μια παρατηρούμενη διαφορά θα μπορούσε να είχε συμβεί τυχαία. Όσο χαμηλότερη είναι η τιμή p , τόσο μεγαλύτερη είναι η στατιστική σημασία της παρατηρούμενης διαφοράς. Η τιμή p μπορεί να χρησιμοποιηθεί ως εναλλακτική λύση ή επιπλέον των προεπιλεγμένων επιπέδων εμπιστοσύνης για hypothesis testing.

A'.2 Z-Score

Στην θεωρία των πιθανοτήτων και την στατιστική, ονομάζεται επίσης κανονική βαθμολογία (Standard score), τιμή z (z-value), κανονικοποιημένη βαθμολογία (Standardized score) ή κανονική βαθμολογία (Normal score). Η βαθμολογία z μετρά την απόσταση μεταξύ μιας παρατήρησης και του μέσου όρου, μετρούμενη σε μονάδες τυπικής απόκλισης (σ). Με άλλα λόγια, η βαθμολογία z είναι ο αριθμός των τυπικών αποκλίσεων που υπάρχουν μεταξύ μιας δεδομένης τιμής και του μέσου όρου του συνόλου δεδομένων. Εάν η βαθμολογία z είναι μηδέν, τότε η βαθμολογία σημείου των δεδομένων είναι ίδια με τη μέση τιμή. Εάν η βαθμολογία z είναι 1, τότε αντιπροσωπεύει μια τιμή που είναι μία τυπική απόκλιση μακριά από τον μέσο όρο. Η βαθμολογία z μπορεί να είναι θετική ή αρνητική. Μια θετική τιμή αντιπροσωπεύει τη βαθμολογία πάνω από τον μέσο όρο (δεξιά ουρά) και μια αρνητική βαθμολογία αντιπροσωπεύει τη βαθμολογία κάτω από τη μέση τιμή (αριστερή ουρά). Χρησιμοποιείται κυρίως

για να περιγράψει την κανονική κατανομή. Η τυπική κανονική κατανομή είναι η κανονική κατανομή με μέση τιμή 0 και τυπική απόκλιση 1 και συμβολίζεται ως $N(0, 1)$. Μια τυχαία μεταβλητή με την τυπική κανονική κατανομή ονομάζεται τυπική κανονική μεταβλητή. Οι συγκεκριμένες τιμές μιας τυπικής κανονικής μεταβλητής ονομάζονται τιμές z . Αν X είναι μια κανονικά κατανομημένη τυχαία μεταβλητή με μέση τιμή μ και τυπική απόκλιση σ , τότε η κατανομή των

$$Z = (X - \mu)/\sigma$$

είναι η τυπική κανονική κατανομή. Για μια συγκεκριμένη τιμή x του X , η $(x - \mu)/\sigma$ ονομάζεται τιμή z . Η τιμή z είναι μια συγκεκριμένη τιμή του Z .

Α'.2.1 BMI z-score

Η ζ-βαθμολογία του ΔΜΣ (Body Mass Index z-score, BMI s.d.) είναι μέτρο σχετικά με το βάρος προσαρμοσμένο για την ηλικία και το φύλο του παιδιού. Δεδομένων της ηλικίας, του φύλου, του ΔΜΣ του παιδιού και ενός κατάλληλου προτύπου αναφοράς, μπορεί να προσδιοριστεί μια ζ-βαθμολογία του ΔΜΣ (ή το ισοδύναμο εκατοστιαίο ποσοστό του ΔΜΣ ανά ηλικία). Πρέπει να σημειωθεί ότι οι ζ-βαθμολογίες υπολογίζονται σε σχέση με εξωτερική αναφορά (εθνική ή διεθνή) και όχι με κάποια εσωτερική αναφορά. Το αποτέλεσμα είναι ότι περισσότερο από το 5% του πληθυσμού θα μπορούσε να είναι, για παράδειγμα, πάνω από το 95ο εκατοστημόριο του ΔΜΣ ανά ηλικία. Ενώ, εάν χρησιμοποιούταν μια εσωτερική αναφορά, ακριβώς το 5% του κάθε δείγματος θα ήταν πάνω από το 95ο εκατοστημόριο και με αυτόν τον τρόπο, συγκεκριμένο οριακό σημείο του ΔΜΣ που ορίζει το 95ο εκατοστημόριο δεν θα ήταν το ίδιο σε όλα τα δείγματα. Για το λόγο αυτό προτιμήθηκε οι ζ-βαθμολογίες να βασίζονται σε εξωτερική αναφορά.

