

ΠΕΡΙΕΧΟΜΕΝΑ	
ΚΕΦΑΛΑΙΟ	
Η ΓΑΛΛΙΚΗ ΣΧΟΛΗ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ	
1.1 Βασικές Αρχές στις οποίες στηρίζεται η ερευνητική σκέψη της Σχολής	13
1.1.1 Στο Πυθαγόρειο θεώρημα	13
1.1.2 Στη ρήση του Επίκουρου	13
1.1.3 Στη ρήση του Σωκράτη	14
1.1.4 Στη ρήση του J.P BENZECRI	14
1.2 Θεωρητικό πλαίσιο στο οποίο στηρίζεται η ερευνητική σκέψη της Σχολής	14
1.2.1 Σχετικά με το Διακύβευμα	15
1.2.2 Σχετικά με τον n -διάστατο Ευκλείδειο διανυσματικό χώρο	15
1.3 Ανάλυση Δεδομένων, μία οικογένεια πολυδιάστατων μη παραμετρικών στατιστικών μεθόδων	17
1.3.1 Ιστορική αναδρομή	17
1.4 Πίνακες δεδομένων	20
1.5 Μια πρώτη προσέγγιση των εννοιών της παραγοντικής ανάλυσης	24
1.5.1 Παραγοντικοί άξονες. Παραγοντικό επίπεδο	28
1.5.2 Διαγωνοποίηση ενός τετραγωνικού πίνακα	33
1.5.3 Δυϊκός χώρος	37
1.5.4 Ανάλυση ενός πίνακα συμπτώσεων με την Παραγοντική Ανάλυση των Αντιστοιχιών	38
ΚΕΦΑΛΑΙΟ II	
ΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ ΤΩΝ ΑΝΤΙΣΤΟΙΧΙΩΝ	
2.0 Γενικά	41
2.1 Μάζα ενός στοιχείου του συνόλου I (ή του J)	43
2.2 Ο πίνακας των "προφίλ" των γραμμών	43
2.3 Ο πίνακας των "προφίλ" των στηλών	45
2.4 Ο δείκτης έλξης-άπωσης	46
2.5 Συμβολή των κελιών	46

2.6 Το νέφος $N(I)_J$ των προφίλ των γραμμών	49
2.6.1 Το simplex των προφίλ των γραμμών	49
2.6.2 Η μετρική του χ^2	50
2.7 Η αδράνεια του νέφους $N(I)_J$ των προφίλ των γραμμών	50
2.8 Παραγοντικοί άξονες στο επίπεδο	51
2.9 Η ανάλυση του νέφους $N(I)_J$ των προφίλ των γραμμών	53
2.10 Ποιότητα προβολής των προφίλ των γραμμών	56
2.11 Η ανάλυση του νέφους $N(J)_I$ των προφίλ των στηλών. Δυϊκή ανάλυση	57
2.12 Τα παραγοντικά επίπεδα	59
2.12.1 Βοήθειες για την ερμηνεία ενός παραγοντικού άξονα	60
2.12.2 Παγίδες με τις δυϊκές σχέσεις	61
2.12.3 Πλήθος ερμηνεύσιμων παραγοντικών αξόνων	62
2.12.4 Γενικές αρχές εφαρμογής της Π.Α.Α	69
2.13 Έλεγχος ανεξαρτησίας του πίνακα συμπτώσεων	69
2.14 Ο τύπος ανασύστασης των δεδομένων	72
2.15 Συμπληρωματικά στοιχεία	72
2.16 Το θεώρημα της ισοδυναμίας των σχετικών κατανομών	73
2.17 Το φαινόμενο Guttman	73
2.18 Η διαφορετική αντίληψη περί παραγοντικής ανάλυσης μεταξύ Γαλλικής και Αγγλοσαξωνικής Σχολής	75
2.19 Η έννοια της πληροφορίας	76
2.20 Αμοιβαία πληροφορία δύο χαρακτηριστικών	78
2.21 Δεσμός δύο χαρακτηριστικών	79
2.22 Ένα απλό αριθμητικό παράδειγμα εφαρμογής της Π.Α.Α	80
ΚΕΦΑΛΑΙΟ ΙΙΙ	
ΑΝΙΟΥΣΑ ΙΕΡΑΡΧΙΚΗ ΤΑΞΙΝΟΜΗΣΗ	
3.0 Γενικά	91
3.1 Απόσταση μεταξύ των στατιστικών μονάδων	94
3.1.1 Ιδιότητες μιας μετρικής	95