Βιβλιογραφία

- [1] E. R. Pulgarón. *Childhood obesity: a review of increased risk for physical and psychological comorbidities*. *Clin Ther*, 35(1):A18–32, 2013.
- [2] W. P. T James. *Overweight and obesity (high body mass index). Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors*. *World Health Organization Geneva*, 1(1), 2004.
- [3] B. A. Swinburn. *The global obesity pandemic: shaped by global drivers and local environments*. *Lancet*, 378(9793):804–14, 2011.
- [4] Y. C. Wang, K. McPherson, T. Marsh, S. L. Gortmaker και M. Brown. *Health and economic burden of the projected obesity trends in the usa and the uk*. *Lancet*, 378(9793):815–25, 2011.
- [5] J. C. Han, D. A. Lawlor και S.Y. Kimm. *Childhood obesity*. 2010.
- [6] S.A. Grover. *Years of life lost and healthy life-years lost from diabetes and cardiovascular disease in overweight and obese people: a modelling study*. *Lancet Diabetes Endocrinol*, 3(2):114–22, 2015.
- [7] R. Dobbs. *Overcoming obesity: an initial economic analysis*. *McKinsey Global Institute*, 2014.
- [8] E. A. Finkelstein, W. C. K. Graham και R. Malhotra. *Lifetime direct medical costs of childhood obesity*. *Pediatrics*, 133(5):854–62, 2014.
- [9] K. Feldman, G. M. B. Solymos, M. P.de Albuquerque και N. V. Chawla. *Unraveling Complexity about Childhood Obesity and Nutritional Interventions: Modeling Interactions Among Psychological Factors*. *Scientific Reports*, 9(18807), 2019.
- [10] L. G. Simons. *Diferences between european americans and african americans in the association between child obesity and disrupted parenting*. *Journal of comparative family studies*, 39(4):589–610, 2008.
- [11] M. Ng. *Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the global burden of disease study 2013*. *Lancet*, 384(9945):766–81, 2014.
- [12] World Health Organization. *Obesity: Preventing and Managing the Global Epidemic*. 2000.

- [13] M. Ng, T. Fleming, M. Robinson, B. Thomson, N. Graetz και C. Margono. *Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet*, 384(9945):746, 2014.
- [14] J. J. Reilly και J. Kelly. *Long-term impact of overweight and obesity in childhood adolescence on morbidity and premature mortality in adulthood: systematic review. Int J Obes (Lond)*, 35(7):891-8, 2011.
- [15] Y. C. Wang, K. McPherson και T. March. *Health and economic burden of the projected obesity trends in the USA and UK. The Lancet*, 378(9793):815-825, 2011.
- [16] S. J. Olshansky, D. J. Passaro και R. C. Hershov. *A potential decline in life expectancy in the United States in the 21st century. N Engl J Med*, 352(11):1138-45, 2005.
- [17] H. J. Tung και M. C. Yeh. *Parenting style and child-feeding behaviour in predicting children's weight status change in Taiwan. Public Health Nutrition*, 17(5):970-978, 2013.
- [18] S. C. Cunnane. *Childhood origins of lifestyle-related risk factors for coronary heart disease in adulthood. Proceedings Ninth IEEE Symposium on Computer-Based Medical Systems*, 1993.
- [19] L. B. Gerald, A. Anderson, G. D Johnson, C. Hoff και R. F. Trim. *Social class, social support and obesity risk in children. Child Care Health Dev*, 20(3):145-63, 1994.
- [20] B. Nod και M. Bigec. *Childhood Obesity Prediction with Artificial Neural Networks. Proceedings Ninth IEEE Symposium on Computer-Based Medical Systems*, 1(1):77-82, 1996.
- [21] Y. Lebingue, M. Migmac, E. Locard και N. Mamelle. *Birth weight and obesity at the age of 6. Study from the growth curves of a population of schoolchildren. Pediatrie*, 48(9):623-32, 1993.
- [22] K. Kancler, M. Bigec, J. Završnik, P Kokol, B. Novak και M. Pasek. *Birth weight and obesity at the age of 6. Study from the growth curves of a population of schoolchildren. Obesity of children in Maribor's region. Pediatrie*, 48(9):623-32, 1992.
- [23] L. Hussain, A. Ali, S. Rathore, S. Saeed, A. Idriss, M. U. Usman, M. A. Iftikhar και D. Y. Suh. *Applying Bayesian Network Approach to Determine the Association Between Morphological Features Extracted from Prostate Cancer Images. IEEE Access*, 7(1):1586-1601, 2019.
- [24] S. S. Tabrizi και N. Sancar. *Prediction of Body Mass Index: A comparative study of Multiple Linear Regression, ANN and ANFIS models. Procedia Computer Science*, 120(1):394-401, 2017.