3.2 Κριτήρια συνένωσης δύο κλάσεων	96
3.3 Διαδικασία δημιουργίας της ιεραρχίας, Κριτήριο του Ward	98
3.3.1 Ο αλγόριθμος της μεθόδου	100
3.3.2 Ταξινόμηση με βάση το κριτήριο συνένωσης δύο κλάσεων του Ward	100
3.4 Κριτήριο διαμελισμού της ιεραρχίας σε r κλάσεις	106
3.5 Ερμηνεία του δένδρογράμματος της CAH	107
3.6 Η μέθοδος VACOR	108
3.7 Η μέθοδος FACOR	109
3.8 Ταξινόμηση βάσει του συντελεστή συσχέτισης	110
3.9 Ταξινόμηση με βάση τις κύριες συνιστώσες	110
3.10 Εντοπισμός των μεταβλητών που χαρακτηρίζουν ένα «αντικείμενο» της ταξινόμησης	110
ΚΕΦΑΛΑΙΟ IV ΜΕΘΟΔΟΛΟΓΙΑ ΤΗΣ ΔΙΕΡΕΥΝΗΣΗΣ ΔΕΔΟΜΕΝΩΝ ΜΕ ΒΑΣΗ ΤΗΝ ΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ	
4.0 Γενικά	121
4.1 Ανάλυση ενός πίνακα δεδομένων	121
4.2 Εφαρμογή της Παραγοντικής Ανάλυσης των Αντιστοιχιών	123
ΚΕΦΑΛΑΙΟ V Η ΜΕΘΟΔΟΣ KARAP ΩΣ ΕΡΓΑΛΕΙΟ ΕΞΟΡΥΞΗΣ ΔΕΔΟΜΕΝΩΝ	
5.0 Γενικά	131
5.1 Σχέση μεταξύ παραγόντων	133
5.2 Σύνδεση μεταξύ των «γραμμών» και των «στηλών» ενός διδιάστατου πίνακα δεδομένων	135
5.3 Πλησιέστερος συγγενής	143
5.4 Γενικά περί εκπαίδευσης δεδομένων με ταξινομητές μηχανικής μάθησης	144
5.4.1 Μηχανές Διανυσμάτων Υποστήριξης (ΜΔΥ). Support Vector Machines (SVM)	145
5.4.2 Εφαρμογή της μηχανής μάθησης SVM	146

5.4.3 Εκπαίδευση των δεδομένων με την Μηχανή Διανυσμάτων Υποστήριξης ΜΔΥ	146
5.5 Η νέα διαδικασία ταξινόμησης των γραμμών ενός πίνακα δεδομένων. Μέθοδος KARAP	147
5.6 Αποσαφήνιση του παραγοντικού επιπέδου 1x2. Συμβολή της μεθόδου KARAP	149
5.7 Συμπεράσματα	153
ΚΕΦΑΛΑΙΟ VI ΜΕΘΟΔΟΣ BENKAR ΑΞΙΟΛΟΓΗΣΗ ΤΗΣ ΑΝΙΟΥΣΑΣ ΙΕΡΑΡΧΙΚΗΣ ΤΑΞΙΝΟΜΗΣΗΣ	
6.1 Γενικά	155
6.2 Δημιουργία της Ανιούσας Ιεραρχικής Ταξινόμησης	156
6.3 Η μέθοδος BENKAR	157
6.4 Εφαρμογή της μεθόδου BENKAR	158
6.5 Εκπαίδευση των δεδομένων με την Μηχανή Διανυσμάτων	163
6.6 Συμπεράσματα	164
ΚΕΦΑΛΑΙΟ VII ΣΥΝΘΕΤΗ ΙΕΡΑΡΧΙΚΗ ΤΑΞΙΝΟΜΗΣΗ	
7.0 Γενικά	165
7.1 Τα βήματα της μεθόδου CHC	168
7.1.1 Εφαρμογή της μεθόδου CHC	169
7.2 CHC vs VACOR	172
7.2.1 Επισημάνσεις και διαφορές μεταξύ των δύο μεθόδων ταξινόμησης	175
7.3 Επιβεβαίωση της χρησιμότητας του όρου «πλησιέστερος συγγενής»	177
7.4 Σημαντικότητα vs Ομαδοποίηση	179
7.5 Παράδειγμα	180
7.6 Μέθοδος CHC vs Αγώνισμα του δεκάθλου	184
7.7 Συμπεράσματα	186
7.8 Αξιολόγηση του συστήματος CHC	186

ΚΕΦΑΛΑΙΟ VIII ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΚΛΑΣΕΩΝ ΤΗΣ ΑΝΙΟΥΣΑΣ ΙΕΡΑΡΧΙΚΗΣ ΤΑΞΙΝΟΜΗΣΗΣ	
8.0 Γενικά	189
8.1 Διαδικασία αξιολόγησης	189
ΚΕΦΑΛΑΙΟ IX ΑΝΑΛΥΣΗ ΣΕ ΚΥΡΙΕΣ ΣΥΝΙΣΤΩΣΕΣ	
9.0 Γενικά	201
9.1 Το κριτήριο του K. Pearson	201
9.2 Διαδικασία κατασκευής και ανάλυσης του νέφους των στατιστικών μονάδων $N(I)$	204
9.2.1 Η ανάλυση του νέφους $N(J)$ των μεταβλητών	207
9.3 Ο κύκλος συσχέτισης	208
9.4 Τα παραγοντικά επίπεδα	209
9.4.1 Ερμηνεία του παραγοντικού επιπέδου	210
9.5 Ανασύσταση των αρχικών δεδομένων από τις q κύριες συνιστώσες	211
9.6 Συμπληρωματικές στατιστικές μονάδες	212
9.7 Κανόνες διατηρητέων συνιστωσών	212
9.8 Διαστήματα εμπιστοσύνης του Anderson	220
9.9 Ο Παραγοντικός δείκτης απλότητας (IFS) του H. F. Kaiser	221
9.10 Ο συντελεστής Alpha του Lee J. Cronbach	222
9.11 Το Test Bartlett	222
ΚΕΦΑΛΑΙΟ X ΣΗΜΕΙΟΜΕΤΡΙΑ	
10.0 Εισαγωγή	225
10.1 Ιστορική αναδρομή	226
10.2 Προβληματισμοί για την εφαρμογή της μεθοδολογίας	227

10.3 Το κλασικό σημειομετρικό ερωτηματολόγιο	228
10.3.1 Τα σημειομετρικά επίπεδα	229
10.4 Παρουσίαση της προτεινόμενης τεχνικής με ένα παράδειγμα	230
10.5 Στατιστική επεξεργασία των δεδομένων	231
10.6 Το σημειομετρικό επίπεδο 2x3	234
10.7 Δημιουργία ψυχογραφήματος	237
10.7.1 Το σημειομετρικό επίπεδο	238
ΚΕΦΑΛΑΙΟ XI ΣΥΝΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ	
11.0 Γενικά	243
11.1 Συνεπαγωγή μεταξύ δυαδικών μεταβλητών	244
11.2 Δείκτης του I.C. Lemn	245
11.3 Η ένταση συνεπαγωγής μεταξύ δύο μεταβλητών	248
11.4 Ένταση συνεπαγωγής και υπό συνθήκη πιθανότητα	250
11.5 Συνεπαγωγή μεταξύ διαβαθμισμένων μεταβλητών	251
11.6 Διαβαθμισμένη μεταβλητή με n κλάσεις	252
11.7 Έννοια της συνοχής	253
11.8 Εντροπία. Συνεπαγωγική συνοχή	254
11.9 Υπολογισμός της εντροπίας του ζεύγους (a,b)	255
11.10 Συνοχή ζεύγους μεταβλητών	256
11.11 Υπολογισμός της συνοχής της τάξης ((a,b),c)	256
11.12 Συνοχή τάξης που αποτελείται από r στοιχεία	257
11.13 Συνεπαγωγή μεταξύ δύο τάξεων A και B	257
11.14 Ιεραρχική ταξινόμηση των τάξεων	258
11.15 Παράδειγμα εφαρμογής της συνεπαγωγικής στατιστικής	259
11.16 Στάδια δημιουργίας της Ιεραρχικής ταξινόμησης	261
ΚΕΦΑΛΑΙΟ XII ΔΙΑΚΡΙΤΙΚΗ ΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ	
12.0 Γενικά	265