- [25] H. A. Hiza, C. Pratt, A. L. Mardis και R. Anand. *Body mass index and health*. Harvard Health Publishing, 2001.
- [26] A. E. Field, E. H. Coakley, A. Must, J. L. Spadano, N. Laird και W. H. Dietz. *Impact of overweight on the risk of developing common chronic diseases during a 10-year period*. *Arch Intern Med*, 161(13):1581-6, 2001.
- [27] L. Williams. *Third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III) final report*. *Circulation*, 106(25):3143-421, 2002.
- [28] S. Grundy, H. Brewer, J. Cleeman, S. Smith και C. Lenfant. *Definition of metabolic syndrome*. *Circulation*, 109(1):433-438, 2004.
- [29] K. Yu, X. Liu, R. Alhamzawi, F. Becker και J. Lord. *Statistical methods for body mass index: A selective review*. *Statistical methods in medical research*. *Statistical Methods in Medical Research*, 27(3), 2016.
- [30] K. Gardner, J. Bird, P. Canning, L. Frizzell και L. Smith. *Prevalence of overweight, obesity and underweight among 5-year-old children in saint lucia by three methods of classification and a comparison with historical rates*. *Child Care Health Dev.*, 37(1):143-9, 2011.
- [31] J. Puder και S. Munsch. *Psychological correlates of childhood obesity*. *Int J Obes (Lond)*, 2(1):37-43, 2010.
- [32] A. Must και R. S. Strauss. *Risks and consequences of childhood and adolescent obesity*. *Int J Obes Relat Metab Disord*, 2(1):2-11, 1999.
- [33] D. M. Janicke. *Impact of psychosocial factors on quality of life in overweight youth*. *Obesity (Silver Spring)*, 15(7):1799-807, 2007.
- [34] D. L. Rofey, J. J. Black, J. E. Phillips, R. Blake και K. Olson. *Childhood obesity: depression, anxiety and recommended therapeutic strategies*. *BMC Medicine*, 18(30), 2011.
- [35] L. Y. Gibson. *An overview of psychosocial symptoms in obese children*. 2011.
- [36] C. Braet. *Patient characteristics as predictors of weight loss after an obesity treatment for children*. *Obesity (Silver Spring)*, 14(1):148-55, 2006.
- [37] E. A. Whitlock. *Effectiveness of weight management programs in children and adolescents*. *Evid Rep Technol Assess (Full Rep)*, 1(170):1-308, 2008.
- [38] D. M. Styne, S. A. Arslanian, E. L. Connor, I. S. Farooqi, M. H. Murad, J. H. Silverstein και J. A. Yanovski. *Pediatric Obesity-Assessment, Treatment, and Prevention: An Endocrine Society Clinical Practice Guideline*. *J Clin Endocrinol Metab.*, 102(3):709-757, 2017.

- [39] R. J. F. Loos και A. Janssens. *Predicting Polygenic Obesity Using Genetic Information. Cell Metab.*, 25(3):535–43, 2017.
- [40] A. E. Locke, B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers, F. R. Day και E. K. Speliotes. *Genetic studies of body mass index yield new insights for obesity biology. Nature*, 518(7538), 2015.
- [41] K. Silventoinen, A. Jelenkovic, R. Sund, Y. M. Hur, Y. Yokoyama, C. Honda και J. Kaprio. *Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COLlaborative project of Development of Anthropometrical measures in Twins (CODATwins) study. Am J Clin Nutr*, 104(2):371–9, 2016.
- [42] J. A. W. Baidal, L. M. Locks, E. R. Cheng, T. L. Blake-Lamb, M. E. Perkins και E. M. Taveras. *Risk Factors for Childhood Obesity in the First 1,000 Days: A Systematic Review. Am J Prev Med*, 50(6):761–779, 2016.
- [43] S. Rayfield και E. Plugge. *Systematic review and meta-analysis of the association between maternal smoking in pregnancy and childhood overweight and obesity. J Epidemiol Community Health*, 71(2):162–173, 2017.
- [44] S. P. Mourtakos, K. D. Tambalis, D. B. Panagiotakos, G. Antonogeorgos, G. Arnaoutis, K. Karteroliotis και L. S. Sidossis. *Maternal lifestyle characteristics during pregnancy, and the risk of obesity in the offspring: a study of 5,125 children. BMC Pregnancy Childbirth*, 21(15):66, 2015.
- [45] J. Yan, L. Liu, Y. Zhu, G. Huang και P. P. Wang. *The association between breastfeeding and childhood obesity: a meta-analysis. BMC Public Health*, 14(1267), 2014.
- [46] R. M. Martin, M. S. Kramer, R. Patel, S. L. Rifas-Shiman, J. Thompson, S. Yang και E. Oken. *Effects of Promoting Long-term, Exclusive Breastfeeding on Adolescent Adiposity, Blood Pressure, and Growth Trajectories: A Secondary Analysis of a Randomized Clinical Trial. JAMA Pediatr.*, 171(7), 2017.
- [47] World Health Organization. *Report of the Commission on Ending Childhood Obesity. Implementation plan: executive summary.* 2017.
- [48] A. G. LeBlanc, P. T. Katzmarzyk, T. V. Barreira, S. T. Broyles, J. P. Chaput, T. S. Church και M. S. Tremblay. *Correlates of Total Sedentary Time and Screen Time in 9-11 Year-Old Children around the World: The International Study of Childhood Obesity, Lifestyle and the Environment. PLoS One*, 10(6), 2015.
- [49] R. Elmesmari, A. Martin, J. J. Reilly και J. Y. Paton. *Comparison of accelerometer measured levels of physical activity and sedentary time between obese and non-obese children and adolescents: a systematic review. BMC Pediatr*, 18(1):106, 2018.

- [50] S. B. D. Torre, A. Keller, J. L. Depeyre και M. Kruseman. *Sugar-Sweetened Beverages and Obesity Risk in Children and Adolescents: A Systematic Analysis on How Methodological Quality May Influence Conclusions*. *J Acad Nutr Diet*, 116(4):638–59, 2016.
- [51] C. S. Costa, B. Del-Ponte, M. C. F. Assuncao και I. S. Santos. *Consumption of ultraprocessed foods and body fat during childhood and adolescence: a systematic review*. *Public Health Nutr*, 21(1):148–159, 2018.
- [52] K. D. Tambalis, D. B. Panagiotakos, G. Psarra και L. S. Sidossis. *Insufficient Sleep Duration Is Associated With Dietary Habits, Screen Time, and Obesity in Children*. *J Clin Sleep Med.*, 14(10):1689–1696, 2018.
- [53] T. J. Saunders, C. E. Gray, V. J. Poitras, J. P. Chaput, I. Janssen, P. T. Katzmarzyk και V. Carson. *Combinations of physical activity, sedentary behaviour and sleep: relationships with health indicators in school-aged children and youth*. *Appl Physiol Nutr Metab.*, 2016.
- [54] L. Li, S. Zhang, Y. Huang και K. Chen. *Sleep duration and obesity in children: A systematic review and meta-analysis of prospective cohort studies*. *J Paediatr Child Health*, 53(4):378–385, 2017.
- [55] J. Stang K. A. Loth. *Parenting style and child feeding practices: potential mitigating factors in the etiology of childhood obesity*. *J Am Diet Assoc*, 111(9):1301–5, 2011.
- [56] L. L. Birch και K. K. Davison. *Family environmental factors influencing the developing behavioral controls of food intake and childhood overweight*. *Pediatr Clin North Am.*, 48(4):893–907, 2001.
- [57] H. R. Clarke, E. Goyder και P. Bissell. *How do parents' child-feeding behaviours influence child weight? Implications for childhood obesity policy*. *J Public Health (Oxf)*, 29(2):132–41, 2007.
- [58] D. R. Musher-Eizenman και S. C. Holub. *Children's eating in the absence of hunger: the role of restrictive feeding practices*. 2006.
- [59] K. Rhee. *Childhood overweight and the relationship between parent behaviors, parenting style, and family functioning*. *The Annals of the American Academy of Political and Social Science*, 615(1):12–37, 2008.
- [60] M. S. Faith, R. I. Berkowitz και V. A. Stallings. *Parental feeding attitudes and styles and child BMI: prospective analysis of a gene-environment interaction*. *Pediatrics*, 114(4):429–36, 2004.
- [61] D. Baumrind. *Parental discipline patterns and social competence in children*. *Youth and Society*, 9(3):239, 1978.
- [62] D. Zlotnik και M. H. Bornstein. *Effects of parenting styles*. *Encyclopedia of Infant and Early Childhood Development*, 2007.