12.1 Διαδικασία κατασκευής και ανάλυσης του νέφους $N(Q)$ των q ομάδων του ποιοτικού χαρακτηριστικού Y	266
12.2 Προσδιορισμός των διακριτικών παραγόντων	267
12.2.1 Διακριτική ικανότητα	268
12.3 Η μελέτη των στατιστικών μονάδων ως προς το νέφος $N(Q)$	268
12.4 Το διακριτικό παραγοντικό επίπεδο	270
12.5 Συμπληρωματική στατιστική μονάδα	270
12.6 Διακριτική ικανότητα κάθε μεταβλητής χωριστά	271
12.7 Παραδείγματα εφαρμογής της –AFD-	272
12.7.1 Τοποθέτηση υπαλλήλων βάσει της αποτελεσματικότητάς τους	272
12.7.2 Τμηματοποίηση της αγοράς	276
12.7.3 Διακριτική ικανότητα τεσσάρων παραγόντων που διαμορφώνουν την ποιότητα ενός κρασιού	279
12.7.4 Εξέταση της ομοιογένειας των απαντήσεων που έδωσαν τελειόφοιτοι μαθητές ως προς το φύλο	283
ΚΕΦΑΛΑΙΟ XIII ΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ ΔΗΜΟΣΚΟΠΗΣΗΣ	
13.1 Περιγραφή των αποτελεσμάτων	287
13.2 Σχετικά με εκλογικές αναμετρήσεις των κομμάτων	292
13.3 Σύνθεση ψηφοφόρων με αδιευκρίνιστη ψήφο Βάσει Φύλου και Ηλικίας	295
13.4 Η ανάλυση των αναποφάσιστων	295
ΚΕΦΑΛΑΙΟ XIV ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΕΩΝ	
14.1 Ανάλυση ενός πίνακα συμπτώσεων	297
14.2 Προσδιορισμός του πλήθους των παραγόντων που είναι σημαντικοί για την ερμηνεία ενός πίνακα δεδομένων	303
14.3 Ταξινόμηση των τίτλων cd-rom ως προς τη ζήτηση που παρουσιάζουν	307
14.4 Ταξινόμηση βάσει του συντελεστή συσχέτιση	314
14.5 Δημιουργία πίνακα τριπλής εισόδου από αρχείο σύμπτωσης $T(nxp)$	317
14.6 Τακτοποίηση νέων ερωτώμενων σε προκαθορισμένες ομάδες	322

14.7 Επεξεργασία τριδιάστατων πινάκων IxJxT	325
14.8 Ανάλυση μιας ακολουθίας k συνεχόμενων πινάκων συμπτώσεων	333
14.9 Η εφαρμογή της μεθόδου karap ως εργαλείο σχεδιασμού διαφημιστικής εκστρατείας ενός τουριστικού προορισμού	336
14.10 Βελτίωση της ερμηνευτικής ικανότητας ενός λογικού πίνακα 0-1 με την χρήση της μεθόδου karap	345
14.11 Εντοπισμός ακραίων απαντήσεων ενός ερωτηματολογίου	347
14.12 Δημιουργία διαφημιστικού μηνύματος με την μέθοδο KARAP	359
14.13 Κωδικοποίηση και ανάλυση μιας λογικής διαδικασίας	364
14.14 Η τεχνική βελτίωσης της πληροφορίας ενός πίνακα BURT	367
14.15 Η στάση των καταναλωτριών στην εμφάνιση ενός νέου προϊόντος	370
14.16 Συγκριτική αξιολόγηση k ομάδων βάσει p κριτηρίων με την χρήση ενός πίνακα αποστάσεων	374
14.17 Προβλήματα με τα οποία δεν ασχολούνται επαρκώς τα τρία μεγάλα κόμματα της Βουλής των Ελλήνων σύμφωνα με τις απόψεις των ψηφοφόρων τους	382
14.18 Σύγκριση πολιτικών κειμένων του ΠΑΣΟΚ, της ΝΔ και του ΚΚΕ πριν από τις εκλογές του 2004	392
14.19 Η ιδεολογική πλατφόρμα επικοινωνίας της ΝΔ του ΠΑΣΟΚ και του ΚΚΕ με τους ψηφοφόρους	395
14.20 Αξιολόγηση της δημοφιλίας ενός ηγέτη κόμματος από τους ψηφοφόρους	401
14.21 Αξιολόγηση προσωπικού μιας επιχείρησης	406
14.21.1 Ως προς την ανταπόκριση στην οργανωσιακή κουλτούρα της επιχείρησης	407
14.21.2 Η ανταπόκριση των 300 εργαζομένων στα 12 κριτήρια	409
14.21.2 Το προφίλ των 300 εργαζομένων που ανταποκρίνονται περισσότερο με συγκεκριμένο ερώτημα	410
14.22 Η διαχρονική εξέλιξη του δείκτη των λιανικών πωλήσεων κατά την χρονική περίοδο 1998-2002	411
14.23 Η Αξιολόγηση του μαθήματος της Ανάλυσης Δεδομένων	414
ΒΙΒΛΙΟΓΡΑΦΙΑ	424