- [63] E. E. Maccoby και J. A. Martin. *Socialization in the context of the family: parent-child interaction*. *Handbook of Child Psychology: Socialization, Personality and Social Development*, 1983.
- [64] L. Chassin, C. C. Presson και J. Rose. *Parenting style and smoking-specific parenting practices as predictors of adolescent smoking onset*. *J Pediatr Psychol.*, 30(4):333-44, 2005.
- [65] L. Steinberg, J. D. Elmen και N. S. Mounts. *Authoritative parenting, psychological maturity, and academic success among adolescents*. *Child Dev*, 60(6):1424-36, 1989.
- [66] S. P. Kremers, J. Brug και H. D. Vries. *Parenting style and adolescent fruit consumption*. *Appetite*, 41(1):43-50, 2003.
- [67] K. van der Horst, S. P. Kremers και I. Ferreira. *Perceived parenting style and practices and the consumption of sugar-sweetened beverages by adolescents*. *Health Educ Res.*, 22(2):295-304, 2007.
- [68] J. L. Joyce και M. J. Zimmer-Gembeck. *Parent feeding restriction and child weight. The mediating role of child disinhibited eating and the moderating role of the parenting context*. *Health Education Research*, 22(2):295-304, 2009.
- [69] M. Shorer, R. David και M. Schoenberg-Taz. *Role of parenting style in achieving metabolic control in adolescents with type 1 diabetes*. *Diabetes Care*, 34(8):1735-7, 2011.
- [70] R. S. Enten και M. Golan. *Parenting style and weight-related symptoms and behaviors with recommendation for practice*. *Nutr Rev.*, 66(2):65-75, 2008.
- [71] K. Rhee, J. C. Lumeng και D. P. Appugliese. *Parenting style and overweight status in first grade*. *Pediatrics*, 117(6):2047-54, 2006.
- [72] J. M. Berge, M. Wall και K. Loth. *Parenting style as a predictor of adolescent weight and weight-related behaviors*. *J Adolesc Health*, 46(4):331-8, 2010.
- [73] A. K. Ventura και L. L. Birch. *Does parenting affect children's eating and weight status?* *Int J Behav Nutr Phys Act.*, 5(15), 2008.
- [74] E. F. C. Sleddens, S. M. P. L. Gerards και C. Thijs. *General parenting, childhood overweight and obesity-inducing behaviors: a review*. *Int J Pediatr Obes.*, 6(2):12-27, 2011.
- [75] N. Darling και L. Steinberg. *Parenting style as context: an integrated model*. *Psychological Bulletin*, 113(3):487-496, 1993.
- [76] Z. B. Yu, S. P. Han, G. Z. Zhu, C. Zhu, X. J. Wang, X. G. Cao και X. R. Guo. *Birth weight and subsequent risk of obesity: a systematic review and meta-analysis*. *Obes Rev.*, 12(7):525-42, 2011.

- [77] P. Chatzidaki. *Study of the correlation between parental style and parental feeding behavior to their children's weight outcomes*. 2019.
- [78] S. O. Hughes, T. G. Power, T. M. O'Connor, J. O. Fisher και T. A. Chen. *Maternal Feeding Styles and Food Parenting Practices as Predictors of Longitudinal Changes in Weight Status in Hispanic Preschoolers from Low-Income Families*. *J Obes.*, 2016.
- [79] S. O. Hughes, T. G. Power, J. O. Fisher, S. Mueller και T. A. Nicklas. *Revisiting a neglected construct: parenting styles in a child-feeding context*. *Appetite*, 44(1):83-92, 2005.
- [80] L. A. Frankel, T. M. O'Connor, T. A. Chen, T. Nicklas, T. G. Power και S. O. Hughes. *Parents' perceptions of preschool children's ability to regulate eating. Feeding style differences*. *Appetite*, 76(1):166-74, 2014.
- [81] S. O. Hughes, R. M. Shewchuk, M. L. Baskin, T. A. Nicklas και H. Qu. *Indulgent feeding style and children's weight status in preschool*. *J Dev Behav Pediatr.*, 29(5):403-10, 2008.
- [82] E. Hennessy, S. O. Hughes, J. P. Goldberg, R. R. Hyatt και C. D. Economos. *Parent behavior and child weight status among a diverse group of underserved rural families*. *Appetite*, 54(2):369-77, 2010.
- [83] A. Tovar, E. Hennessy και A. Pirie. *Feeding styles and child weight status among recent immigrant mother-child dyads*. *Int J Behav Nutr Phys Act.*, 9(1):62, 2012.
- [84] S. L. Hoerr, S. O. Hughes, J. O. Fisher, T. A. Nicklas, Y. Liu και R. M. Shewchuk. *Associations among parental feeding styles and children's food intake in families with limited incomes*. *The International Journal of Behavioral Nutrition and Physical Activity*, 6(1):55, 2009.
- [85] S. O. Hughes, T. G. Power και M. A. Papaioannou. *Emotional climate, feeding practices, and feeding styles: an observational analysis of the dinner meal in Head Start families*. *Int J Behav Nutr Phys Act.*, 8(1):60, 2011.
- [86] N. Shloim, L. R. Edelson, N. Martin και M. M. Hetherington. *Parenting styles, feeding styles, feeding practices, and weight status in 4-12 year-old children: a systematic review of the literature*. *A Systematic Review of the Literature*, 6(1):1849, 2015.
- [87] A. K. Ventura και L. L. Birch. *Does parenting affect children's eating and weight status?* *Int J Behav Nutr Phys Act.*, 5(1):15, 2008.
- [88] J. S. Gubbels, S. P. J. Kremers και A. Stafleu. *Association between parenting practices and children's dietary intake, activity behavior and development of body mass index: the KOALA Birth Cohort Study*. *Int J Behav Nutr Phys Act.*, 8(1):18, 2011.
- [89] L. Webber, L. Cooke, C. Hill και J. Wardle. *Child adiposity and maternal feeding practices: a longitudinal analysis*. *Am J Clin Nutr.*, 92(6):1423-8, 2010.

- [90] S. O. Hughes, T. G. Power, T. M. O'Connor και J. Orlet Fisher. *Executive functioning, emotion regulation, eating selfregulation, and weight status in low-income preschool children: how do they relate*. *Appetite*, 89(1):1-9, 2015.
- [91] K. Silva Garcia, T. G. Power, J. O. Fisher, T. M. O'Connor και S. O. Hughes. *Latina mothers' influences on child appetite regulation*. *Appetite*, 103(1):200-207, 2016.
- [92] M. A. Hernandez-Valero, A. V. Wilkinson και M. R. Forman. *Maternal BMI and country of birth as indicators of childhood obesity in children of Mexican origin*. *Obesity (Silver Spring)*, 15(10):2512-9, 2007.
- [93] J. Van Hook, E. Baker, C. E. Altman και M. L. Frisco. *Canaries in a coal mine: immigration and overweight among Mexican origin children in the US and Mexico*. *Soc Sci Med*, 74(2):125-34, 2012.
- [94] M. B. McCullough και A. K. Marks. *The immigrant paradox and adolescent obesity: examining health behaviors as potential mediators*. *J Dev Behav Pediatr.*, 35(2):138-43, 2014.
- [95] B. M. Popkin και J. R. Udry. *Adolescent obesity increases significantly in second and third generation U.S. immigrants: the National Longitudinal Study of Adolescent Health*. *J Nutr.*, 128(4):701-6, 1988.
- [96] T. G. Power, T. M. O'Connor, J. Orlet Fisher και S. O. Hughes. *Obesity risk in children: the role of acculturation in the feeding practices and styles of low-income hispanic families*. *Child Obes.*, 11(6):715-721, 2015.
- [97] A. Lake και T. Townshend. *Obesogenic environments: exploring the built and food environments*. *J R Soc Promot Health*, 126(6):262-7, 2006.
- [98] B. Wansink. *Environmental factors that increase the food intake and consumption volume of unknowing consumers*. *Annu Rev Nutr.*, 24(1):455-79, 2004.
- [99] J. R. Koiter. *Visualizing Inference in Bayesian Networks*. 2006.
- [100] R. T. Bayes. *An Essay Toward Solving a Problem in the Doctrine of Chances*. 1763.
- [101] S. J. Russell και P. Norvig. *Artificial Intelligence A Modern Approach*. 2003.
- [102] J. Pearl. *Fusion, propagation, and structuring in belief networks*. *Artificial Intelligence*, 29(3):241-288, 1986.
- [103] G. F. Cooper. *The computational complexity of probabilistic inference using Bayesian belief networks*. *Artificial Intelligence*, 42(2):393-405, 1990.
- [104] M. Henrion. *Propagation of uncertainty by Bayesian networks by probabilistic logic sampling*. *Machine Intelligence and Pattern Recognition*, 5(1):149-163, 1988.
- [105] R. Fung και K. C. Chang. *Weighting and integrating evidence for stochastic simulation in Bayesian networks*. *arXiv*, 1989.

- [106] R. D. Shachter και M. A. Peot. *Simulation approaches to general probabilistic inference on belief networks. Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence*, 1989.
- [107] R. Fung και B. D. Favero. *Backward simulation in Bayesian networks. Uncertainty Proceedings 1994*, 1994.
- [108] J. Cheng και M. J. Druzdzel. *An adaptive importance sampling algorithm for evidential reasoning in large Bayesian networks. Journal of Artificial Intelligence Research*, 13(1):155-188, 2000.
- [109] C. Yuan και M. J. Druzdzel. *An Importance Sampling Algorithm Based on Evidence Pre-propagation. arXiv*, 2003.
- [110] R. A. Howard και J. E. Matheson. *Influence diagrams. Readings on the Principles and Applications of Decision Analysis*. 1984.
- [111] C. Lacave και F. J. Diez. *A review of explanation methods for Bayesian networks. Cambridge University Press*, 2002.
- [112] J. L. Rodgers και W. A. Nicewander. *Thirteen ways to look at the correlation coefficient. The American Statistician*, 42(1):59-66, 1988.
- [113] E. Hellinger. *Die Orthogonalinvarianten quadratischer Formen von unendlich vielen Variablen*. 1907.
- [114] J. Burbea και C. R. Rao. *Entropy differential metric, distance and divergence measures in probability spaces: A unified approach. Journal of Multivariate Analysis*, 12(3):575-596, 1982.
- [115] P.C. Kraaijeveld. *An Interactive Generator of Diagnostic Bayesian Network Models*. 2005.
- [116] L.L.C. BayesFusion. *GeNIe Modeler User Manual*. 2021.
- [117] J. M. Tanner και R. H. Whitehouse. *Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. Arch Dis Child*, 51(13):170-9, 1976.
- [118] World Health Organization. *What is overweight and obesity?* 2018.
- [119] C. C. Robinson, B. Mandleco, S. F. και C. H. Olsen Hart. *The Parenting Styles and Dimensions Questionnaires (PSDQ)*. 2001.
- [120] A. Antonopoyloy και G. Tsitsas. *Investigation of the parental typology of the Greek mother: Weighting of the questionnaire: Parenting Styles and Dimensions Questionnaire (PSDQ)*. 2011.
- [121] D. Musher-Eizenman και S. Holub. *Comprehensive Feeding Practices Questionnaire: validation of a new measure of parental feeding practices. J Pediatr Psychol.*, 32(8):960-72, 2007.

- [122] M. Michou, D. Panagiotakos, E. Mamalaki, M. Yannakoulia και V. Costarelli. *Development and Validation of the Greek version of Comprehensive Parental Feeding Questionnaire. Mediterranean Journal of Nutrition and Metabolism. Mediterranean Journal of Nutrition and Metabolism*, 1(2):2, 2018.
- [123] G. F. Cooper και E. Herskovits. *A Bayesian method for the induction of probabilistic networks from data. Machine Learning*, 9(1):309-347, 1992.
- [124] D. Heckerman, D. Geiger και M. D. Chickering. *Learning Bayesian Networks: The Combination of Knowledge and Statistical Data. Machine Learning*, 20(1):197-243, 1995.
- [125] P. Spirtes, C. Glymour και R. Scheines. *Causation, Prediction, and Search*. 1993.
- [126] C. Glymour και M. Druzdzel. *Causal Inferences from Databases: Why Universities Lose Students*. 1999.
- [127] P. W. Cheng. *From covariation to causation: A causal power theory. Psychological Review*, 1(140):367-405, 1997.
- [128] N. Friedman, D. Geiger και M. Goldszmidt. *Bayesian Network Classifiers. Machine Learning*, 29(1):131-163, 1997.
- [129] R. Castillo. *Culture and mental illness: A client-centered approach*. 1997.
- [130] L. C. Vander Gaag και U. Kjaerulff. *Making Sensitivity Analysis Computationally Efficient*. 1999.
- [131] S. Mougiakakou, A. Prountzou, K. Zarkogianni και K. S. Nikita. *Prediction of glucose profile in children with type 1 diabetes mellitus using continuous glucose monitors and insulin pumps. Hormone Research*, 70:22-23, 2008.
- [132] I. Valavanis, S. G. Mougiakakou, K. A. Grimaldi και K. S. Nikita. *A multifactorial analysis of obesity as CVD risk factor: use of neural network based methods in a nutrigenetics context. BMC Bioinformatics*, 11:453, 2010.
- [133] K. Zarkogianni, A. Vazeou, S. G. Mougiakakou, A. Prountzou και K. S. Nikita. *An insulin infusion advisory system based on autotuning nonlinear model-predictive control. IEEE Transactions on Biomedical Engineering*, 58(9):2467-2477, 2011.
- [134] K. Zarkogianni, K. Mitsis, E. Litsa, M. Arredondo, G. Fico, A. Fioravanti και K. S. Nikita. *Comparative assessment of glucose prediction models for Patients with Type 1 Diabetes Mellitus applying sensors for glucose and physical activity monitoring. Journal of Medical and Biological Engineering and Computing*, 53(12):1333-1343, 2015.
- [135] K. Zarkogianni και K. S. Nikita. *Editorial: Special issue on emerging technologies for the management of diabetes mellitus. Journal of Medical and Biological Engineering and Computing*, 53(12):1255-1258, 2015.

- [136] K. Zarkogianni, E. Litsa, K. Mitsis, P. Wu, C. Kaddi, C. Cheng, M. Wang και K. S. Nikita. *A Review of Emerging Technologies for the Management of Diabetes Mellitus. IEEE Transactions on Biomedical Engineering*, 62(12):2735–2749, 2015.
- [137] L. Spanoy, K. V. Dalakleidi, K. Zarkogianni, P. Anastasia, K. S. Nikita, V. Vasiliki, M. Alevizaki και E. Anastasioy. *Ketonemia and ketonuria in gestational diabetes mellitus. Hormones*, 14(4):644–650, 2015.
- [138] K. Dalakleidi, K. Zarkogianni, A. Thanopoulou και K. S. Nikita. *Comparative assesment of statistical and machine learning techniques towards estimating the risk of developing type 2 diabetes and cardiovascular complications. Expert Systems*, 2017.
- [139] K. Zarkogianni, M. Athanasiou, A.C. Thanopoulou και K.S. Nikita. *Comparison of machine learning approaches towards assessing the risk of developing Cardiovascular disease as a long-term diabetes complication. IEEE Journal of Biomedical and Health Informatics*, 2018.
- [140] G. Fico, L. Hernandez, J. Cancela, A. Dagliati, L. Sacchi, A. Martinez-Millana, J. Posada, L. Manero, J. Verdú, A. Facchinetti, M. Ottaviano, K. Zarkogianni, K. S. Nikita, L. Groop, R. Gabriel-Sanchez, L. Chiovato, V. Traver, J. F. Merino-Torres, C. Cobelli, R. Bellazzi και M. T. Arredondo. *What do healthcare professionals need to turn risk models for type 2 diabetes into usable computerized clinical decision support systems? Lessons learned from the MOSAIC project. BMC medical informatics and decision making*, 19(1):1–16, 2019.
- [141] M. Athanasiou, K. Sfrintzeri, K. Zarkogianni, A. C. Thanopoulou και K.S. Nikita. *An explainable XGBoost-based approach towards assessing the risk of cardiovascular disease in patients with Type 2 Diabetes Mellitus. 2020.*
- [142] K. Zarkogianni, M. Athanasiou, K. Mitsis, E. Chatzidaki, N. Polychronaki, K. Perakis, D. Vergeti, D. Antonopoulou, E. Papachristou, V. Chioti, A. Voutetakis, E. Kalafatis, P. Pervanidou, C. Kanaka-Gantenbein και K. Nikita. *A comprehensive approach to empower self-management of health in childhood obesity based on gamification mechanisms and biofeedback. Diabetes technology and therapeutics*, 23:A57–A58, 2021.
- [143] K. Zarkogianni, M. Athanasiou, K. Mitsis, I. Vasilakis, I. Kosteria, K. Perakis, D. Vergeti, D. Antonopoulou, E. Papachristou, V. Chioti, I. Tokou, C. Kanaka-Gantenbein και K. S. Nikita. *The ENDORSE pilot trial. Diabetes technology and therapeutics*, 23:A68–A68, 2021.